

UNIVERSIDADE FEDERAL FLUMINENSE

JOSÉ MORISTA CARNEIRO DA SILVA

**BONE-SCAN: A COMPUTER-AIDED DETECTION  
(CADE) AND DIAGNOSIS (CADX) SYSTEM FOR BONE  
SCINTIGRAPHY.**

Niterói

2022

JOSÉ MORISTA CARNEIRO DA SILVA

**BONE-SCAN: A COMPUTER-AIDED DETECTION  
(CADE) AND DIAGNOSIS (CADX) SYSTEM FOR BONE  
SCINTIGRAPHY.**

Dissertation presented to the Computing Graduate Program of the Universidade Federal Fluminense in partial fulfillment of the requirements for the degree of Master of Science in Computer Science. Area: Computer Science.

Supervisor:  
AURA CONCI

Co-supervisor:  
CLÁUDIO TINOCO

Niterói  
2022

Ficha catalográfica automática - SDC/BEE  
Gerada com informações fornecidas pelo autor

S586b Silva, José Morista Carneiro da  
Bone-scan: a Computer-aided Detection (CAdE) and Diagnosis  
(CAdx) system for bone scintigraphy / José Morista Carneiro  
da Silva ; Aura Conci, orientadora ; Cláudio Tinoco Mesquita,  
coorientador. Niterói, 2022.  
122 f. : il.

Dissertação (mestrado)-Universidade Federal Fluminense,  
Niterói, 2022.

DOI: <http://dx.doi.org/10.22409/PGC.2022.m.10242638600>

1. Medical image processing. 2. Bone scintigraphy. 3. Bone  
metastasis. 4. Produção intelectual. I. Conci, Aura,  
orientadora. II. Mesquita, Cláudio Tinoco, coorientador. III.  
Universidade Federal Fluminense. Instituto de Computação.  
IV. Título.

CDD -

José Morista Carneiro da Silva

## **Bone-scan: a Computer-aided detection (CAdE) and diagnosis (CAdx) system for bone scintigraphy.**


Master's dissertation submitted to obtain the title of Master in Computer from Universidade Federal Fluminense - UFF

Work approved in February, 2022:



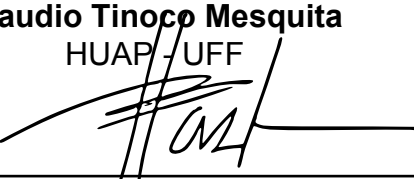
---

**Aura Conci**  
IC - UFF



---

**Claudio Tinoco Mesquita**  
HUAP - UFF



---

**Alexandre Plastino**  
IC - UFF



---

**Leandro Augusto Frata Fernandes**  
IC - UFF



---

**Luiz Eduardo S. Oliveira**  
UFPR

Niterói-RJ, Brazil



# Acknowledgements

The initial thanks are directed to my family: my mother, Tania de Farias Carneiro; my brother, Raul Carneiro Morista da Silva, my father, José Morista da Silva, my grandparents Laura Farias Carneiro, Gilberto Guimarães Carneiro, Neli Pereira da Silva and Valdete Morista da Silva for all the understanding, help and affection they provided me throughout my life and formation.

To my fiancée Mariane Ferreira Coelho, for all the love, incentive and help in life and at the process of doing this work.

To my friends, I thank for the companionship and support during the course.

To professor Aura, who has oriented and guided me not only in this work, but in many others, always with extreme competence and dedication.

To the nuclear medicine partners participating in the project, especially Dr. Cláudio, for all his availability and help provided.

Finally, I would like to thank all the professors and researchers who have been part of my academic career, for teaching me not only the contents of the disciplines, but also the value and importance of knowledge.

*"Machines take me by surprise with great frequency."  
(Alan Turing)*

# Resumo

Esta pesquisa apresenta o uso de técnicas de visão computacional e aprendizado de máquina para implementar um sistema automatizado capaz de auxiliar profissionais especializados na compreensão de exames de cintilografia óssea no formato DICOM. Para atingir os objetivos desta dissertação, algumas ideias originais são aplicadas; a primeira está relacionada ao uso de cada pixel da imagem proveniente dos exames como objeto de classificação. Com essa metodologia, mesmo com um pequeno conjunto de dados de exames é possível obter uma grande quantidade de entradas para serem utilizadas como treinamento e teste, extraíndo assim o máximo das informações disponíveis no pequeno conjunto de dados inicial. Outra ideia original apresenta o uso de operações que normalmente são aplicadas como filtros de pré-processamento (como Mediana, Gaussiana e outros) para serem usadas como features dos pixels presentes no conjunto de dados. A base de dados de cintilografias ósseas utilizada neste trabalho foi composta por 42 imagens de cintilografias ósseas advindas de 21 pacientes reais. Ao fim das tarefas de processamento de imagem e extração de features, foi obtido um conjunto de dados de aprendizagem de 2513386 registros, 64 features e 1 atributo de classe, este representando um pixel metastático ou não. A fim de reduzir a cardinalidade do vetor de features para cada pixel, a técnica Principal Component Analysis foi empregada levando a um novo conjunto de dados com 25 componentes principais. Os classificadores utilizados nesta dissertação foram Naive Bayes, Random Forest, Support Vector Machines e Multilayer perceptrons. Apesar do pequeno número de exames de cintilografia óssea utilizadas neste trabalho, os resultados de generalização obtidos foram promissores, apresentando como melhores resultados 77% de revocação e 84% de precisão para o classificador Random Forest e 78% de revocação e 84% de precisão no algoritmo Multilayer perceptrons.

**Palavras-chave:** Processamento de imagens médicas, Cintilografia óssea, Metástases ósseas.

# Abstract

This work presents the use of image analysis and machine learning techniques to implement a system to aid specialized professionals on analyze bone scans exams in DICOM format. In order to achieve this, some original ideas are applied. The first one is related to the use of each pixel of an exam as object of interest to classification. Another original idea is the use of operations that are normally applied in pre-processing as features for the dataset. With both, even with a small dataset was possible to obtain enough amount of entries to be used for training and testing.

Each feature vector used is initially composed of 64 features and 1 target attribute, this last one representing if the pixel is a metastasis or not. The used bone scans set was composed of 42 bone scan images from 21 patients. At the end of the image processing and feature extraction tasks a learning dataset of 2,512,386 records is computed. In order to reduce the cardinality of the vector of features, the Principal Component Analysis technique was employed leading to a new feature set with 25 components per object to be classified. The Naive Bayes, Random Forest, Support Vector Machines and Multilayer perceptrons classifiers were tested.

Despite the small number of bone scans available for this work, the Random Forest classifier achieves the results of 77% , 84% and 0.96 for Precision, Recall and Receiver Operation Characteristic respectively.

**Key-words:** Medical image processing, Bone scintigraphy, Bone metastasis.

# List of Figures

Figure 1 – A X-ray digital imaging machine and a X-ray exam indicating bone metastases [95]. . . . .	25
Figure 2 – A BS machine and a whole-body BS from a patient with metastatic bone disease. . . . .	26
Figure 3 – A CT scanner and an exam of this type that indicates the presence of bone metastases [95]. . . . .	27
Figure 4 – MRI machine and a MRI exam with suspicion of bone metastases disease [95]. . . . .	28
Figure 5 – Machine used to perform PET scans and an exam of this type with bone metastases diagnosis [96]. . . . .	29
Figure 6 – Ventral, dorsal and cranium bone scan images. . . . .	30
Figure 7 – Gaussian surface with same standard deviation in both directions. . . . .	42
Figure 8 – Window size and the median filter computation. . . . .	45
Figure 9 – Elements of the Gabor filters. . . . .	47
Figure 10 – A template to be found in a image. . . . .	48
Figure 11 – SVM hyperplanes for binary classification. . . . .	57
Figure 12 – Perceptron model. . . . .	58
Figure 13 – Relation among the used elements. . . . .	60
Figure 14 – Diagram of the application functionalities. . . . .	61
Figure 15 – View of the login system screen. . . . .	62
Figure 16 – View of the account creation system screen. . . . .	63
Figure 17 – View of the system screen of patients management. . . . .	64
Figure 18 – System screen related to exams visualization and management. . . . .	64
Figure 19 – Manual annotation of a bone scan image. . . . .	68
Figure 20 – Original image and ground truth provided by physicians. . . . .	69
Figure 21 – Transforming 2d array into 1d one. . . . .	70
Figure 22 – Original, simple and after median filter thresholding images . . . . .	73
Figure 23 – Template matching operation on processed images. . . . .	74
Figure 24 – Histogram of Figure 6 ventral scintigraphic presented from 0 to 255. . . . .	75
Figure 25 – Original x Equalization images and their histograms. . . . .	76
Figure 26 – Original x CLAHE images and their histograms. . . . .	77
Figure 27 – Original, Otsu result and ground truth images of the same exam. . . . .	78
Figure 28 – Original and Gaussian filtering with standard deviation 3, 5 and 7. . . . .	79
Figure 29 – Original and Median filtering with windows of 3x3 and 5x5. . . . .	80
Figure 30 – Original image and it after Entropy filtering. . . . .	82

Figure 31 – Original and after Robert filter images. . . . .	83
Figure 32 – Original and after Sobel filter images. . . . .	84
Figure 33 – Original and after Variance filter images. . . . .	85
Figure 34 – Original and Gabor filter result with angles of 0 and 90 degrees. . .	87
Figure 35 – Classes distribution of the generated dataset. . . . .	89
Figure 36 – PCA analysis plot. . . . .	90
Figure 37 – Processing results with dark background. . . . .	94
Figure 38 – Processing results with white background. . . . .	95
Figure 39 – Original, automatic classification output and ground truth. . . . .	96
Figure 40 – Example of an exported report. . . . .	97
Figure 41 – Evaluation of the completeness of the results obtained by the application.	102
Figure 42 – Likert scale evaluating the classification of true positives. . . . .	103
Figure 43 – Likert scale evaluating the classification of false positives. . . . .	103
Figure 44 – Likert scale evaluating the classification of false negatives. . . . .	103
Figure 45 – Original, results and ground truth of cranium. . . . .	108
Figure 46 – Threshold operation of first version. . . . .	119
Figure 47 – Detected regions and metastasis areas in $mm^2$ . . . . .	121

# List of Tables

Table 1 – Results for Lund group using CNN and 2,164 patients [8] . . . . .	35
Table 2 – Best confusion matrix of Larisa group (970 exams) [10]. . . . .	36
Table 3 – Cross-Validation training results . . . . .	92
Table 4 – Comparing related works . . . . .	99
Table 5 – Results of questions about patients management . . . . .	101
Table 6 – Results of questions about exams management . . . . .	101
Table 7 – Results of questions about exams processing . . . . .	102
Table 8 – Results for training set. . . . .	122
Table 9 – Results for test set. . . . .	122

# List of abbreviations and acronyms

AHE	Adaptive Histogram Equalization
AI	Artificial intelligence
ANN	Artificial Neural Network
API	Application Programming Interface
ASM	Active shape models
AUC	Area under the curve
BS	Bone scan
BSI	Bone scan Index
CADe	Computer-aided Detection
CADx	Computer-aided Diagnosis
CDF	Cumulative Distribution Function
CLAHE	Contrast Limited Adaptive Histogram Equalization
CNN	Convolution Neural Network
CSS	Cascading Style Sheets
CSV	Comma Separated Values
CT	Computerized Tomography
DICOM	Digital Imaging and Communications in Medicine
DR	Dimensionality Reduction
DT	Decision Tree
EANM	European Association of Nuclear Medicine
FN	False Negative
FP	False Positive
HTML	Hyper Text Markup Language

HTTPS	Secure Hyper Text Transfer Protocol
HUAP	Hospital Universitário Antônio Pedro
IA	Image Analysis
IC	Computation Institute
IP	Image Processing
JSON	JavaScript Object Notation
JWT	JSON Web Token
KCV	K-fold Cross Validation
KNN	K-Nearest Neighbors
LDH	Lactate dehydrogenase
LSAC	Level set active contour
MIT	Massachusetts Institute of Technology
MLP	Multi Layer Perceptrons
MRI	Magnetic Resonance Imaging
NB	Naive Bayes
NIST	National Institute of Standards and Technology
PCA	Principal Component Analysis
PET	Positron Emission Tomography
PGC	Program for Graduation in Computer at IC-UFF
PX	Pixel
ReLU	Rectified Linear Unit
REST	Representational State Transfer
RF	Random Forest
RGB	Red-Green-Blue
ROC	Receiver Operation Characteristic Curve
SASS	Syntactically Awesome Style Sheets

SEAC	Standards Eastern Automatic Computer
SQL	Structured Query Language
SL	Supervised learning
SVM	Support Vector Machines
TN	True Negative
TP	True Positive
TS	Typescript
UFF	Universidade Federal Fluminense
URL	Uniform resource locator
UL	Unsupervised learning
US	Ultrasound
WWW	World Wide Web

# List of symbols

$\sigma$	Standard deviation
$\sigma^2$	Variance of a distribution
$x$	Spatial plane coordinate on horizontal axis
$y$	Spatial plane coordinate on vertical axis
$f$	Two-dimensional function f
$H$	Histogram of a monochromatic image
$G$	Number of gray level intensities of a image.
$N$	Horizontal image resolution
$M$	Vertical image resolution
$H_c$	Cumulative histogram
$H_{eq}$	Global equalized histogram
$\alpha$	Clip factor to CLAHE
$\beta$	Clip factor to CLAHE
$\mu$	Mean of a distribution
$e$	Euler-Mascheroni Constant
$\pi$	PI constant
$\sigma_\omega^2$	Within class variance
$\omega_a$	Probability of a class a after separated by a threshold
$P$	Probability function
$Grad$	Gradient of a function
$\lambda$	Wavelength of the harmonic oscillator
$\theta$	Orientation of the normal to the parallel stripes of a Gabor function
$\gamma$	Aspect ratio or ellipticity of the Gabor function support

$\mu$	Greek letter mu
$\theta$	Greek letter theta
$\psi$	Phase angle
$R$	Subregion of an image
$\varphi_{s_i s_j}$	Cross-correlation of functions $f_1$ and $f_2$
$V$	Data vector
$v$	Arbitrary value
$D$	Dimension of a data vector
$d$	Arbitrary dimension of a data vector
$N_r$	Number of rows of a dataset
$n$	Arbitrary Number of elements of a dataset
$Z$	Centered and standarized matrix
$E$	Matrix of eigen vectors
$D_g$	Diagonal matrix
$\lambda_n$	nth Eingenvalue
$TN$	True negative
$FP$	False positive
$TP$	True positive
$FN$	False negative
$C$	Class/label
$s_j$	Input signal from a input j
$w_{ij}$	Weight of a input j from a neuron i
$U$	Output unit from a DICOM conversion
$S_v$	Stored value of a DICOM file
$m$	Rescale slope
$b$	Intercept point with ordinate axis

$L$	Constant for range conversion
$\infty$	Infinity representation
$el$	Element of a feature
$el_{norm}$	Normalized element of a feature
$el_{min}$	Minimum element of a feature
$el_{max}$	Maximum element of a feature
$\gamma_p$	Adjusted parameter of Kernel trick function

# Contents

<b>1</b>	<b>INTRODUCTION</b>	<b>20</b>
<b>1.1</b>	<b>Objective of this work</b>	<b>20</b>
<b>1.2</b>	<b>Work organization</b>	<b>21</b>
<b>1.3</b>	<b>About the medical partnership to this work</b>	<b>22</b>
<b>1.4</b>	<b>Importance of the work</b>	<b>22</b>
<b>2</b>	<b>BONE METASTASIS AND BONE SCINTIGRAPHY</b>	<b>23</b>
<b>2.1</b>	<b>Exams for early diagnosis</b>	<b>24</b>
2.1.1	Blood analysis	24
2.1.2	X-ray	24
2.1.3	Bone scan (BS) or Scintigraphy	25
2.1.4	Computerized tomography (CT)	26
2.1.5	Magnetic resonance imaging (MRI)	27
2.1.6	Positron emission tomography (PET)	28
<b>2.2</b>	<b>Bone scan procedure details</b>	<b>29</b>
<b>2.3</b>	<b>Medical files: DICOM standard</b>	<b>31</b>
<b>3</b>	<b>RELATED WORKS</b>	<b>32</b>
<b>3.1</b>	<b>Works up the first decade of this century</b>	<b>32</b>
<b>3.2</b>	<b>Works from 2011 to now</b>	<b>33</b>
<b>4</b>	<b>IMAGE ANALYSIS</b>	<b>37</b>
<b>4.1</b>	<b>Image analysis steps</b>	<b>37</b>
4.1.1	Image acquisition	38
4.1.2	Image enhancement	38
4.1.3	Feature extraction	38
4.1.4	Segmentation	38
<b>4.2</b>	<b>Image operations</b>	<b>39</b>
4.2.1	Histogram equalization	39
4.2.2	Contrast limited adaptive histogram equalization (CLAHE)	40
4.2.3	Variance and variance filter	41
4.2.4	Gaussian functions and Gaussian filter	41
4.2.5	Otsu thresholding	43
4.2.6	Median filter	44
4.2.7	Entropy filter	45
4.2.8	Roberts filter	45

4.2.9	Sobel filter . . . . .	46
4.2.10	Gabor filters . . . . .	46
4.2.11	Template Matching . . . . .	47
<b>4.3</b>	<b>Feature vector . . . . .</b>	<b>49</b>
<b>5</b>	<b>MACHINE LEARNING . . . . .</b>	<b>50</b>
<b>5.1</b>	<b>Supervised or unsupervised learning . . . . .</b>	<b>50</b>
<b>5.2</b>	<b>Ground truth and gold standard . . . . .</b>	<b>50</b>
<b>5.3</b>	<b>Dimensional reduction . . . . .</b>	<b>51</b>
<b>5.4</b>	<b>Principal component analysis . . . . .</b>	<b>52</b>
<b>5.5</b>	<b>Splitting data for training and testing . . . . .</b>	<b>53</b>
<b>5.6</b>	<b>Performance evaluation . . . . .</b>	<b>53</b>
<b>5.7</b>	<b>Learning techniques . . . . .</b>	<b>54</b>
5.7.1	Naive Bayes . . . . .	55
5.7.2	Random forest . . . . .	55
5.7.3	Support vector machines . . . . .	56
5.7.4	Artificial neural network . . . . .	57
<b>6</b>	<b>IMPLEMENTED SYSTEM . . . . .</b>	<b>59</b>
<b>6.1</b>	<b>An overview of the technologies used . . . . .</b>	<b>59</b>
<b>6.2</b>	<b>Operations available to physicians . . . . .</b>	<b>60</b>
<b>6.3</b>	<b>Identification and authentication of users . . . . .</b>	<b>61</b>
<b>6.4</b>	<b>System security . . . . .</b>	<b>63</b>
<b>6.5</b>	<b>Patient management . . . . .</b>	<b>63</b>
<b>6.6</b>	<b>DICOM meta data . . . . .</b>	<b>65</b>
<b>6.7</b>	<b>Used dataset description . . . . .</b>	<b>67</b>
<b>6.8</b>	<b>Ground Truth . . . . .</b>	<b>68</b>
<b>6.9</b>	<b>Implemented approach assumptions . . . . .</b>	<b>69</b>
<b>6.10</b>	<b>Feature vector composition . . . . .</b>	<b>71</b>
6.10.1	Intensity of absorption . . . . .	71
6.10.2	Pixel position . . . . .	71
6.10.3	Tag identifying pixels in pelvic region . . . . .	71
6.10.4	Histogram equalization features . . . . .	74
6.10.5	CLAHE feature . . . . .	76
6.10.6	Otsu feature . . . . .	78
6.10.7	Gaussian features . . . . .	79
6.10.8	Median features . . . . .	80
6.10.9	Entropy feature . . . . .	81
6.10.10	Roberts feature . . . . .	82
6.10.11	Sobel feature . . . . .	83

6.10.12	Feature of Variance . . . . .	84
6.10.13	Gabor features . . . . .	85
<b>6.11</b>	<b>Dataset processing and number of records . . . . .</b>	<b>87</b>
6.11.1	Data normalization . . . . .	89
6.11.2	Dataset reduction . . . . .	90
<b>6.12</b>	<b>Training results . . . . .</b>	<b>90</b>
6.12.1	Discussion about decisions made . . . . .	92
<b>6.13</b>	<b>Computation of the metastatic area . . . . .</b>	<b>93</b>
<b>6.14</b>	<b>Implementation output . . . . .</b>	<b>93</b>
<b>7</b>	<b>RESULTS AND CONCLUSIONS . . . . .</b>	<b>98</b>
<b>7.1</b>	<b>Comparison with related works and previous versions . . . . .</b>	<b>98</b>
<b>7.2</b>	<b>User Survey . . . . .</b>	<b>100</b>
7.2.1	Evaluation of patient management of functionalities . . . . .	100
7.2.2	Evaluation of exam management functionalities . . . . .	101
7.2.3	Evaluation of exam processing tasks and learning . . . . .	101
<b>7.3</b>	<b>Conclusion . . . . .</b>	<b>104</b>
7.3.1	Contributions of this dissertation . . . . .	104
7.3.2	About the originality of the work . . . . .	104
7.3.2.1	Using innate pre-process technique as features . . . . .	105
7.3.2.2	Maximizing the samples to be analyzed for each exam . . . . .	105
7.3.3	Final considerations . . . . .	106
<b>7.4</b>	<b>Future works . . . . .</b>	<b>106</b>
	<b>References . . . . .</b>	<b>109</b>
	<b>APPENDIX . . . . .</b>	<b>117</b>
	<b>APPENDIX A – PREVIOUS DEVELOPMENTS . . . . .</b>	<b>118</b>

# 1 Introduction

Computed-aided Detection (CADE) and Diagnosis (CADx) systems have been important areas of research in recent decades. However, these areas have been worked on separately. CADE systems do not present the final characteristics of diseases and CADx systems do not detect its locations. Moreover, both do not present good levels of automation. As a result, these systems are not yet widely used in clinical settings [1] [2].

The purpose of this work is to develop a new system for detection and diagnosis of metastases from gamma radiation acquired by external scanners. This gamma scanner produces scintigraphy exams. The goal is to develop a single system for metastatic identification and characterization in a way that the level of bone affected area could be comparable over the time.

Metastases are defined as the activities associated with the occurrence of new secondary cancers in patients outside the region of the primary cancer [10]. Bone is a frequent site of metastases and usually indicates a short-term prognosis for cancer patients. Once the cancer has spread to the bones, it can rarely be cured, however, it can often still be treated with the aim of slowing their growth [12].

Identifying metastatic bones requires a systematic inspection of scintigraphy exams by trained and experienced specialists. However, in the majority of the occasions, this process is performed only through a pure visual analysis, making impossible carry out a quantitative assessment of the issue extension.

The aforementioned way of examination also makes impractical to compare the progression of the patient when the time goes by, because evaluations based only on visual inspection may not be faithfully reproducible at a future examination done by same person and even less coincident with the evaluation of another professional.

The inability to quantify the findings imposed by this purely visual analysis directly impacts the routine of nuclear physicians and specialists, promoting greater difficulty in monitoring the evolution of the disease in patients and consequently, in conducting their respective treatments.

## 1.1 Objective of this work

This work aims to verify the viability of Image Analysis (IA) techniques and Machine Learning (ML) to design an application for aid in the analysis and quantification of bone metastasis areas on scintigraphic exams. In other words, to develop a union of

CADe and CADx ideas in one implementation for the detection and diagnosis of bone scans.

In order to achieve the goal, the developed solution must act as a facilitator at the various stages related to the diagnosis process. It can support the users in the initial steps on the organization of the acquired data (patients' exams and records); it must go up to tasks such as automatic segmentation of regions of possible metastatic activity and, finally write a medical report about the number of metastatic points found in a exam.

## 1.2 Work organization

The next chapter presents important concepts about what bone metastatic diseases are and their implications, as well as the current techniques for diagnosing this disease, highlighting bone scintigraphy as one of the main exams for this purpose.

Chapter 3 presents and discusses relevant works in the field of computational solutions to diagnose bone metastases; pointing out the gaps and possibilities that still exist and the relevance of the work developed in this dissertation.

Chapter 4 introduces important concepts about digital images processing and analysis, describing in details some important techniques from this field and how they contribute to this work.

Chapter 5 considers machine learning concepts and techniques. This chapter describes some fundamentals of computational learning and points out the methods and algorithms used in this dissertation to pre-processing, classification and evaluation of the generalization capacity of the implemented models.

Chapter 6 shows in its initial sections an overview of the system architecture, technologies and components used to develop the solution. Next, it describes the desired functionalities for the proposed solution and how physicians can register and add patients in the application, ensuring their information security and privacy. Still in this chapter, the methodology applied to the implemented CADe and CADx system is presented; the generation of the features vector and the various IA techniques used for its composition; the pre-processing operations applied on the learning dataset and the process of training and testing the classifiers used in this dissertation.

Chapter 7 presents and discusses the results obtained from the implemented application, debating the adequacy of all the aspects developed, from the information automatically collected from the images, the methodology used to classify and quantify the points of the image and up to the opinion of users about it. This chapter also discusses conclusions and desirable and ongoing improvements of the implementation to show potential future works on the same or related areas.

There is also one appendix showing former developments.

### 1.3 About the medical partnership to this work

The work developed in this dissertation had the guidance of resident physicians from the Hospital Universitário Antônio Pedro (HUAP) and students from the postgraduate course in nuclear medicine at the Universidade Federal Fluminense (UFF). These colleagues advised the development in the stages related to the understanding and manipulation of medical data, providing, with their expertise, subsidies in the construction of the system.

This research was also approved by the Ethics Committee of the Brazilian Federal Government under the number CAAE: 45645121.7.0000.5243, title: "Optimization of a MACHINE LEARNING method for classification of scintigraphic images", in Portuguese: "Otimização usando método de MACHINE LEARNING para classificação de imagens cintilográficas" on 01/06/2021.

### 1.4 Importance of the work

Although there are other previous works related to aiding on the detection and diagnosis of various diseases in the visual computing area of the Graduation Course on Computing (PGC), at the Institute of Computing (IC) of UFF, this is the first with nuclear images at the IC, introducing a new research line in PGC .

Moreover, providing a way to better detect and quantify bone metastases can significantly improve the conduct of treatments, which can be decisive in cases of possible treatment for cancer spread and vital for maintenance of patients life in most cases.

The bone scan are a very low-contrast type of imaging. Despite the high expertise a nuclear medicine professional must have, there are inherent limitations to the human eye when faced with images from gamma radiation to perform a quantitative count of the findings in the exam. These highly complex task can be more easily accomplished with the help of a computer. For example, once correctly classified in the exam all areas of metastases, a computer can quickly count the pixels present in them.

Once the presence of secondary tumors in bone regions can be quantified, it is possible to efficiently measure the advance or reduction of the affected area, enabling better monitoring of the response to treatments performed by the patient. Consequently, providing this computational aid can be a way to bring great benefits to the medical community and patients in the confrontation to the effects of this serious cancer condition.

## 2 Bone metastasis and bone scintigraphy

When cancer cells break away from a tumor, they can move through the bloodstream or lymph vessels to other parts of the body. Then, these cells can settle in an organ far from its original position and start a new tumor. The original tumor that cells break away from is called primary and the new tumor is called the secondary tumor. When this happens, it is called metastatic cancer.

Nearly all types of cancer have the ability to metastasize, but whether they present this behavior depends on a variety of individual factors. The most common sites for cancers to metastasize are the brain, lungs, liver and bones. Metastases can occur in three ways: They can grow directly into the tissue surrounding the tumor; Cells from a cancer can travel through the bloodstream to distant locations; or these cells can travel through the body lymphatic system to nearby or distant lymph nodes.

Secondary tumors in the bone region are called bone metastases [12]. For instance, consider a patient with prostate cancer. Prostate cancer cells from the primary tumor can break away and get into the bloodstream. Once in the blood, the cancer cells can travel to the bone and form a new (secondary) tumor. This is same-wise known as bone metastasis [16]. It is important to remember that this secondary tumor is made up of abnormal prostate cancer cells, not abnormal bone cells. The result of this process is referred to as prostate cancer that has metastasized to the bone or metastatic prostate cancer. Prostate, breast, and lung cancers are most likely to spread to the bone. However, there are other primary cancers with bone metastasis [16].

When cancer cells metastasize to the bone, they can cause changes to the bone [16] and are classified as osteolytic, osteoblastic or mixed, according to the primary mechanism of interference with normal bone remodeling [11]. Osteolytic is characterized by destruction of normal bone [11]; osteoblastic (or sclerotic) is characterized by deposition of new bone; and mixed occurs if a patient has both osteolytic and osteoblastic lesions, or if an individual metastasis has both osteolytic and osteoblastic components [11].

These changes can cause many symptoms to patients. Bone pain is often the first symptom of cancer that has spread to the bone. Fractures also became more common, breaks might happen with a fall or injury, but a weak bone can also break during normal activities [14]. Bone metastases can also cause a release of calcium into the bloodstream, this problem is called hypercalcemia and if untreated, this high calcium levels may cause a coma [14].

The median survival of patients after first bone metastasis by breast cancer is 20 months. This is in marked contrast to those with metastases of cancer in the liver, 3

months, or with bone metastases from non-small cell lung cancer, 6 months [11].

In prostate cancer, men with a good performance status and only bone metastasis (that is, with one affecting predominantly the axial skeleton), have a median duration with disease control (after androgen blockade) of 4 years or a median survival of 53 months. Oppositely, those with visceral disease and poor performance status have 12 months of average survival [11].

## 2.1 Exams for early diagnosis

There are some cases where physicians may find bone metastasis before the patient has symptoms. In some cancers, where bone metastasis is common, exams might be done to make sure the cancer has not spread to bones before a treatment plan starts. Some of the exams that can be used to allow this diagnosis are: blood analysis; X-ray; computerized tomography; magnetic resonance; bone scan and positron emission tomography.

### 2.1.1 Blood analysis

Cancer in the bones can cause high levels of chemicals in the bloodstream, these being identifiable in specific blood analysis. Although blood tests are not best suited to diagnose bone cancer, they may be helpful once a diagnosis is made. For example, high levels of chemicals in the blood such as alkaline phosphatase and lactate dehydrogenase (LDH) can suggest that the cancer may be more advanced in the patient [87].

Exams like blood cell counts and blood chemistry tests are also commonly done before and throughout the conduction of some bone metastasis treatments (such as chemotherapy) in order to monitor the patient's overall health [87].

### 2.1.2 X-ray

This is the most used imaging type for bone examination in general. Often a simple X-ray can show advanced cases of bone metastasis [15]. Classic X-ray (radiography) performed in 2 orthogonal projections is most often the first test performed in the early diagnosis of metastases. On the basis of X-ray, one can not only obtain information on the size and location of the bone tumor, but also assess the risk of pathological fracture. X-ray allows to determine whether the tumor is osteolytic, osteoblastic, mixed, or if there is a change due to tumor like change.

Figure 1 shows a X-Ray imaging machine and an example of an exam of this type with indicatives of bone metastasis. The X-ray examination is quick and cheap,

giving it a crucial advantage in accessibility and popularity. In each case the metastatic diagnosis strive to clarify the nature of the change further imaging methods like bone scintigraphy are used [86].



Figure 1 – A X-ray digital imaging machine and a X-ray exam indicating bone metastases [95].

### 2.1.3 Bone scan (BS) or Scintigraphy

A bone scan is a nuclear medicine exam in which a radioactive fluid that is attracted to bone cells affected by cancer is injected into the patient's vein. After a few hours, a scanner takes a series of images (Figure 2) of the patient's body and any areas of cancer in the bone will appear as highlighted areas [15]. This exam is a method of imaging with the highest sensitivity (95%) regarding bone metastases [86]. Figure 2 shows a BS machine and a bone scintigraphy from a patient suffering from bone metastases.

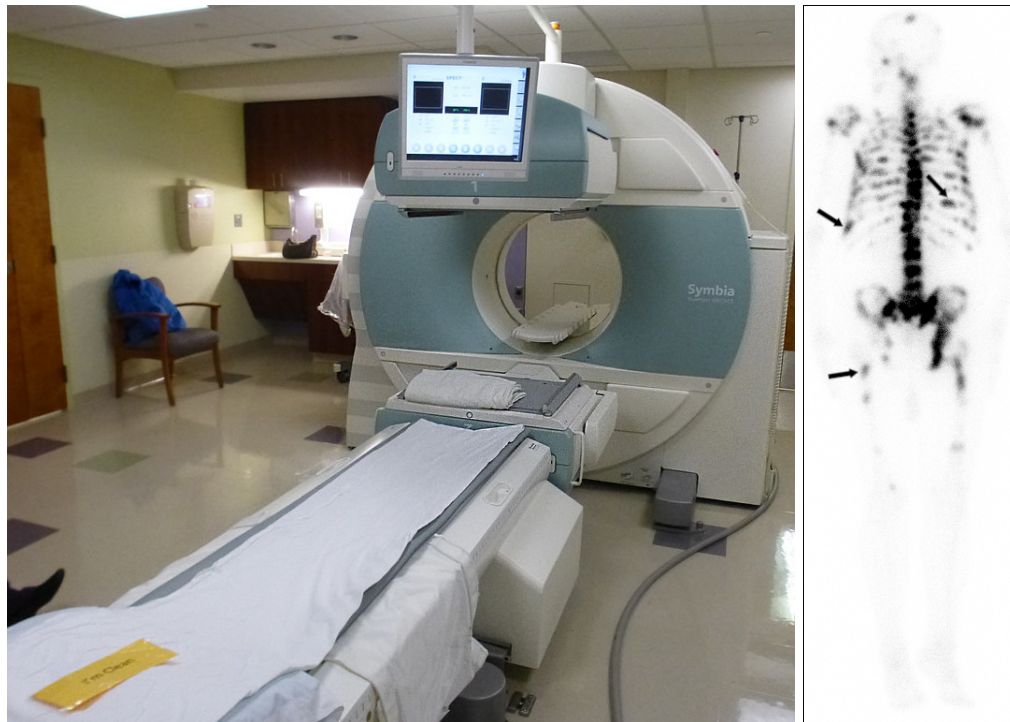


Figure 2 – A BS machine and a whole-body BS from a patient with metastatic bone disease.

#### 2.1.4 Computerized tomography (CT)

The term “computerized tomography”, or CT, refers to a digital image procedure in which a narrow beam of x-rays is sent to a patient and quickly rotated around the body, producing signals that are processed by the machine’s computer to generate cross-sectional images, or “slices”, of the body. Once a number of successive slices are collected by the machine’s computer, they can be digitally “stacked” together to form a 3-dimensional image of the patient [88]. A CT machine and an example of a CT exam with indicatives of bone metastasis is shown at Figure 3.

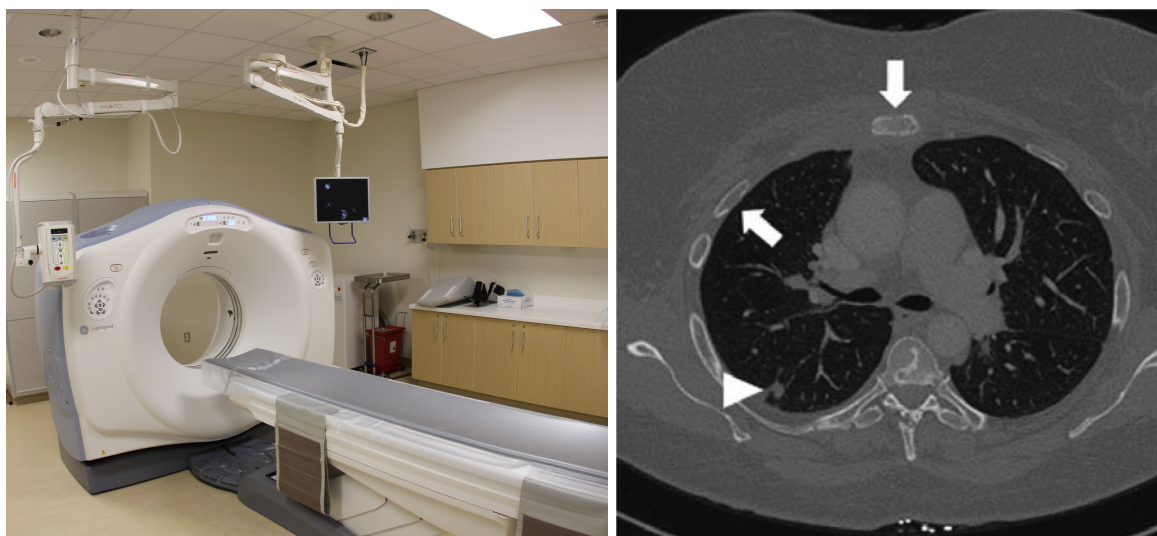


Figure 3 – A CT scanner and an exam of this type that indicates the presence of bone metastases [95].

This imaging procedure gives more detailed internal pictures of a body than a conventional X-ray by taking lots of acquisitions from different angles. It can be used to show images with more resolutions and details of a specific bone or area [15]. It is mainly used as a complementary imaging technique to diagnose bone metastases, being particularly recommended in dealing with discrepancies between the outcomes of X-ray and scintigraphy [86].

### 2.1.5 Magnetic resonance imaging (MRI)

Magnetic resonance imaging (MRI) is a medical imaging technique that uses a magnetic field and computer-generated radio waves to create detailed images of the organs and tissues. MRIs employ powerful magnets which produce a strong magnetic field that forces protons in the body to align with that field. When a radiofrequency current is then pulsed through the patient, the protons are stimulated, and spin out of equilibrium, straining against the pull of the magnetic field. When the radiofrequency field is turned off, the MRI sensors are able to detect the energy released as the protons realign with the magnetic field [89].

MRI is currently one of the most sophisticated methods for imaging bone tumors. Diagnostic images determine not only the degree and type of the damage to the bone but also give the opportunity to assess the soft tissues adjacent to the tumour. During whole body MRI examination there is greater chance of diagnosing metastases than with the use of scintigraphy. Nevertheless, this method is far less available due to equipment limitations [86]. A machine used to perform this exam and a MRI exam indicating metastatic bone cancer is shown at Figure 4.

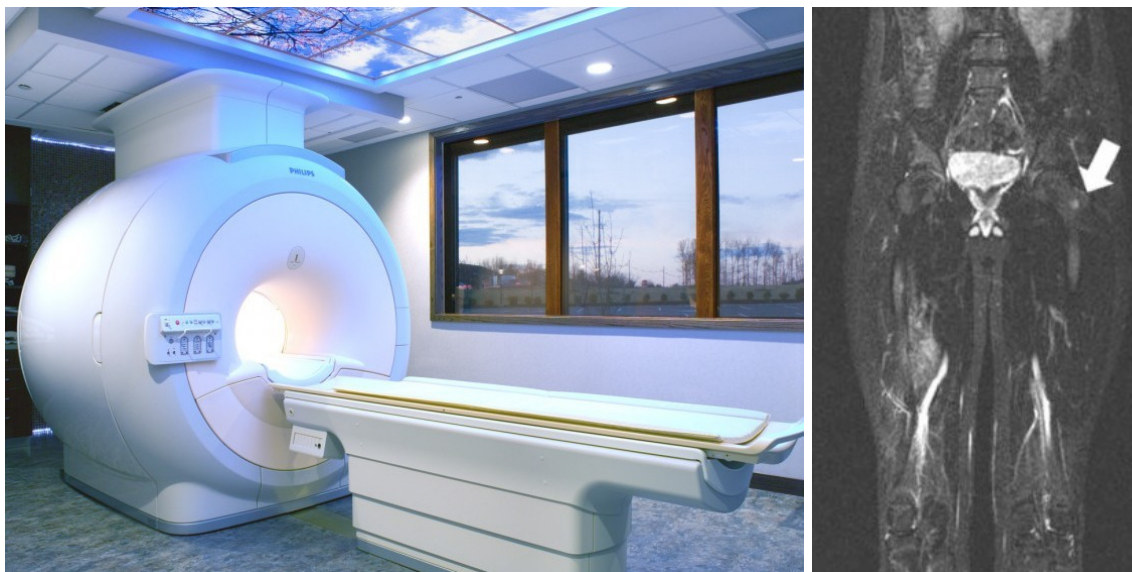


Figure 4 – MRI machine and a MRI exam with suspicion of bone metastases disease [95].

#### 2.1.6 Positron emission tomography (PET)

Positron emission tomography (PET) is a type of nuclear medicine procedure that measures metabolic activity of the cells of body tissues. PET is actually a combination of nuclear medicine and biochemical analysis. It is used mostly in patients with brain or heart issues or cancer and helps to visualize the biochemical changes taking place in the body.

PET can be used to detect metastases by creating a 3-dimensional representation of the patient's body with the use of radio waves that show them up after an injection of nuclear tracers. Although PET scans are not routinely used to look for bone metastases [15], by combining PET and CT, high-resolution images can be obtained with their diagnostic value being comparable to that obtained with the use of the MRI whole-body exam [86]. Figure 5 shows a PET machine and a exam of this type with bone metastasis. There is a new technology combining PET and CT into one scanner, known as PET/CT. It promises great advances in the diagnosis and treatment of lung cancer, epilepsy, Alzheimer's disease and coronary artery disease [90].



Figure 5 – Machine used to perform PET scans and an exam of this type with bone metastases diagnosis [96].

## 2.2 Bone scan procedure details

Although there are several exams for diagnosing bone metastases, as discussed in previous section, bone scan or bone scintigraphy is still one the most common (and the one with highest sensitivity (95%) to diagnose metastases) procedures of nuclear medicine, according to the European Association of Nuclear Medicine (EANM) [13]. It is usable for both men and women [17] and when other examination methods are unable to provide a reliable diagnosis, scintigraphy becomes the most suitable means of making a final conclusion [17].

Bone scan is a type of nuclear radiology procedure because in it a tiny amount of a radioactive substance is used to assist in the examination of the bones. This substance, called a radio nuclide (radioactive nuclide, radioisotope or radioactive isotope) has excess of nuclear energy, making it unstable. These are species of atoms that emit radiation as they undergo radioactive decay (or tracer) being collected within the bone tissue at spots of abnormal physical and chemical change [19].

The radionuclide emits a type of radiation (called gamma radiation) through the emission of gamma rays. This radiation is detected by a scanner, which processes the information into a picture of the bones [19]. The amount of the radionuclide injected into the patient's vein in this examination is small and there is no need for precautions against radioactive exposure. However, the injected tracer may cause some slight discomfort.

Allergic reactions are rare, but may occur [19]. Figure 6 shows examples of images obtained with this procedure. Bone scan images are usually presented with foreground in white color and background in black, however these colors were inverted in the figures showed in this dissertation. This inversion was made to allow better visibility of the findings in the exams in printed texts.

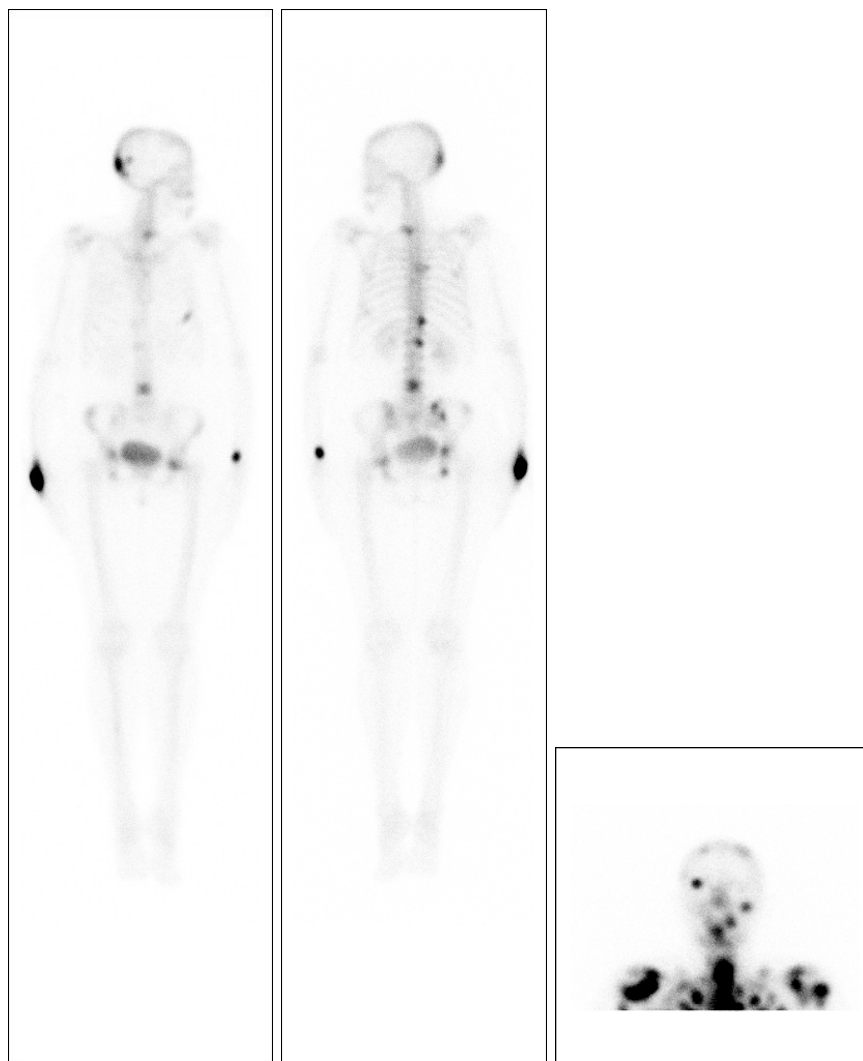


Figure 6 – Ventral, dorsal and cranium bone scan images.

During the exam, after lying the patient on the examination table, the scanner starts taking images from different directions. During these acquisitions, the patients are requested to lie face down and up so that images can be taken. These are from the whole body or from small regions (usually cranium) [18]. Commonly full body dorsal (toward the back) or Ventral (toward the belly) images have a 256 x 1024 resolution of pixels and cranium images have 256 x 256 pixels.

The areas where the radionuclides are concentrated are called "hotspots" and may indicate the presence of conditions such as arthritis, malignant (cancerous) bone

tumors, metastatic bone cancer, bone infections, bone trauma (invisible on ordinary X-rays) and other bone conditions [19].

## 2.3 Medical files: DICOM standard

The Digital Imaging and Communications in Medicine (DICOM) format is recognized by the International Organization for Standardization (ISO 12052) [20] as the current international standard for medical imaging and related information.

DICOM defines a format that can be shared, as well as the data and quality required for clinical use [20]. Since its first publication in 1993, this standard has been responsible for an evolution in the practice of radiology, allowing the replacement of X-ray film by a fully digital workflow [20].

DICOM is implemented in many programming tools. It is used to store outputs for almost all scanners used in radiology, cardiology and radiotherapy devices (X-rays, CTs, MRIs, ultrasound, among others). It is also increasingly adopted in devices from other medical domains, such as ophthalmology and dentistry [20].

Files defined in the DICOM standard follow a model commonly composed of information entities that represents the patient, study, series, equipment, frame of reference, data type and the effective values of the instance. These entities, after reading and decoding the file, are accessible through unique keys called *Tags* in a simplified key-value model [21].

## 3 Related works

Up to now, some investigations have been conducted concerning the application of machine learning and image processing with emphasis on the diagnosis of cancer metastasis from bone scan images. All found works are commented in the following lines, they are organized in chronological order considering their publication year (and grouped according the decade of publication).

### 3.1 Works up the first decade of this century

Analyzing the works carried out on the automation of classification and quantification of bone metastases, the literature treats mainly the introduction and application of Bone Scan Index (BSI) using dedicated software. This index was proposed by investigators of Memorial Sloan-Kettering Cancer Center (New York, NY, USA) and developed during 1977 as a promising means of measuring bone metastasis [93].

BSI is a quantitative assessment of bone scan data that represents the total tumor burden as the fraction of total skeleton weight. In preliminary studies [92], the BSI has shown clinical utility as a *prognostic and progression biomarker* to assess the association between therapy and bone burden. Despite demonstrating clinical utility, the application and dissemination of the BSI in clinical practice have been limited because of the laborious manual calculations [91].

In 2006, a Swedish group developed a CADx system called *Exini Diagnostics AB bone* based on Image Processing and Artificial Neural Network (ANN). This tool was able to completely automate the analysis of whole-body bone scans in order to help calculate BSI [4]. The ANN is an machine learning algorithm that is trained using examples derived from large databases. This learning technique are data-driven, (i.e. learning by training) then a database consisting of a large number of bone scans must be used in it. That study was first conducted with anterior and posterior bone scan images from a total of 200 patients of which 100 were used for training while the remaining 100 were used for validation (400 images formed the used dataset) [4]. The project was further improved using a bigger dataset, composed by bone scans acquired from 810 patients [5].

*Exini Diagnostics* used three basic operations in the image processing: skeleton segmentation, hotspot detection and feature extraction-classification. Regarding to segmentation, two different approaches were used to segment the skeleton. In the first published version, simple threshold was used to separate the bones from the background.

In later versions, active-shape models (ASM) were used, which is a statistical approach to find different skeleton parts [5]. A region-specific threshold algorithm was used to detect potential hotspots. The algorithm was based on simple static evaluation of the mean and standard deviation of all pixel in a specific region. Clusters of pixels with values of this features above a fixed threshold and a defined area are considered hotspot. The used cluster size must be equal or greater than 6 pixels in the 2006 version and bigger than 13 pixels in 2009 to be labeled as potential hotspots [5]. IA algorithms were implemented to extract useful parameters from the segmented regions. The size, shape, intensity and localization of a hotspot were calculated, as well as the intensity characteristics of the region in which the hotspot was located. Fourteen features were used in the first CADx to describe the anterior and posterior images, while this number was extended to 45 features to describe each hotspot in later versions of the software, with additionally 26 features to describe the whole bone scan examination [5]. This features includes: area, width, height, maximal and minimum value, standard deviation, location at skeleton, symmetry, bladder and area ratio of hotspot to region [4].

About the results obtained from the first version of *Exini*, the automated method correctly identified 28 of the 31 patients with metastases in the test group, i.e., a sensitivity of 90%. A false positive classification of metastases was made in 18 of the 69 patients not classified as having metastases by the experts, resulting in a specificity of 74% [4]. An improvement between the first and the later *Exini CADx* version is that the specificity was increased from 74% to 89%. True negative interpretations were made for 34 of the 38 patients classified as not having bone metastases. The second CADx system made correct interpretations for 19 of the 21 patients with bone metastases, showing the same level of sensitivity of 90% as in the first CADx [5].

## 3.2 Works from 2011 to now

*BONENAVI software* is a bone scan classification tool introduced in Japan by the **FUJIFILM RI Pharma** (Tokyo) with a database from Japanese patients with prostate, breast and other types of cancer [3]. They used ANN algorithm as the *Exini software* but trained in a multi center database. The first BONENAVI version[6] (2012) was trained using a set of 904 patients from a single Japanese institution and the revised version of the software (2013) was trained with bone scans coming from 1,532 patients of nine institutions [7]. The diagnostic accuracy of the softwares was investigated and the ROC analysis based on the ANN value showed significant improvement from *Exini*. In men, the area under the curve (AUC) that was 0.877 for *Exini*, go now to 0.912 for the first version of BONENAVI and 0.934 for the last BONENAVI release [7]. In women, the AUC that was 0.831 for *Exini* increases to 0.910 for first BONENAVI and 0.932 for the

later BONENAVI version. The sensitivity and specificity based on the last version of BONENAVI was 90% and 84% for men and 93% and 85% for women [7].

Five years ago, a solution called *Computer-aided diagnosis system for whole-body bone scintigraphy scans* (CADBOSS) was developed as a supplementary software to facilitate physicians decision making on bone metastasis. The dataset used 130 images of whole body bone scans from 60 patients who underwent bone scintigraphy for bone metastatic diseases from 2003 to 2013. Of those 130 images, 100 contained metastases. Ten fold cross validation was used during the training and test phases [94]. CADBOSS consists of 4 main operations: hotspot segmentation, feature extraction, feature selection, and classification. The method used for hotspot segmentation was the *Level Set Active Contour* (LSAC). After detection of hotspots with segmentation, feature extraction was carried out. The best features was selected and decreased by using the Principal Component Analysis (PCA) method, because too many features would negatively affect the system results. In the final stage, images are classified by the ANN as presenting or do not presenting metastases [94].

There is a preprocessing phase in CADBOSS tool. As the sizes of the images is quite high for directly inputting them to the neural network, first of all, the whole body images were reduced to  $200 \times 700$  pixels and a set of sub images were created by dividing each image into 25 equal pieces horizontally as well as vertically. Thus, were obtained 625 sub-images with a size of  $8 \times 28$  pixels. In order to remove the impact of the background pixels still present in these sub-images, the average value of each one of these was transferred to a matrix. Before applying whole-body bone scans to the ANN, the images were transformed to be represented by features and optimized for be used in classification phase. Later, only the ten (10) most important features were selected among the 625 initial values computed and send to the PCA. Finally these 10 values were then used in the ANN as the input patterns to be classified [94].

The reported results obtained by the CADBOSS advertises are that it was able to correctly identify the presence or absence of metastasis in 120 of the 130 images. The accuracy, sensitivity, and specificity of the software were 92.30%, 94%, and 86.67% respectively [94].

At same time (i.e. in 2016) another relevant work emerges at Lund University (Sweden). It presented the use of convolutional neural networks (CNNs) for hotspot classification of scintigraphic scans [8]. The motivation of this was to provide results obtained by a fully automated method, without the needs of hand-designed heuristics. The used dataset in this work was provided by *Exini CADx* in the form of image patches of already found hotspots collected from exams of 2,164 patients. Due to execution time restrictions, only the hotspots found in the spine were used to train the CNN due to the high availability of data in that area [8] and as they were considered to be easiest for

classification. Thus, the used dataset consisted of a total of 10,428 examples where 3,170 of them belonged to the “high risk” class, while the remaining 7,258 examples were from “low risk” class. The dataset was splitted into training set, validation set and test set, each one having 4,171, 3,128, and 3,129 hotspot images respectively [8]. The used CNN was implemented using the `Keras` tool and because of the slight unbalance between the two classes in the training data, under sampling was used, in order to have a good combination of faster training and good performance. The numerical performance results of the trained CNN of this work on both the validation set and test set are presented in Table 1.

Table 1 – Results for Lund group using CNN and 2,164 patients [8]

	Validation set	Test set
Accuracy (%)	87	89
Sensitivity (%)	98	98
Specificity (%)	61	65
AUC	0.95	0.95

In 2017, another project from the Lund University studied the use of CNNs to classify bone scan hotspots as metastatic or non-metastatic. The dataset for this contained a total of 10,427 hotspots provided by *Exini*, of which 3,169 are positive (high risk of metastasis) and 7,258 are negative (low risk of metastasis) [9]. The hotspots were segmented, using image processing methods already available in *Exini*, from whole-body bone scans of patients with confirmed prostate cancer and suspected metastatic disease. First, in order to find optimal hyper-parameters, approximately 20% of the data were separated to form a test set, used to calculate the final performance. The remaining data were subsequently split into equally sized ‘folds’ for K-fold cross-validation with the number of folds 3 (i.e.  $K = 3$ ). For all results obtained in this work the AUC was used as a performance index. The final test set performance was calculated by taking the optimum hyper-parameters obtained after a random hyper-parameter search task and retraining a network using the training and validation data combined, thereby extending the number of training examples. At the end of this work, the best performing ensemble of CNNs gave an AUC score of 0.97 [9].

More recently, in 2020, a study investigates the application of a deep learning (CNN) to build a computerized solution that automatically identifies whether a patient is suffering from bone metastasis or not, based on patient’s whole body bone scans. In such work, was performed a retrospective review of 970 exams of whole body scintigraphy images from 817 different male patients who visited *Nuclear Medicine Department of Diagnostic Medical Center* in order to have diagnostic in Larisa, Greece, from June 2013 to June 2017. The selection criterion was prostate cancer patients who had undergone whole-body scintigraphy due suspected bone metastatic [10]. The initial dataset of

970 images did not contain only bone metastasis, there were patients suffering from other degenerative lesions. Due to this fact, as well as aiming to cope with a two-class classification problem in the study, a pre selection phase, concerning images of healthy and cancerous (malignant tumors) patients was accomplished. Specifically, 586 of 970 whole-body scintigraphy images of men from 507 different patients were selected and diagnosed, of those, 368 bone scans were from men with bone metastasis and 218 from men without bone metastasis [10]. Then, the CNN method for bone metastasis classification proposed in that work includes four major processing steps: data pre processing, feature computations from the selected bone scan data, training phase for CNN learning and validation (or testing, that is the phase where the tool did the evaluation of the classified results). Regarding to the **data curating** that work applied the following steps: data normalization; data shuffle; data augmentation and data split.

The Data normalization is used to ensure that the dataset values are in the same scale (i.e. within [0, 1]), that is the min-max normalization was applied. The data shuffle avoided or eliminated unbiased sampling in machine learning, in it, a random-number generator was used to reorder the bone scan images. In the data augmentation used in that research, the number of images used for learning processes was followed by an image augmentation processing of rotation, enlargement, reduction, range changes and flip [10]. The data split divides the dataset into three groups for training, validation (that would allow the training process to be improved) and a testing group. The first data split takes place by removing 15% of the total dataset and saving for later use as testing. The remaining 85% of the dataset is then split again into an 80/20 ratio, where the small portion is used as validation set [10]. In that study, a deep-layer network with 3 convolutional pooling layers, 1 dense layer followed by a dropout layer, as well as a final output layer with one node was used. Several CNN architectures were tested, leading to the one that performed best under all hyperparameter selection cases and regularization methods. The final output was: recall of 98%, accuracy of 97% and precision of 95% [10]. The best confusion matrix obtained for that work is presented at Table 2.

Table 2 – Best confusion matrix of Larisa group (970 exams) [10].

	Malignant in ground truth	Healthy in ground truth
Malignant as output (%)	61	0
Healthy as output (%)	1	26

According to both sections of this chapter, it is possible to state that the developed of bone scan aiding tools has had a notable advances in the last years, but there are still opportunities for new approaches, implementation strategies and to promote better interaction between the computer system and medical community.

## 4 Image analysis

A single channel or *one band image* may be defined as a two-dimensional function,  $f(x, y)$ , where  $x$  and  $y$  are spatial (plane) coordinates. The value of  $f$  at any pair of coordinates  $(x, y)$  is called the intensity of the image at that point [45]. When  $x, y$  and  $f(x, y)$  are all finite and discrete quantities, it is called a **discrete image** [29]. All discrete images can have their  $x$  and  $y$  variables mapped to integers [74]

A *digital image* contains an integer and limited number of points or pixels:  $M \times N$ . Pixels are the smallest individual element in an image, holding values that represent the brightness of the corresponding area when it is presented in the computer screen or the intensity of an acquired physical element from a sensor (like the x-ray or  $\gamma$ -ray for some types of medical imaging). Typically, pixels of a digital image or raster map are represented in a form of a two-dimensional array where each cell,  $f(x, y)$ , stores an integer or a real (floating point) value. Moreover, these  $f(x, y)$  values can be remapped to a level of gray and presented for visual interpretations as of those presented in Figure 6. This representation is named a monochrome or gray level digital image [45].

The image processing (IP) considers transformation of digital images by computer algorithms and programs [28]. It can be considered as a sub area of signal processing, which focuses primarily on signals perceptible by human eyes (generically named a 2D signal or an image). Digital image processing allows users to take a digital image as an input and perform different approaches on it to generate a new image as output [69].

However, digital images can be processed in a way that the output is other kind of information. For instance the answer of these questions: There is metastasis in this point? Or what is the total area of this image that can be considered metastatic? This kind of result is not an image and such types of outputs are the main goals of Image Analysis.

Today, there is almost no medical diagnostic that is not impacted in some way by digital image applications. This dissertation makes extensively use of IP and IA techniques. Due to this, next sections of this chapter present the main used steps and operations that will be used in the CADe and CADx development.

### 4.1 Image analysis steps

Image processing (IP) and image analysis (IA) approaches present some steps from the signal acquisition or the image retrieval from a database to its uses in high level systems. Such steps compose the workflow sequence of a specific system [72]. The IA

and IP concepts used are commented in next subsections and their implementations details are presented in Chapter 6.

#### 4.1.1 Image acquisition

Although image acquisition could be as simple as using images from a database previously organized and already stored in digital form, generally for medical applications (like this one) such a step involves lots of effort including approvals from ethical committee, privacy considerations and permissions of institutions and patients involved [28]. In this work are employed the DICOM files of examination previously done on patients of the UFF university hospital (Hospital Universitário Antônio Pedro - HUAP) from the nuclear medicine group.

#### 4.1.2 Image enhancement

Image enhancement is the step that uses procedures for improving the quality of the information content of original data before processing in order to turn it more suitable than the original for a specific application [73]. There are various enhancement techniques. Common applications include contrast enhancement, domain transformations, spatial and frequency filtering [28].

#### 4.1.3 Feature extraction

In IA, IP and Pattern Recognition, the term *feature extraction* (or *feature computation*) consists of transform a visual information in digital one, that could be numerical, categorical or Boolean. Feature extraction includes methods for computing abstract information that helps on making decision about every image point. The resulting features will be elements of the feature array (or feature vector) used often for some decision criteria about areas or points of the images related to assignments needed to be done or to a classification of a pattern [28].

#### 4.1.4 Segmentation

Segmentation is defined as the process of partitioning a digital image into multiple parts (segments). The goal of segmentation is to simplify or change the representation of an image into something that is more meaningful and easier to be analyzed [78].

The result of image segmentation is a set of parts that cover the entire image. Each of the pixels in a segment are similar according to some homogeneity criteria. For instance, one criterion could be see if "a given pixel corresponds to a metastatic point of

the body". And homogeneity criteria examples are: "these cluster of pixels are from a non metastatic area", or "this set of pixels belong to the background of the image"[75].

In this research a correct segmentation of metastatic area and its quantification is the main goal because it is directly related to the diagnostic of the examination represented by the image on analysis.

## 4.2 Image operations

In the subsections of this section, the filters, algorithms and strategies applied in this dissertation are described.

### 4.2.1 Histogram equalization

Image histogram is a graphical representation of the tonal distribution of an image. It shows the number of pixels for each tone [77]. In other words, a histogram  $H(f(x, y))$  of a monochromatic image with  $G$  gray-levels is represented by an array of  $1 \times G$  elements. Histograms are simple to compute [28] and can be used for contrast improvement, threshold (on image segmentation to objects and background) and other tasks [29].

Histogram Equalization (HE) is an example of one possible application of histograms. Equalization aims to achieve contrast enhancement. The goal of this is to create an image with more equally distributed levels over the whole  $G$  gray scale [29]. The algorithm to perform the equalization is [29]:

1. Initialization: For an  $N \times M$  image of  $G$  gray-levels, create an array  $H$  of length  $G$  with 0 values :  $H[g] = 0$  for all  $g = 0, 1, 2, \dots, G - 1$ .
2. Compute the histogram: Scan every pixel  $(x, y)$  of the image. Identify its value  $g = f(x, y)$  and increment the correspondent position  $g$  of  $H$ . That is, any time the pixel  $(x, y)$  has intensity  $g$  then do:

$$H[g] = H[g] + 1 \quad (4.1)$$

3. After, compute the cumulative histogram  $H_c$ :

$$H_c[0] = H[0], \quad H_c[p] = H_c[g - 1] + H[g], \quad g = 1, 2, \dots, G - 1 \quad (4.2)$$

4. Set the equalized gray level:

$$H_{eq}[p] = \text{round}\left(\frac{G - 1}{NM} H_c[g]\right) \quad (4.3)$$

5. Rewrite the image: replacing in each pixel  $(x,y)$  its original gray-levels  $g$  to the new  $p$  setting:

$$p = H_{eq}[g], \quad f(x, y) = p \quad (4.4)$$

#### 4.2.2 Contrast limited adaptive histogram equalization (CLAHE)

Histogram Equalization (HE) is a relatively simple method for contrast enhancement. However, the procedure is based on the assumption that the image quality is uniform over all areas and one unique gray scale mapping provides similar enhancement for all regions of the image. However, when distributions of gray scales change from one region to another, this assumption is not valid. In this case, an Adaptive Histogram Equalization (AHE) technique can significantly outperform the standard Equalization [42].

The main idea in AHE is to find the mapping for each pixel based on its neighborhood gray scale distribution [42]. The number of times that this calculation should be repeated is the same as the number of pixels of the image leading to an extensive increase of computation requirement.

The approach that deals between global histogram equalization and fully adaptive equalization is the Regional Equalization. In this, the image is divided into a limited number of regions and the same histogram Equalization technique is applied to pixels at each one of this regions [42].

The problem with this approach is that due to differences between mappings of two adjacent regions, pixels from the left and right (or top and down) sides of a border can be mapped in very different gray levels causing a significant blocking effect. An alternative to solve this problem is to approximate the mapping required for each pixel by choosing a small number of contextual regions. Thus, the mapping for a pixel is obtained by using a weighted-sum of the mappings of its nearest regions [42].

There are also cases where the gray scale distribution is highly localized. In such cases, the mapping function may include segments with high slopes, resulting that two very close gray levels might be mapped to significantly different ones. This issue is solved by limiting the contrast that is allowed through the histogram equalization operation [42]. The combination of this contrast limiting approach with the aforementioned adaptive histogram equalization results in what is referred as Contrast Limited Adaptive Histogram Equalization (CLAHE).

This method starts by dividing the image into several non-overlapping regions of almost equal sizes; after dividing, the histogram of each region is calculated; then, based on a desired limit for contrast expansion, a clip limit for clipping histograms is obtained. Then, each histogram is redistributed in such a way that its height does not

go beyond the clip limit. Finally, cumulative distribution function (CDF) of the resultant contrast limited histograms are determined for gray scale mapping [42].

One approach to limit the maximum slope is to use a clip limit as maximum of all histograms. This slope upward can be related to what is referred as clip factor, in percent, as follows:

$$\beta = \frac{M}{N} \left( 1 + \frac{\alpha}{100} (s_{max} - 1) \right) \quad (4.5)$$

In this case, for clip factor of zero percent ( $\alpha = 0$ ) the clip limit becomes exactly equal to  $\frac{M}{N}$ , which results into an identity mapping by evenly distributing pixel regions into all possible gray scales.

#### 4.2.3 Variance and variance filter

Variance evaluates the degree of variability or dispersion of a dataset. Its general objective is to quantify how heterogeneous the data are. This variability is normally calculate around the data central value [27]. The variance is the square of the standard deviation, i.e. the second central moment of a distribution, and it is often represented by  $\sigma^2$  [82].

The variance filter  $\sigma^2(x, y)$  of a pixel  $(x, y)$  of a region  $N \times M$  of an image is computed by the square root of the ratio between the sum of squares of the differences from the gray level of each pixel of this region to the region average gray level and the number of elements  $N \times M$  of this region:

$$\sigma^2 = \frac{1}{NM} \sum_{x=1}^N \sum_{y=1}^M (f(x, y) - \mu)^2 \quad (4.6)$$

where  $\mu$  represents the mean gray level of the region:

$$\mu = \frac{1}{NM} \sum_{x=1}^N \sum_{y=1}^M f(x, y) \quad (4.7)$$

#### 4.2.4 Gaussian functions and Gaussian filter

In one dimension, that is for an array of values,  $f(x)$ , the Gaussian function is the probability density function of the normal distribution, sometimes also called the frequency curve [81]. Gaussian equation is presented in terms of the standard deviation  $\sigma_x$  and mean or average  $\mu_x$  of the distribution values [23]:

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}} \quad (4.8)$$

For two-dimensional Gaussian filters, there are two standard deviations. They can present the same values in each direction or do not. First, let's consider a normal distribution with equal standard deviations  $\sigma = \sigma_x = \sigma_y$  in both directions and average  $\mu_x$  and  $\mu_y$  respectively. This Gaussian filter appearance is presented on Figure 7 and is defined by the equation:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x-\mu_x)^2 + (y-\mu_y)^2}{2\sigma^2}} \quad (4.9)$$

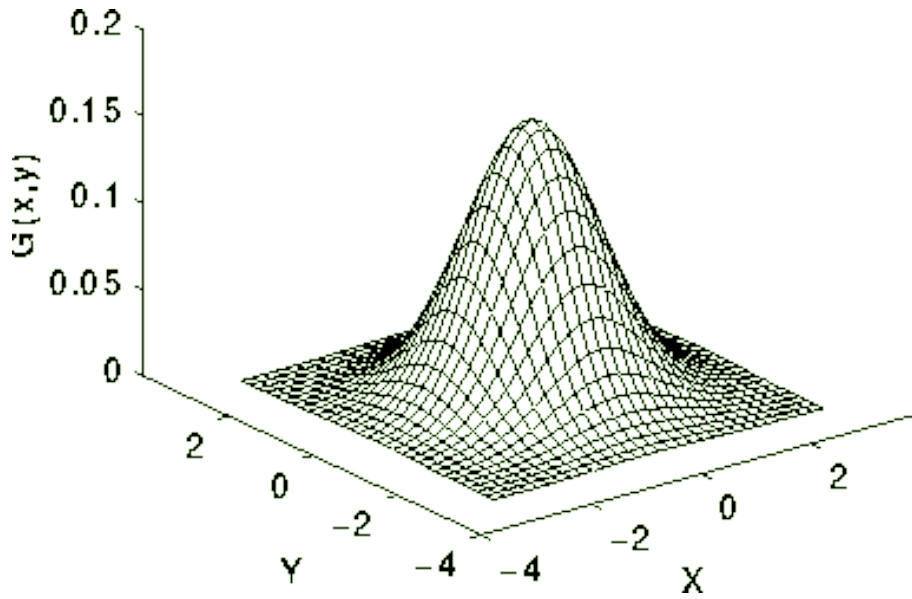


Figure 7 – Gaussian surface with same standard deviation in both directions.

The equation of 2d Gaussian filter for diverse standard deviations in each direction, that is for the cases where  $\sigma_x \neq \sigma_y$  is given by:

$$G(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2} - \frac{(y-\mu_y)^2}{2\sigma_y^2}} \quad (4.10)$$

The Gaussian filter in spatial domain is computed through a two-dimensional convolution. It smooths the distribution, resulting in blurring the image on processing, acting as a low-frequency filter [23].

The effect of Gaussian smoothing is similar to the median and average filters. The Gaussian results for the central pixel gentler smooth the region than the result of the mean or average filtering [23]. The degree of smoothing is determined by the standard deviation  $\sigma$  of the Gaussian. In a digital representation of this filter a larger standard deviation in the Gaussian functions require more elements in the convolutional kernels in order to the discrete form of this continuous functions be accurately represented [23].

### 4.2.5 Otsu thresholding

When for an image  $f(x, y)$  there are only two possible values for each  $(x, y)$  pixel, it is called a *black-and-white* image or a *binary* image. Binary images are useful in many image processing applications due to their simplicity and effectiveness for communication of visual information. Most of the time the values of  $f(x, y)$  are refereed as 0 and 1, even when for the image representation the white color is represented by the 255 levels in a 1 byte pixel representation of the gray levels [43].

The transformation of a gray level to a binary image generally involves two steps: (1) determination of the gray tone  $T$  to be used for threshold (according to some objective criteria) and (2) the assignment for each pixel to the class of 0 or 1. If the pixel intensity is greater than the threshold  $T$  then it is assigned to the 1 class, otherwise, to the 0 class [43]. Consequently, the main problem in this transformation is the choice of the best threshold value  $T$ .

Over the years some algorithms have been investigated and proposed to define this optimal value. The Otsu's threshold technique is a histogram based method which searches for the threshold that minimizes the variance inside each classes and maximizes the variances between two classes [44].

The method proposed by Otsu to define the threshold tries minimize the overlapping of the class distributions [44]. The within class variance is defined as:

$$\sigma_w^2 = \omega_b(t)\sigma_b^2(t) + \omega_a(t)\sigma_a^2(t) \quad (4.11)$$

where  $\omega_a$  and  $\omega_b$  are the probabilities of the classes  $a$  and  $b$  when separated by a threshold  $t$ ,  $\sigma_a^2$  and  $\sigma_b^2$  are variances of these two classes. The probability  $P$  is calculated for each pixel value of the two clusters  $a$  and  $b$  using the cluster probability functions expressed as [80]:

$$\omega_a(t) = \sum_{i=0}^{t-1} P(i) \quad (4.12)$$

$$\omega_b(t) = \sum_{i=t}^{L-1} P(i) \quad (4.13)$$

For a given  $t$ , to minimize the variance within the classes  $\sigma_w^2$  is equivalent to maximize the variance between classes,  $\sigma_b^2(t)$ , because they are related by [44]:

$$\sigma_b^2(t) = \sigma^2 - \sigma_w^2(t) = \omega_a(t)(\mu_a - \mu_T)^2 + \omega_b(t)(\mu_b - \mu_T)^2 = \omega_a(t)\omega_b(t) [\mu_a(t) - \mu_b(t)]^2 \quad (4.14)$$

where:

$$\mu_a(t) = \frac{\sum_{i=0}^{t-1} ip(i)}{\omega_a(t)} \quad (4.15)$$

$$\mu_b(t) = \frac{\sum_{i=t}^{L-1} ip(i)}{\omega_b(t)} \quad (4.16)$$

and  $\mu_T$  is the total mean level of the original image:

$$\mu_T = \sum_{i=0}^L ip(i) \quad (4.17)$$

Once the class probabilities,  $\omega_a, \omega_b$ , and class means,  $\mu_a, \mu_b$ , can be easily computed by iterative algorithms, Otsu technique is an effective image threshold solution.

#### 4.2.6 Median filter

In probability theory, the median divides the higher half of a probability distribution from the lower half. The median of a finite list of numbers can be found by ordering the values and selecting the middle one. Lists are often constructed to be odd in length to secure uniqueness of middle position [29].

Median filtering is a non-linear smoothing method in which the idea is to replace the current point in the image by the median of its neighborhood pixel values. The median of the neighborhood is not affected by individual noise spikes and so median filtering eliminates impulse noise quite well [29].

The application of this filter is performed through computation of the median in a slidding window. The size of this window represents the neighborhood where the median is calculated. Figure 8 shows how the median algorithm works.

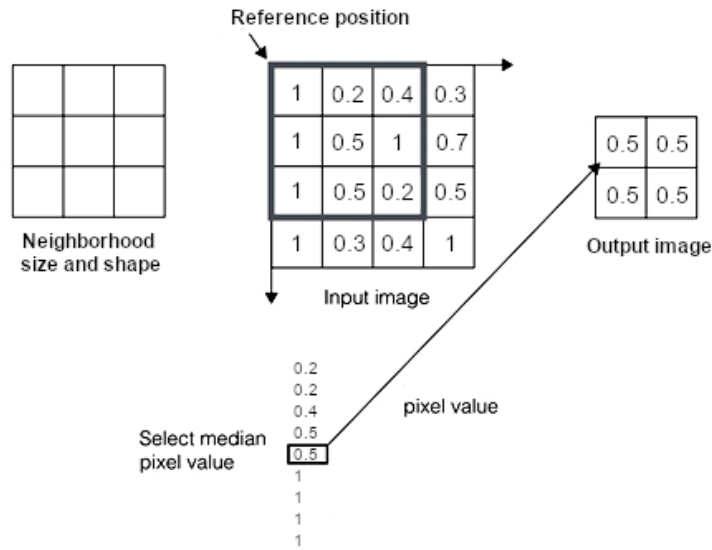


Figure 8 – Window size and the median filter computation.

#### 4.2.7 Entropy filter

In information theory, entropy is evaluated as the logarithm of the number of possible outcomes for a message [24]. In image analysis and processing, entropy is defined by equation 4.18:

$$Entropy(V) = - \sum P(V) \log_2(V) \quad (4.18)$$

where  $P(V)$  is the degree of randomness of a given set of pixels.

For an image, local entropy is related to the definition of a certain neighborhood. This neighborhood can be defined by a rectangular window (Figure 8) or a circular region. The entropy filter can then be computed by Equation 4.18 where  $V$  can be considered described by the gray level of each pixel,  $f(x, y)$ , and the probability  $P(V)$  is the number of occurrence of this gray level in the neighborhood divided by the number of pixels of the region (for all image it is the normalized histogram, or the histogram divided by the total number of pixels) [25].

#### 4.2.8 Roberts filter

The Roberts filtering is one of the first edge detector used in image processing [26]. It is defined by the masks shown in 4.19, their convolution with  $f(x, y)$  and later the combination of their results, producing a new image.

$$Grad_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad Grad_y = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad (4.19)$$

Note that one of the masks corresponds to the result of the 90° rotation of the other. In this way, masks can be applied separately to the input image to measure the gradient component in orthogonal directions. To calculate the absolute magnitude of the gradient, it is necessary to combine the measurements obtained for both of them:

$$|Grad| = \sqrt{[Grad_x(f(x, y))]^2 + [Grad_y(f(x, y))]^2} \quad (4.20)$$

#### 4.2.9 Sobel filter

Like Roberts, the Sobel filtering also applies two-dimensional convolution operations on the image in order to highlight regions of high spatial gradient. It is used to find the approximate magnitude of the absolute gradient from the gray scale of each point of the input image. The convolution masks of Equation (4.21) are designed to respond in a maximum way to the edges in horizontal and vertical directions.

$$Grad_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}, \quad Grad_y = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix} \quad (4.21)$$

The results obtained can be combined by adding the outputs as described on Equation 4.20 or by finding the maximum gradient magnitude. The border orientation,  $\theta$  at each point is:

$$\theta = \tan^{-1} \left( \frac{Grad_y(f(x, y))}{Grad_x(f(x, y))} \right) \quad (4.22)$$

#### 4.2.10 Gabor filters

A two-dimensional Gabor function can be defined as a combination of harmonic oscillator adjusted to some frequency and orientation by a Gaussian [31]. Figure 9 shows a graphic representation of these aspects. By setting the parameters of frequency, orientation and bandwidth, the Gabor equations can be considered as an adjustable filter for specific applications.

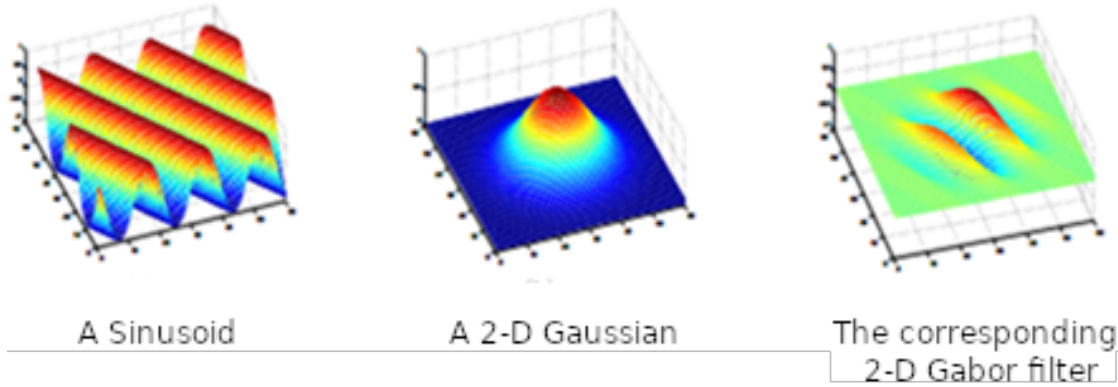


Figure 9 – Elements of the Gabor filters.

Gabor equations present real and imaginary components in orthogonal directions. These components enable combination of the harmonic oscillator as a complex number or individual uses as *sine* and *cosine*. The Gabor equations in the combined exponential (*C* mode), real (*R* mode) and imaginary parts are respectively:

$$Gabor_C(x, y; \lambda, \theta, \psi, \sigma, \gamma) = e^{\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right)} e^{i\left(2\pi \frac{x'}{\lambda} + \psi\right)} \quad (4.23)$$

$$Gabor_R(x, y; \lambda, \theta, \psi, \sigma, \gamma) = e^{\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right)} \cos\left(2\pi \frac{x'}{\lambda} + \psi\right) \quad (4.24)$$

$$Gabor_I(x, y; \lambda, \theta, \psi, \sigma, \gamma) = e^{\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right)} \sin\left(2\pi \frac{x'}{\lambda} + \psi\right) \quad (4.25)$$

where:

$$x' = x \cos(\theta) + y \sin(\theta), \quad y' = y \cos(\theta) - x \sin(\theta). \quad (4.26)$$

In the Gabor function,  $\lambda$  represents the wavelength of the harmonic oscillator (sine or cosine),  $\theta$  represents the orientation of the normal to parallel ranges of the Gabor function,  $\psi$  is the phase angle,  $\sigma$  is the standard deviation of the Gaussian and  $\gamma$  is the spatial aspect ratio and specifies the ellipticity of the Gabor function support. Some notations replace  $\lambda$  by the frequency of the harmonics. Gabor filters are then very controlled filters. They are widely used in image processing for definition of features in specific areas and computation of very local aspects. Most of the time these filters are used in a form of a set of filters or a bank of filters.

#### 4.2.11 Template Matching

Template matching operation consists in going through an image searching for the subset of pixels that has the highest degree of similarity with a set of pixels referred

as a template or pattern. Template matching can be seen as a very basic form of object detection. Using template matching, it is possible to detect objects in an input image using a "template" containing the object desired to detect [39].

To find the match  $R_{obj}$  position in the search image  $I$ , the template  $R$  is sided from left-to-right and top-to-bottom across  $I$ . At each  $(x, y)$  location, a correlation value is calculated to represent the degree of match [39]. Figure 10 shows in the square (left) an example of the pattern to be found in a image with a number of other objects (right) by this technique.

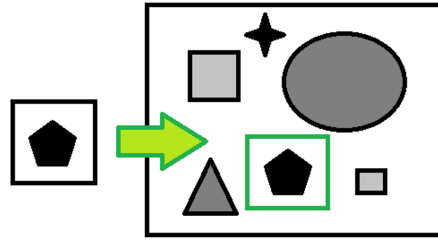


Figure 10 – A template to be found in a image.

This method also relies on evaluating a *quality scores* function over many rectangular sub regions  $R$  of the image  $I$  and taking its maximum as the object's location  $R_{obj}$  [40]. Formally, this search is written as presented at 4.27, where  $R$  ranges over all rectangular regions in the image  $I$  [40].

$$R_{obj} = \operatorname{argmax}_{R \subseteq I} f(R) \quad (4.27)$$

Although there are many *quality score* functions available, one very common and that often presents very good results is the normalized cross correlation. The correlation between two signals or simply the cross correlation is a standard tool for evaluating the degree to which they are similar [41]. It presents two advantages: the first advantage is that the cross correlation is simple to compute and the second advantage is that it is independent of translations and scaling in the intensity domain, or contrast invariant [41].

The cross-correlation  $\varphi_{f_1 f_2}$  of real continuous functions  $f_1$  and  $f_2$  is defined by:

$$\varphi_{f_1 f_2}(t) = \int_{-\infty}^{\infty} f_1(t - \tau) f_2(\tau) d\tau \quad (4.28)$$

In the discrete domain, the correlation of two real time series  $s_i : i = 0, 1, \dots, M - 1$

and  $s_j : j = 0, 1, \dots, N - 1$  is by analogy to equation 4.28 given by:

$$\varphi_{s_i s_j}(t) = \sum_{j=\max(0,k)}^{\min(M-1+k,N-1)} s_{j-k} s_j \quad (4.29)$$

where  $k = -(M + 1), \dots, 0, \dots, (N-1)$ .

When normalized, the cross correlation values computed are re-scaled in other to present unity maximum result. The main advantage of using the normalized cross correlation as a quality score over the regular cross correlation is that it is less sensitive to linear changes in the amplitude of illumination in both pattern and image [41].

### 4.3 Feature vector

A feature vector can be defined as an  $n$ -dimensional array of numerical features that represent some object. In machine learning and pattern recognition, a feature is an individual quantified property or characteristic of a phenomenon [64]. All techniques described in previous subsection can be considered a candidate to compose the vector to represent the patterns present in the bone scan on analysis, in order to help on classifying their areas.

The feature vectors are the basic building blocks of dataset classification by computer techniques. The quality of the features used in a problem has a major impact on insights gained when that information is used for machine learning. At the same time, huge feature vectors could be an issue related to computational effort and processing time.

## 5 Machine learning

Before the reporting of the method and techniques applied in this dissertation, basic concepts of machine learning are introduced in this chapter. In the data science application, learning methods can be informally considered as a means of building mathematical models to help understand data.

The word *learning* in this case refers to fact that for classification based in a set of features, it is possible to use parameters that can be adapted to the observed data and a known classification of them, in a way that the decision model can be considered to *learn* from the set of pairs data-class. Once these models have been fit to previously known set of data, they can be used to predict the class of newly observed data [38].

### 5.1 Supervised or unsupervised learning

In a basic level, machine learning operations can be categorized into two types: *supervised* learning and *unsupervised* learning.

Supervised learning (SL) involves modeling the relationship among computed features from data and associated labels of these data; once the model of association is learned, it can be used to apply labels to new (unlabeled) data. This is further subdivided into classification tasks and regression tasks. In classification, the labels are discrete categories, while in regression, the labels are continuous quantities [38].

Unsupervised learning (UL) involves modeling the features of a dataset without reference to any label. UL is often described as “letting the dataset speak for itself”. This modeling includes tasks such as clustering and dimensional reduction. Clustering algorithms identify distinct groups of data, while dimensional reduction algorithms search for more succinct representations of them [38].

The main focus of this work is the use of supervised learning techniques to perform the classification of each bone scan image. However, the use of unsupervised techniques of dimensional reduction is also addressed.

### 5.2 Ground truth and gold standard

Supervised learning (SL) tasks need a reference to “learn”, it is represented by the name “Ground truth”. Ground truth can be defined as an ideal expected result. In other words, is a way to use expert opinions to create a set of examples with output

labels considered adequate. It might involve hand-labeling examples in order to collect desirable answer data for system training [84].

In the context of Artificial Intelligence (AI) systems, the Gold Standard refers to a set of data that has been manually prepared or verified and that represents “the objective truth” as closely as possible [83]. The term “Gold standard” refers to a benchmark that is available under reasonable conditions. It is not always the perfect test, but the best available one representing a standard for learning. In other words, by giving models plenty of data to learn from, the central hypothesis of AI systems is that it is possible to automatically identify ways to recognize patterns accurately. The data that AI systems are fed with, needs to be unambiguous, otherwise they can’t make any sense from the process [83].

### 5.3 Dimensional reduction

Dimensional reduction (DR) refers to the transformation of high-dimensional arrays into a meaningful representation of same array with reduced number of elements or dimension. Ideally, the reduced representation should have a dimension that corresponds to the most important datum [57]. The intrinsic dimension can be defined as the minimum number of parameters needed to accomplish a task with the feature array employed for representing the dataset [58]. DR is important in many domains, since it mitigates problems such as storage area and computational resources related to the dimension of the used sub spaces [59].

The expression “curse of dimensionality” refers to various phenomena that arise when analyzing and organizing data in high-dimensional spaces that do not occur in low-dimensional settings. Mainly in dynamic programming, if the problem have more features than observations than it run the risk that a massively over fitting could result in terrible out performance [65, 66].

The problem of DR can be defined as follows: Assume we have a dataset represented in a  $n \times D$  matrix  $V$  consisting of  $n$  data vectors  $v_i$ , ( $i \in 1, 2, \dots, n$ ) with dimension  $D$ . Assume further that this dataset has intrinsic dimension  $d$  (where  $d < D$ , and often  $d \ll D$ ) [57]. DR techniques transform the dataset  $V$  with dimensional  $D$  into a new dataset with dimension  $d$ , while retaining the representative degree of the data as much as possible. In general, neither the geometry of the data manifold, nor the intrinsic dimension  $d$  of the dataset  $V$  are known. Therefore, DR is an ill-posed problem that can only be solved by assuming certain properties of the data.

There are many ways to achieve DR, but most of the DR techniques fall into one of two classes: Feature elimination or feature extraction. The Feature Elimination techniques reduce the feature space by eliminating features. The disadvantage is that

by eliminating features, it is also eliminated any benefits those dropped variables could bring [67].

In a problem with  $D$  variables, Feature Extraction refers to the possibility to create  $D_{new}$  independent variables, where each “new” independent variable is designed as combination of some “old” variables. Ideally, these new variables must be created in a specific way that is possible to order them by how well they predict the results. This makes possible to keep as many of the new variables as we wanted, by dropping the least important ones [67].

## 5.4 Principal component analysis

Principal component analysis (PCA) is fundamentally a feature extraction DR algorithm. However it is also useful as a noise filtering in visualization tools [38]. It creates a set of uncorrelated components that are rank ordered by variance (the first component has higher variance than the second, the second has higher variance than the third, and so on) [68].

An algorithm for conducting PCA works as following [67]:

1. Organize the data into  $n$  rows and  $d+1$  columns, where one column corresponds to the dependent variable (usually denoted  $y$ ) and  $d$  columns where each corresponds to an independent variable (the matrix of which is denoted  $V$ ).
2. Standardize the data in order to obtain it's normal or Gaussain distribution (i.e., distribution with null mean and standard deviation equals 1 ). This centered and standardized matrix will be called  $Z$ .
3. Take the matrix  $Z$ , transpose it, and multiply the transposed matrix by  $Z$ . The resulting matrix ( $Z^T Z$ ) is the covariance matrix of  $Z$ .
4. Calculate the eigen vectors and corresponding eigen values of  $Z^T Z$ . The eigen decomposition of  $Z^T Z$  results  $ED_g P_v^{-1}$ , where  $E$  is the matrix of eigen vectors and  $D_g$  is the diagonal matrix with eigen values on the diagonal and values of zero everywhere else.
5. Take the eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_p$  and sort them from largest to smallest. In doing so, sort the eigen vectors in  $E$  accordingly. This sorted matrix of eigen vectors will be called  $E^*$  where the columns of  $E^*$  should be the same columns of  $E$ , but in a different order. Also, note that these eigen vectors are independent of one another.

6. Lastly, calculate  $Z^* = ZE^*$ . This leads to a new matrix  $Z^*$  which is a centered/standardized version of  $V$  but now each observation is a combination of the original variables, where the weights are determined by the eigen vector. Also, because each eigen vectors in  $P_v^*$  are independent of one another, each column of  $Z^*$  is also independent of one another.

A central issue in PCA is the decision about the number of principal components to be retained. A common method to determine this amount is calculate the proportion of variance explained for each feature, then sort features by proportion of variance explained and plot the cumulative proportion of variance explained as more features are keep [67].

After applying this technique it is possible to end up with a reduced number of independent variables and still retain a great amount of the explained variance.

## 5.5 Splitting data for training and testing

One possible technique for dataset split is the one called K-fold Cross-Validation (KCV). This is currently one of the most adopted techniques for evaluating the performance of a model on a set of information [61].

The cross-validation technique consists of dividing the database into  $k$  parts, each consisting of samples of size  $N_r/k$ , where  $N_r$  represents the total number of records in the dataset. After this division,  $k-1$  parts are used, in turn, as a training set, while the remainder of the samples are used as a validation set [60].

An advantage of the KCV method over simple or stratified division of a database between training and testing is the repeated use of all data available for the construction of a learning model and to test it. This process can considerably reduce the risk of an unfavorable split [61].

## 5.6 Performance evaluation

For evaluation of an implemented solution performance, commonly used evaluators are specificity, precision, recall and area under receiver operating characteristic curve (ROC).

To validate the degree of relevance of a binary classifier, specificity is defined as:

$$Specificity = \frac{TN}{TN + FP} \quad (5.1)$$

where  $TN$  is the amount of true negatives samples and  $FP$  is the number of samples wrongly classified as positive.

The precision (also called positive prediction value in the literature) refers to the ratio of relevant instances to the total instances defined as part of the positive binary class:

$$Precision = \frac{TP}{TP + FP} \quad (5.2)$$

where  $TP$  is the amount of true positives samples classified.

The recall (also commonly called sensitivity) is the ratio of the amount of positive elements classified correctly to the total of elements of this class present in the data.

$$Recall = \frac{TP}{TP + FN} \quad (5.3)$$

where  $FN$  symbolizes the positive elements wrongly classified as part of the negative binary class.

The ROC is a graphical representation that illustrates the performance of a binary classification system as its discrimination threshold varies. The area under this curve (AUC) represents the degree of separability. In other words, it outlines how well the model is able to distinguish between the classes. The higher the AUC, the better is the prediction capacity of the model.

The ROC curve is plotted with the recall value against the value of 1 minus the specificity, this corresponds to the ratio between the correctly classified negative class values in relation to the total number of values in this class. In the ROC curve, recall is on the vertical axis while  $(1 - specificity)$  is on the horizontal axis.

## 5.7 Learning techniques

There are a large amount of supervised learning techniques, among them the potentially more related to the type of problem considered in this dissertation are: Random Forest (RF), Support Vector Machines (SVM), Naive Bayes and Multilayer Perceptrons (MLP).

In this list, the absence of lazy classifiers like K-Nearest Neighbors (KNN) are justified by the nature of the methodology applied in this dissertation. In preliminary tests carried out, the computational time to execute these algorithms was considered impracticable, since a lazy classification for the vast set of pixels of interest requires a lot of computational time.

### 5.7.1 Naive Bayes

Naive Bayes (NB) classifiers are based on Bayesian classification methods. These techniques rely on Bayes's theorem, which describes the relationship of conditional probabilities of statistical quantities. In Bayesian classification, the goal is to find the probability of a label given some observed features, which can be wrote as  $P(C|features)$ . Bayes's theorem tells how to express this in terms of quantities [38]:

$$P(C|features) = \frac{P(features|C)P(C)}{P(features)} \quad (5.4)$$

These classifiers perform the prediction of elements through calculating the probabilities of the instance belonging to each of the classes available. Learning in these classifiers can be considerably simplified by assuming that the resources present in the feature array are conditionally independent of each other, this is:

$$P(V|C) = \prod_{i=1}^n P(V_i|C) \quad (5.5)$$

where  $V = (V_1, \dots, V_n)$  is a vector of attributes and  $C$  is a class [33].

Although this assumption is not consistent with the reality of the problems in most occasions, the resulting classifier family called Naive Bayes provides very good results in practice, even being competitive with other classical methods in terms of accuracy, training and processing time [34].

In the literature, there are many variations of the basic algorithm, this variations are generally related to implementations of the resource and probability calculation involved. The Gaussian's variation looks more adequate for this work. In this, the assumption is that data from each label  $v_i$  is drawn from a simple Gaussian distribution, thus, the probability of resources is assumed to be:

$$P(v_i|C) = \frac{1}{\sqrt{2\pi\sigma_C^2}} e^{-\frac{(v_i - \mu_C)^2}{2\sigma_C^2}} \quad (5.6)$$

where the parameters of mean,  $\mu_C$ , and standard deviation,  $\sigma_C$ , are estimated using the maximum probability of each class  $C$ .

### 5.7.2 Random forest

The Random Forest (RF) technique is based in the Decision Tree classifier. Decision trees (DT) are structures in which there are three different types of nodes: probability nodes, decision nodes and final nodes. A probability node shows the probabilities that the element has to presents certain results. A decision node represents a

test to be performed and a consequent decision regarding an attribute value and a final node shows the class found by a decision path [62].

The essence of the DT prevails in the division of datasets into their sections that indirectly emerge a DT with root nodes at the top. The stratified DT model leads to the final class result by traversing the tree nodes.

During the DT decision-making process it is essential to consider the relevance and consequences of each feature in order to assign the appropriate resources to the nodes closest to the root. To solve this problem of choice and division of these attributes, there are some fundamental splitting parameters such as entropy, information gain, Gini index and others.

The Gini index is a well established and popular splitting parameter to build decision trees. This index, also known as Gini impurity, calculates the amount of probability that a specific trait has to be miss classified when randomly selected. It is determined by subtracting from one the sum of squares of the probabilities,  $P_i$ , of each class.

$$Gini = 1 - \sum_{i=1}^n (P_i)^2 \quad (5.7)$$

The Gini index is a real number that varies between the values 0 and 1. Value zero, 0, expresses the purity of the classification, that is, all elements belong to a specified class or only one class exists in it. A value of 1 indicates the random distribution of elements across classes. The index value of 0.5 represents an equal distribution of elements across classes.

Random forests are a combination of tree predictors so that each tree depends on the values of an independently sampled random vector with the same distribution for all trees in the forest. After a large number of trees are generated, they vote for the most popular class. These forests generally correct the bad habit of overfitting from decision trees [63].

### 5.7.3 Support vector machines

Support Vector Machines (SVMs) are defined as a set of supervised learning techniques used in tasks of classification, regression and extreme point detection (outliers). A support vector machine constructs a hyperplane or a set of hyperplanes in a high or infinite dimensional space [35].

Intuitively, a good separation is achieved by the hyperplane that has the greatest distance to the closest training data point of any class, this is called the functional margin. In general, the larger this margin, the smaller is the error of the classifier [36]. Figure

11 illustrates the basic behavior of the algorithm. Note that in this figure, Hyperplane1 does not separate the classes. Hyperplane2 separate them, but with a small margin and Hyperplane3 separates them with the maximum margin.

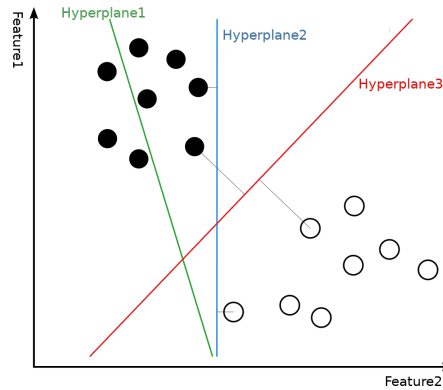


Figure 11 – SVM hyperplanes for binary classification.

The original maximum margin hyperplane method builds an excellent classifier for linearly separable problems. However, this becomes very inefficient for problems whose distribution falls outside this category. The nonlinear SVMs classifiers solves this by applying the kernel trick to maximum margin hyperplanes [37]. The proposal is to adjust the method through a nonlinear kernel function, allowing the algorithm to adjust the maximum margin hyperplane in a transformed resource space.

A kernel function accepts inputs in the original lower dimensional space and returns the dot product of the transformed vectors in a higher dimensional space. The selected function varies between problems and implementations, being a parameter to be defined for the use of the algorithm.

#### 5.7.4 Artificial neural network

Artificial neural networks (ANN) are computational algorithms that present a mathematical model inspired by the structure of intelligent organisms. To achieve good performance, these networks employ a massive interconnection of simple computational cells, called neurons, perceptrons or processing units [32].

This atomic unit of the model is composed of three elements:

- A set of synapses characterized by its own weight. In this way, for each input signal  $s_j$  in an input  $j$  of a neuron  $i$ , there is a weight  $w_{ij}$  [32].
- Adder operation, which is responsible for performing the weighted sum of the respective neuron synapses.

- A step function, which is responsible for restricting the possible range in the output amplitude [32].

Another component commonly found in perceptrons is bias. This corresponds to a constant value selected to increase or decrease the activation function input value. The Figure 12 illustrates the model of a mathematical neuron.

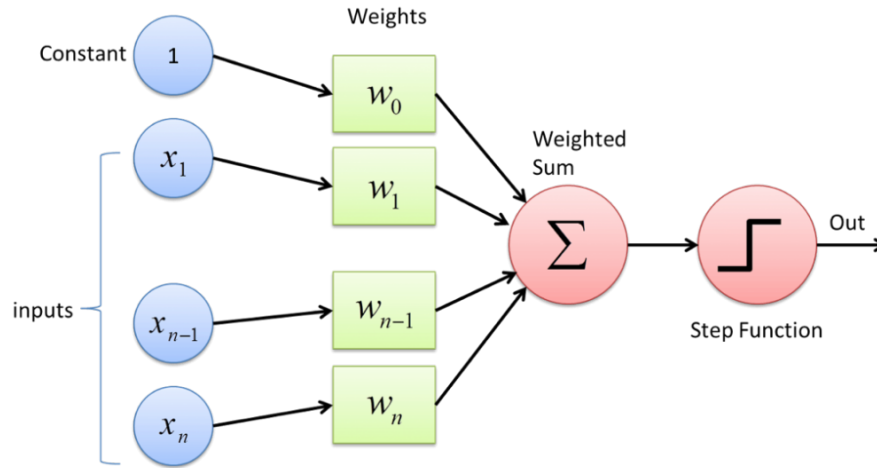


Figure 12 – Perceptron model.

When using this technique, there are some important parameters to be defined, including the general architecture in the multilevel network, bias value, number of layers implemented and what activation function will be applied.

The Rectified Linear Unit (ReLU) function is an example of such activation functions. The output of a neuron  $i$  that employs this function is defined through the equation 5.8.

$$f(v) = \begin{cases} 0 & \text{if } v < 0, \\ 1 & \text{if } v \geq 0. \end{cases} \quad (5.8)$$

In the literature, this neuron is referred by the McCulloch-Pitts model, describing the so-called all-or-nothing property, in which the response takes the value 1 if the induced input is non-negative and 0 otherwise [32].

In the next chapter the reported techniques will be tested in order to define the best possible one for decisions about the metastatic areas in the CADe, CADx implementation.

## 6 Implemented system

This chapter presents details of the developed system for aiding physicians on the task of detect possible areas of metastasis and to perform their quantification based on the bone scan examination. Such details are related to all programmed aspects from how the users interact with the application up to other invisible to users tasks (as internal processing and decision related to programming image recognition and classification decisions). That is, this chapter presents elements from the interface technologies up to fundamental decisions based on experimentation that had to be accomplished for proper parameter and implementation definitions.

### 6.1 An overview of the technologies used

The application's interface uses the World Wide Web (WWW) and was implemented using ReactJs, Hyper Text Markup Language (HTML), Cascading Style Sheets (CSS), Syntactically Awesome Style Sheets (SASS) and TypeScript (TS) technologies. Available languages for descriptions and communication between the machine and the user in the system interface are English and (Brazilian) Portuguese.

This front-end application communicates through requests under the Secure Hypertext Transfer Protocol (HTTPS) with an application programming interface (API). The body of the requests is written in JavaScript Object Notation (JSON) and the methods available for communication follow the Representational State Transfer (REST) architecture.

The API, mostly implemented with NodeJs, TS, ExpressJs and Python technologies is responsible for handling and processing the various requests entering from the web interface, regarding authentication, management and processing of exams and patients.

To perform operations related to image analysis (IA) and machine learning (ML), the API executes sub processes (or child processes) written in the Python language with the help of popular libraries and tools such as: Pydicom; Skimage; Opencv; Scipy; Sklearn and Pandas. Pydicom was used for reading and processing files in DICOM format. The Skimage, Opencv and Scipy are employed to execute the set of operations related to the digital image analysis and processing. Sklearn and Pandas are used for the tasks of classification and learning.

The application's back-end also communicates with a PostgreSQL database cluster in order to provide persistence of information related to users, patients and

exams. Figure 13 shows a diagram of the relations among the used tools in the complete implemented architecture.

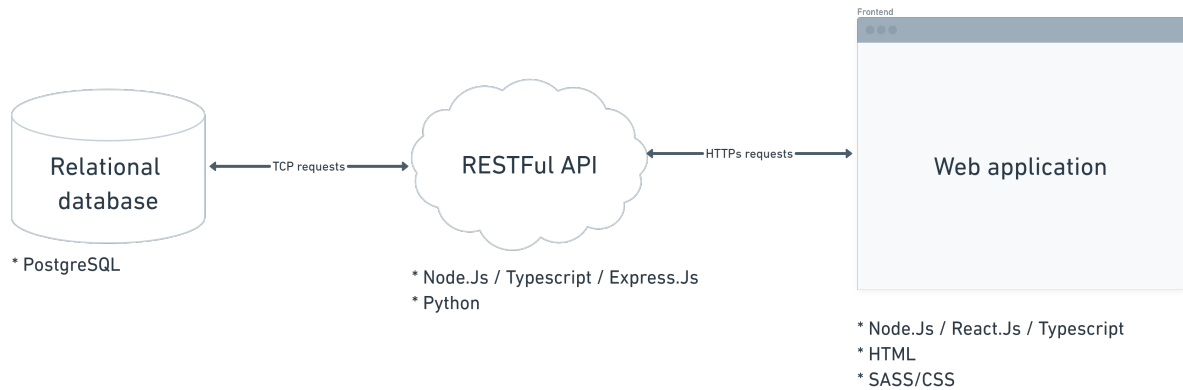


Figure 13 – Relation among the used elements.

The source code of this project is open source, free and accessible at the following online repositories: [github.com/josemorista/bm-server](https://github.com/josemorista/bm-server) (Back-end) and [github.com/josemorista/bm-web](https://github.com/josemorista/bm-web) (Front-end). The implemented web application is also available online for public access at [bm-diag.org](https://bm-diag.org).

In order to achieve bone metastasis classification and to compute the number pixels labeled as with metastasis, the system performs in the back-end a large set of IA and ML operations that will be described in details in future sections of this work.

## 6.2 Operations available to physicians

In the application there are some sequences of operations or functionalities available to the users. Figure 14 shows an use-case diagram of them. In this diagram, the «include» relationship represents functionalities that are implicitly related to the execution of another, and the «extend» relationship represents a functionality available but not required to the execution of another.

Public available functionalities (i.e. to all users even do not yet authenticated) in the implemented application are **sign-up** and **sign-in**. They allow users to create their accounts and authenticate users into the system. Once authenticated, the users may access a set of private operations in the application, such as:

- Include, query or visualize patients;
- Include and visualize exams of previously added patients;

- Upload and process (identify, classify and count areas) bone scans exams files in the DICOM format;
- Visualize automated metastasis classification results (medical reports) and export them in text format.

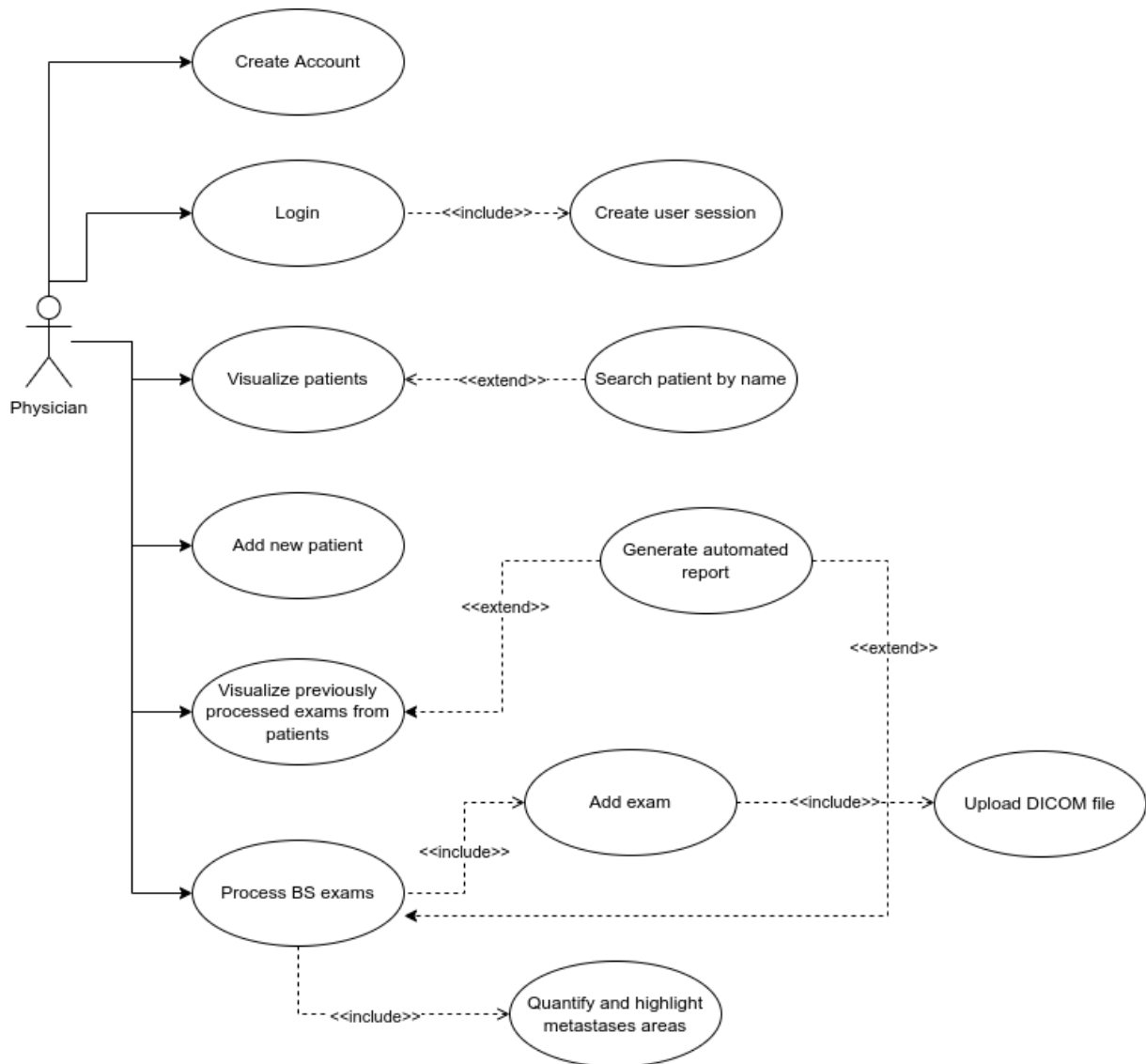


Figure 14 – Diagram of the application functionalities.

### 6.3 Identification and authentication of users

In order to access the full set of features present in the application, the professionals (users) must first be authenticated into the platform. Figure 15 presents the system screen related to this authentication operation. This operation can be done by introducing their identifier (email address) and password. These data must be authenticated

(recognized as valid) by the system. This aspect is necessary to ensure the security and privacy of sensitive data related to exams and personal information of users and their patients within the platform.

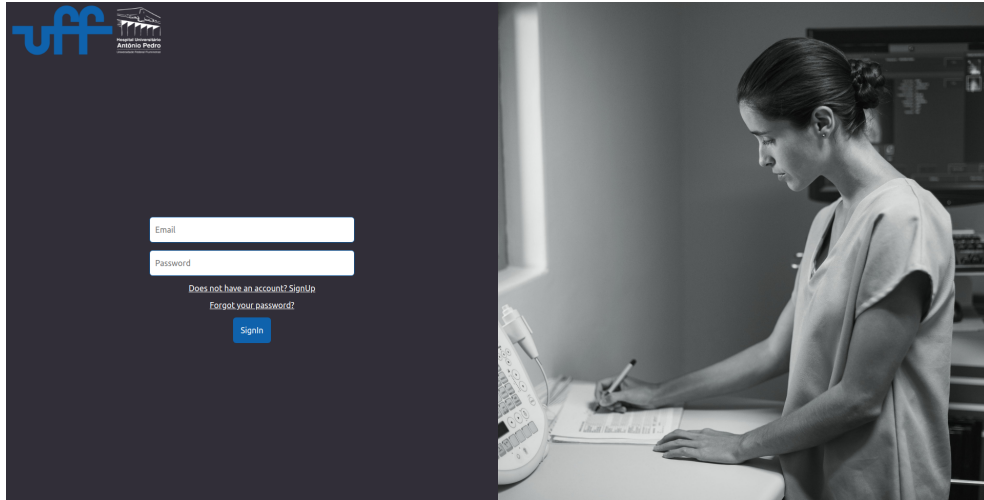


Figure 15 – View of the login system screen.

For new permission of system users (user permissions), the developed application offer a **route** with support for the operation of inclusions of new professionals (physicians or medical doctors) as users of the system. Figure 16 presents the system screens related to **new users inclusion operation**. For this, the new users must provide the following mandatory information:

- Name: This information is collected to provide a more personal user experience in the application. The system use this name in order to interact with the user more directly and friendly;
- Email address: Used as a unique identifier to access the application.
- Password: User secret key, this key is encrypted using the Bcrypt algorithm before being stored in the database.
- Institution: The institution to which the professional is linked.
- Occupation: Specialization, profession or current occupation of the user.

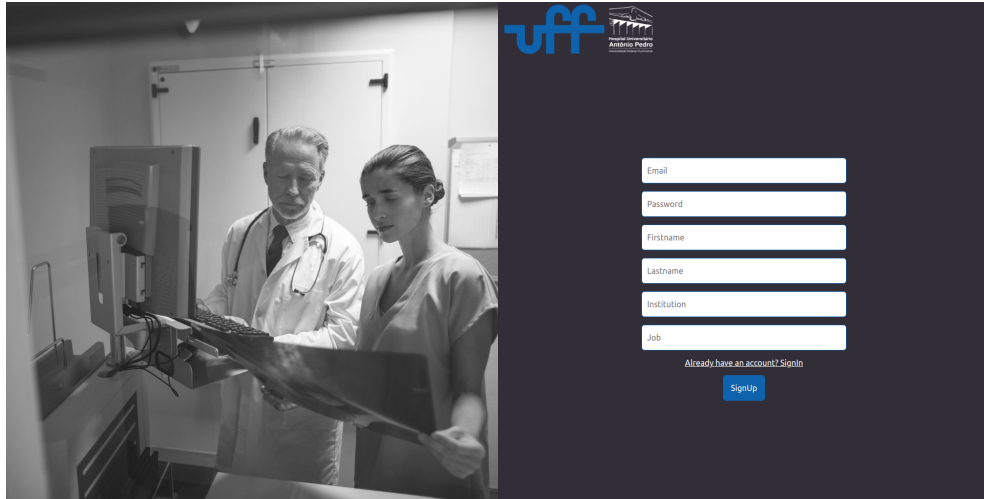


Figure 16 – View of the account creation system screen.

## 6.4 System security

Security and access control are managed through a Bearer Token and JSON Web Token (JWT) strategy. The user's JWT is created and signed by the API at the time of the user authentication. This token has a limited lifetime of 30 minutes, being extensible through security routes. This JWT contains the professional's unique identifier in the system and must be attached to the HTTP "**Authorization**" header in all private route requests. Through this security mechanism, it is possible to guarantee that each professional will only be able to view and manipulate data referring only to their profile and the patients added by them. These include: upload DICOM files, create, edit and view patients information; create, edit and process exams.

## 6.5 Patient management

Once authorized, the user (physician) can use functionalities to create, edit, list, remove patients (Figure 17 shows the system screen for it) and process their exams for automated detection and diagnosis of metastasis. In order to add a patient into the system, the following information is considered mandatory:

- Patient's full name: that is used for research and identification;
- Date of birth: that is used to calculate the patient's age.
- Gender: information collected as an individual attribute for future statistic uses.
- Description: that is a free text field for medical and complementary observations about the patient.

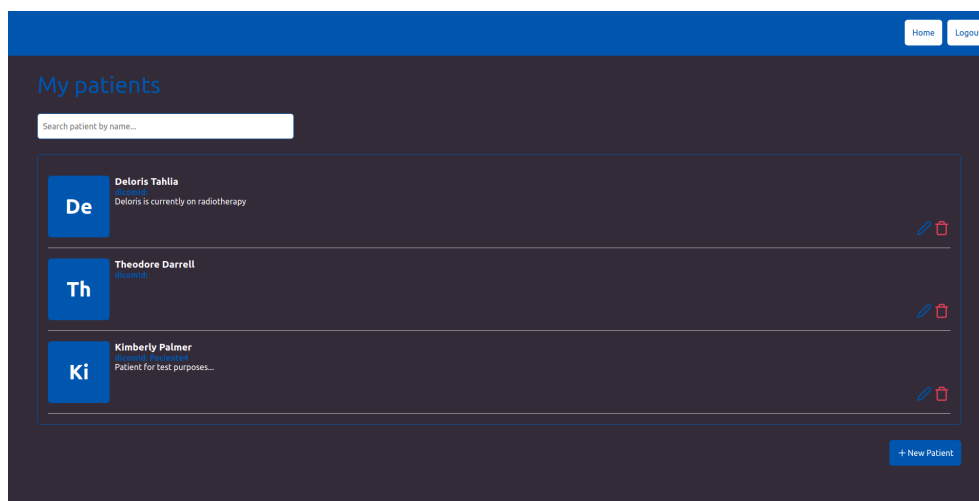


Figure 17 – View of the system screen of patients management.

Once a patient was added, by clicking in the patient's displayed information, the user is redirect to the screen related to the exams of that patient. In this screen it is possible to visualize already processed exams from this person or start uploading and processing new bone scan exams files of this patient. The option of choose exams to analyze is managed by the screen presented in Figure 18. In this screen, the available exams are presented ordered by their realization dates, from the more recent to the oldest. Already processed exams are assign with a **green** dot and exams do not yet processed with a **yellow** one.

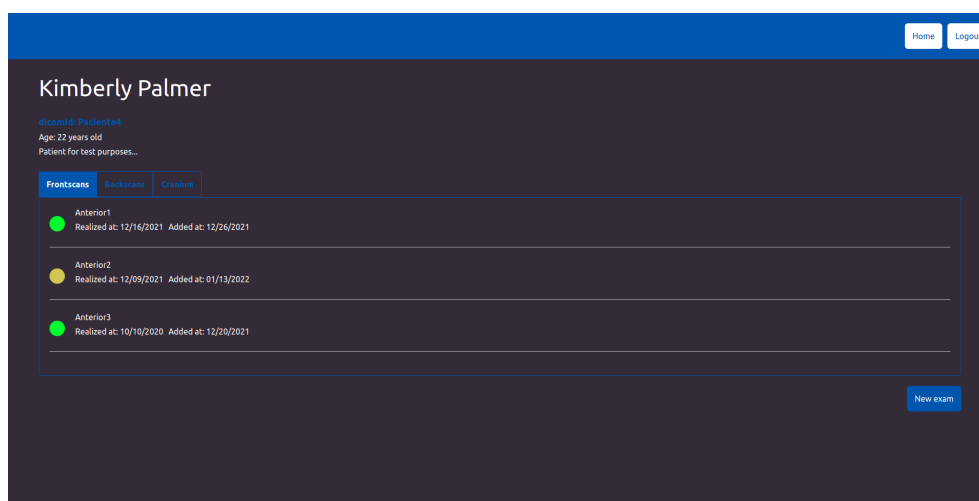


Figure 18 – System screen related to exams visualization and management.

After clicking at a available exam, the user enter the screen where the DICOM file exam is processed and classified by the application. The next sections of this chapter

detail the manipulations of the DICOM file for its image presentation and parameters used in image processing and learning computations steps for detection and classification of the findings.

## 6.6 DICOM meta data

In order to execute the transformation and processing of the files, the API needs a copy of the file in its internal system. For this purpose, exams in DICOM file format must be uploaded through HTTPs requests. Once the DICOM was uploaded to the server, the application can then start the exam processing and transformations.

The first operation performed deals with reading and extracting the necessary information from the file. To perform this task, was used the Python library PyDicom. This tool can easily recover the metadata and pixel values available in a DICOM file content and expose it as a Python data structure called dictionary (this is composed by a set of values of different types accessible by keys).

The most important recovered information for image processing is the pixels content accessible by the *Pixel\_data* key. These containing the patient's radio tracer absorption values identified by the scanner. They can be digitally represented as a two dimensional data structure and are the effective image of the bone scan exam. They are referred in this dissertation as matrix of the pixels or image data.

This matrix has a width of 256 pixels and a height of 1024 pixels. As matrices are identified first by their rows, and then by their columns, they have 256 x 1024 resolution or are composed by 262,144 elements (pixels). This resolution is for full body scans ( i.e. ventral and dorsal patient images). Each of this elements can presents integer values from 0 to 4,000.

Other useful information extracted from the DICOM file are:

- *PatientId*: It is a patient identifier presents in the exam metadata, this information is attached to the patient data.
- *PixelSpacing*: It is an information used to calculate the real area of pixels. This is a number that does a relation of a pixel size unit on millimeters, *mm*.
- *RescaleSlope* and *RescaleIntercept*: they are linear coefficients for conversion between the stored values and their equivalent measured units.
- *Bits Allocated*, *Bits Stored*, *Pixel Representation*, and *Sample per Pixel* are tags that describe the stored value of pixels further [22].

*RescaleIntercept* and *RescaleSlope* are DICOM tags that specify the linear transformation from pixels size stored on disk representation to their size in memory representation. Both can form the equation that performs this transformation, considering that abscissa refers to the horizontal ( $S_v$ ) axis and the ordinate refers to the vertical ( $U$ ) axis of a standard two-dimensional linear graph. It is presented by Equation 6.1.

$$U = m S_v + b \quad (6.1)$$

where  $U$  is in output unit,  $m$  is the rescale slope,  $S_v$  is the stored value, and  $b$  is the point of linear interception with the ordinate axis.

One reason to use Equation 6.1 is related to the need of a different representation on stored (file) information versus the real information, but one must be converted in each other. The DICOM conversion can be related to a different range of values need for this representation versus storage easiness in DICOM file. For example, for CT images, whose pixel values are measured in Hounsfield units, a pixel can have negative values. However, CTs in DICOM files are commonly stored in files as an unsigned integer. As a consequence, it is common for CT DICOM files to have a negative  $b$  intercept value in Equation 6.1. The linear scaling factor,  $m$ , is also applied in cases where a pixel may have a large range of values while storing the values with as few bits as possible and at the same time avoiding quantization errors [22].

*PixelSpacing* is an important tag as well. It is defined as an array where the first value is the row spacing in  $mm$ , that is the spacing between the centers of adjacent rows, or vertical spacing and the second value is the column spacing in millimeters,  $mm$ , that is the spacing between the centers of adjacent columns, or horizontal spacing. Thus, in order to calculate the area in square millimeter units of each pixel, the multiplication of these two values is taken as show at Equation 6.2.

$$PixelArea = PixelSpacing[0] \times PixelSpacing[1] \quad (6.2)$$

It is worth note that some of the tools used in this dissertation to collect the learning features do not support operations over the original bone scan values range (e.g. from 0 to 4,000). Thus, in order to use these tools, the bone scan retrieved from the exams needed to be converted to 8-bit single channel images (that is its pixels content must range from 0 to 255). This converted images will be referred as 8-bit original images.

To perform this conversion, first, a clip operation was applied. In this clip operation, when there are values bigger than the specified threshold of 500 they are set to the maximum possible value in a 8-bit representation (this is 255). This threshold value was obtained by analyzing the available bone scan images histograms. Although this

clip may look like a huge loss of information, in practice, very few pixels in the exams exceed this threshold, in 42 bone scan images analyzed, 2229 pixels exceed the value of 500, resulting in the average of 53 pixels clipped per exam.

After performing the clip, a linear transformation operation is applied to the intensities of the exam so that the values present themselves at the range [0,255]. The equation applied is described below:

$$Converted = \frac{255 \text{ PixelValue}}{L} \quad (6.3)$$

where  $L=500$ .

After performing this initial calculations to retrieve DICOM information and ensure proper conversions and measurement standards, the next step deals with the collection of features from the available DICOM files that will be later used at the learning steps.

## 6.7 Used dataset description

This dissertation makes use of a set of bone scans exams already acquired in the *Antônio Pedro University Hospital* (HUAP) of UFF. The aquisition was done according to the HUAP protocol for this exam. The files arrived already stored as DICOM (that is following the specification of the DICOM standard by the hospital scan equipment). The scanner used to perform the exam was a GE Millennium® MPR Gamma Camera.

The total number of exams available is 42, they are from 21 different patients (21 bone scans on ventral and 21 on dorsal positions from each patient). From these, 8 patients have diagnoses of bone metastasis while the remainder 13 patients have healthy radio tracer distributions.

It is also important to note that although there are some patients with cranium bone scans, the focus of this dissertation will be only the full body images. The reason for this is the very small number of images of this type available in the dataset. Due to this fact, inclusion of such captures in the CAD development could not be beneficial.

It is noteworthy that as a measure to guarantee the privacy of these contributors, all the sensitive information and any other data that makes possible the identification of the individuals were turned anonymous in this study. The names of each one of these volunteer patients were DELETED. In addition to the DICOM files of the exams, the medical reports and a qualitative diagnoses of each of the findings present in those exams were also included.

## 6.8 Ground Truth

In order to effectively train the classifiers and evaluate the output, it is necessary to have a ground truth of the exams. These ground truth is considered a proper one if obtained through medical consensus of the experts participating in this project. These experts were responsible for the task of manually annotate (identifying some pixels as belonging to a metastatic point) the images by highlighting the metastatic regions). Although desirable, due to the Covid-19 pandemic, only one M.D. in fact has made the used GT until now. However, as soon as a better dataset is available, a refined version of this work will be easily possible. In this dissertation, the used GT were composed of 42 annotated images from the 42 available bone scan exams.

The results of this annotations were the digital images with metastatic regions segmented and labeled as foreground pixels. To do this operation in the images, the physicians could make use of any digital image processing software they are familiar. For instance, the web application APEER. Available at <https://www.apeer.com/annotate>. This online tool provides an easy to use interface for precisely annotate images [85]. This software provides brush and selection tools so that the only thing the user has to do for annotating is zoom-in into the digital image and draw-over the desired annotation regions. When all the desired areas are highlighted, it is possible to export an image in a .tiff format containing only the annotated regions as not null pixels. Figure 19 shows a screenshot of this application.

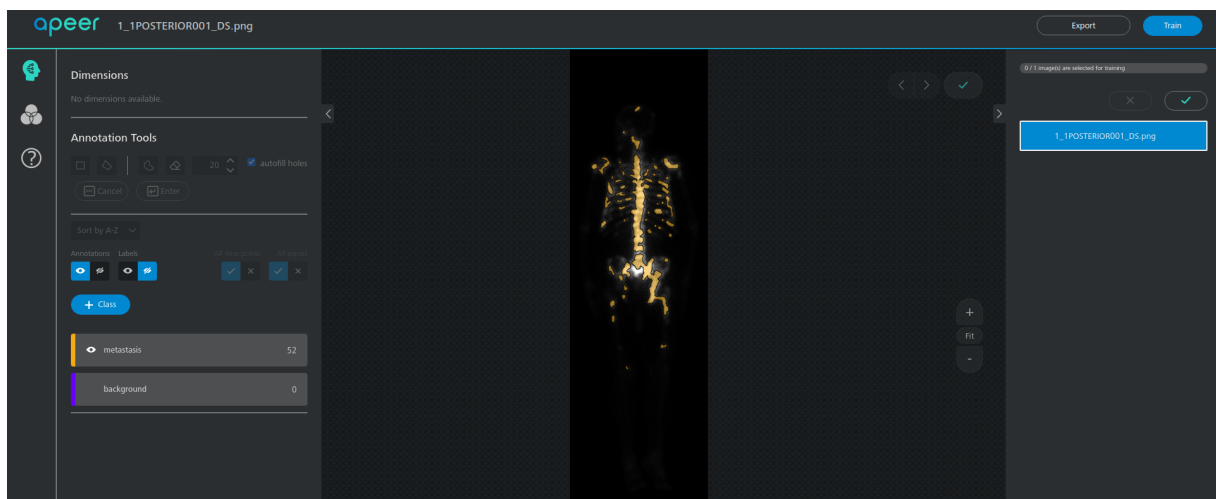


Figure 19 – Manual annotation of a bone scan image.

After the end of this manual annotation, each ground truth looks as that presented at right in Figure 20. As discussed in section 5.2 the quality of the classifiers is heavily dependent on the quality of the ground truth provided. This annotations were precisely

revised by the partner physicians, aiming to leave the least amount of relevant metastatic pixels unclassified.

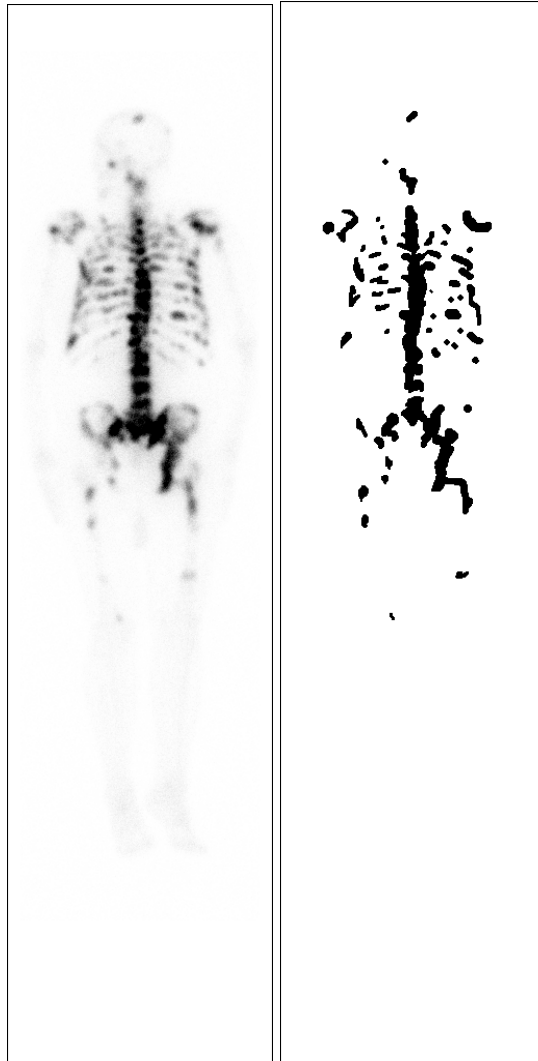


Figure 20 – Original image and ground truth provided by physicians.

## 6.9 Implemented approach assumptions

In this work it is proposed to carry out the classification at each non-null pixel of the exam, classifying it as metastatic or non-metastatic without user interaction. In other words, the idea is the design of an application full of very adequate default options. Such a decision is based in previous implementation done in the tentative of arrive to the present applications to aid in metastatic classification of bone scan exams [78]. These previous works are described in Appendix A.

The methodology applied in this dissertation is that each foreground pixel present in the bone scan DICOM file pixel data will be considered an record of the learning database and will have its own feature vector. The pixel is considered foreground when

the pixel intensity value in the original image in the range  $[0, 4,000]$  is not equal to zero. This means that all pixel with null intensity values in the original related bone scan image will not be part of the learning dataset, even if by applying some IA filters this positions receive values different of zero.

It is worth note that this strategy besides providing a much larger dataset to be trained by the classifiers, also suits well for the main goal of this dissertation that is to provide metastasis classification at a level that would be extremely difficult for the human eye, that is examining the images at pixel level.

The implemented strategy to collect the features will be compose the vectors with elements coming from the original intensities matrix, positional attributes and other several intensity values features obtained by the application of all techniques and filters described in section 4.2 with the following commented parameters and strategies.

To avoid possible confusions, it is important to point out that although some of the techniques and operations that will be applied at this methodology are commonly used for image processing tasks such as image enhancement, edge detection, image restoration and others. In the here used strategy, the results produced by these techniques will be used as features and do not have their results visualized or used by next IP steps as usually done in IP, IA and Pattern Recognition applications.

Moreover, as these features present their outputs as two dimensional arrays, in order to create a regular vector of features they are submitted to operations to be transformed in an array of one dimension with same number of elements. This is done by a flatten operation. Figure 21 shows how this operation proceeds.

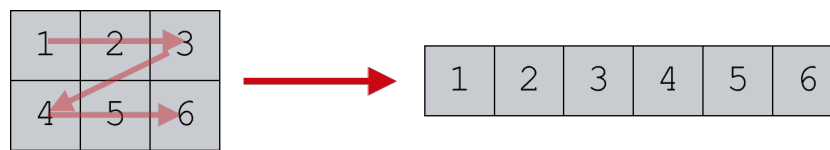


Figure 21 – Transforming 2d array into 1d one.

This operation leads to a type of re-indexing pixels used extensively in programming image processing algorithms called linear indexing, which consists of a 1-D sequence of non-negative integers based on computing offsets from coordinates  $(0, 0)$  [28]. When applied at the bone scan images and its results after filtering, the flatten process transforms the images originally stored at a two dimensional data structures  $(256 \times 1024)$  to one dimensional arrays of 262,144 elements.

## 6.10 Feature vector composition

This section presents the applied characteristics and details of how is calculated each feature of the vector used in the next learning stage, as well as their respective implementations and tuning parameters of the methodology.

### 6.10.1 Intensity of absorption

The first feature considered refers to the intensities collected by the scanner related to the radionuclides absorption by the body tissues represented in the scintigraphy exam. This information is the starting point for several of the next attributes, it is represented by the intensity of each pixel or each element of the image array.

Thus, the first feature corresponds to the result of the flatten operation when applied to the image of intensity values of the bone scan exam, resulting in a array of 262,144 elements. Each one presents content that can varies representing a integer number (intensities in the files) between 0 and 4,000, this being the minimum and maximum value of pixels from the bone scan file.

### 6.10.2 Pixel position

Another important attribute to consider is the position of the pixel to be classified. This is due to the fact that even if there are hotspots in an exam, depending on the location inside the image, they may be located in expected concentration regions, could be external objects or even irregularities in the exam. Examples of expected concentration regions in the body of the patient regions are the bladder and patient hand (where the injection of the liquid for promoting the  $\gamma$ -ray emission were done).

The position attributes considered refer to the horizontal and vertical distance of the on analysis pixel to the top left corner of the bone scan image, that is, their origin or the (0,0) image pixel.

### 6.10.3 Tag identifying pixels in pelvic region

As mentioned for the pixel position features in Subsection 6.10.2, the coordinates of the point on analysis (hots pot) is a determining factor for its classification. One of the most common cases of false positive entries in the exam are those pixels located in the pelvic region of the patient.

A greater accumulation of radionuclides in this region is an expected result since they tend to naturally concentrate in the patient's bladder region (and then eliminated i.e. excreted when urinating). Thus, would be interesting to the classifier the existence

of an attribute to represent if the pixel to be classified belongs or not to this particular body region.

In order to achieve such result, the first step is to identify the location and extent of the pelvic region in the bone scan image. Then it will be easy to recognize whether a pixel coordinate belongs (or does not) to this region. To perform this task, the implemented strategy was the use of the template matching operation to identify this body part. This operation was commented in Subsection 4.2.11.

Initially, a rectangular clipping of a bone scan image containing only this portion of the exam (i.e. the pelvic region) was considered. As this regions appearance does not change to much from one to other patient due to the low contract and resolution characteristics of this type of exams, that in this specific case could be seem as an advantage. This unique area (after some experimentation) are considered adequate for identifies if a given pixel belong to the limits of the region (it means the pelvic region bounding box) [40].

With the template matching technique the software was able to correctly identify the pelvic region in all of the 42 bone scan exams with a single pelvic image template. It is important to note that the image used as the template was manually adjusted to be the most distinct as possible from other patient's body regions. To achieve this, the horizontal limits (or horizontal extension) of the bounding box of the pattern was defined by 2 lines: The inferior one contains the beginning of the patient's legs (patients' femurs) and the upper horizontal limit is the beginning of his waist. The vertical borders (vertical extension) of the pattern are the average distance from the bones of the hands and pelvis bone limits. The bounding box of this region is used as the pattern or the window operated by normalized correlation computations (that is the matching sliding window).

In order to standardize the images and facilitate their match, some prepossessing operations were applied both in the target image from the exam and the pattern used for the template matching. To address the image intensities differences and the patients body, a simple threshold operation was applied in both the target and template image (8-bit bone scan image). The threshold value after manual experimentation and adjustments was defined to 12. With this value the software is able to remove some of the least visible pixels of the exam that are most often not related to the patient's bone. The result of this operation is also later processed with a median filter with window size 3x3 to smooth the pattern.

Figure 22 shows the image prepossessing operations performed in the images before the correlation application to find the region of best match. Figure 23 shows the template matching operation done in the template and the template window operated over it by sliding this template over the image in order to detect the best matching

position on the image on analysis.

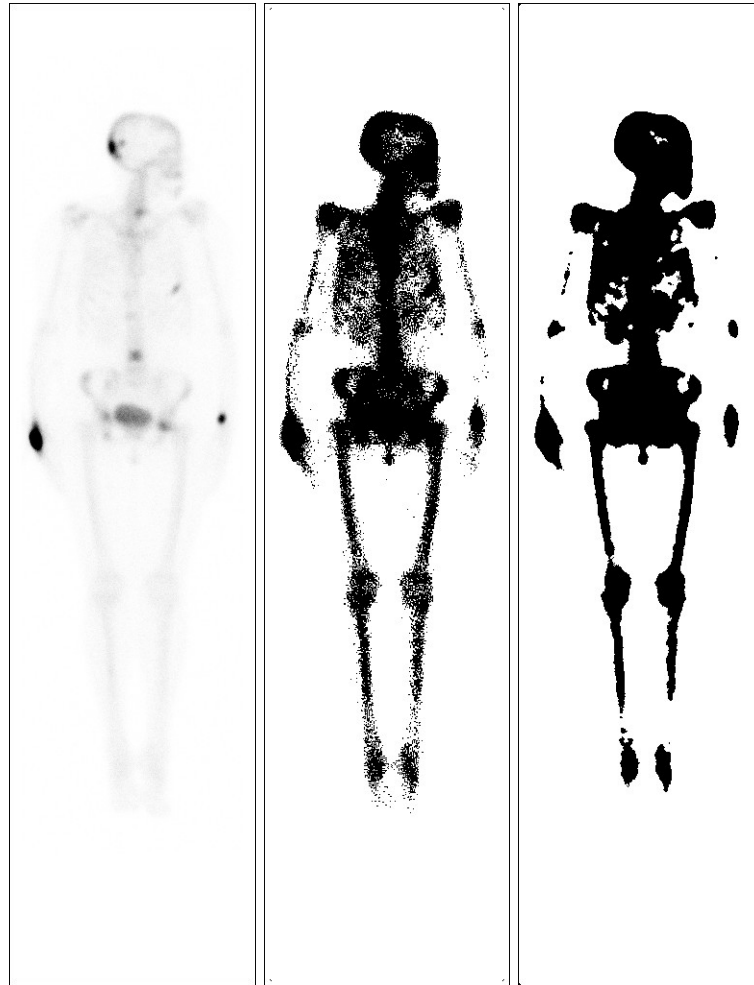


Figure 22 – Original, simple and after median filter thresholding images

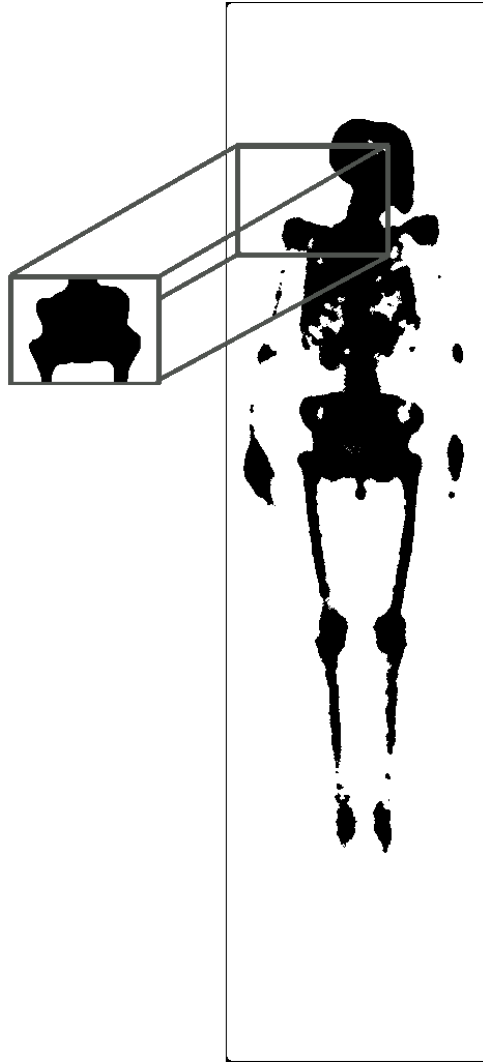


Figure 23 – Template matching operation on processed images.

After determining the location and extent of the bounding box containing the pelvic region in the image on study, for each non-null pixel of the original bone scan image, a binary attribute (i.e a **tag**) of value 1 was set for this feature if the pixel are in the bounding box region of the pelvis, or 0 otherwise.

#### 6.10.4 Histogram equalization features

The Figure 24 shows the histogram of the ventral bone scan presented at Figure 6. An obvious observation is that the number of pixels with values equal or close to ZERO (background pixels) is considerably higher than other intensities since background and non-hotspots are the predominant areas of this type of image.

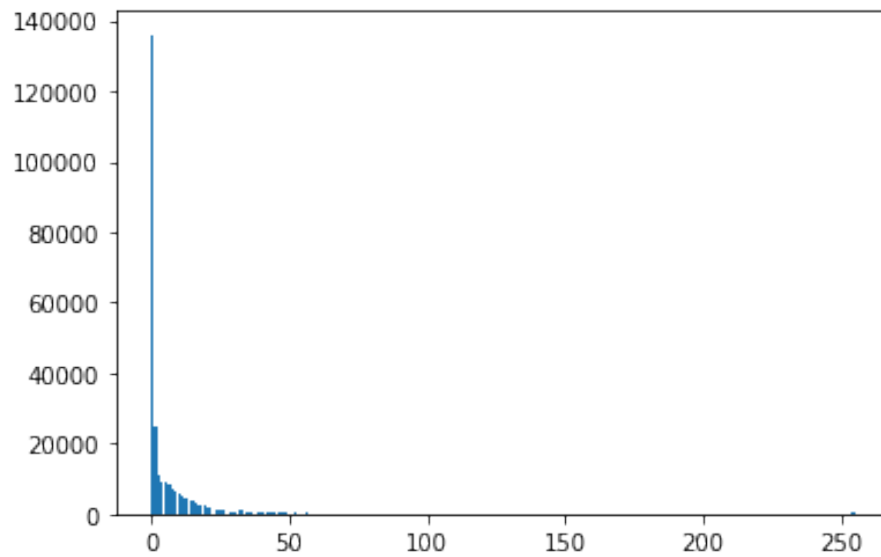


Figure 24 – Histogram of Figure 6 ventral scintigraphic presented from 0 to 255.

The image and histogram results obtained by applying the Histogram equalization technique to the 8-bit original image coming from this exam are presented at Figure 25. In this implementation the equalization algorithm was applied using the `equalizeHist` function, available in the OpenCv Python library.

As the result of Histogram equalization operation each pixel of the output image will receive a new integer gray level value of range 0 to 255. This new gray level intensity for each pixel is considered as a feature in this implementation. Meaning that the Histogram equalization feature of each pixel of the resulting image that are not background in the original image is the Histogram equalization feature of it. The flatten values of the filter result were appended to the features vector of each non zero pixel.

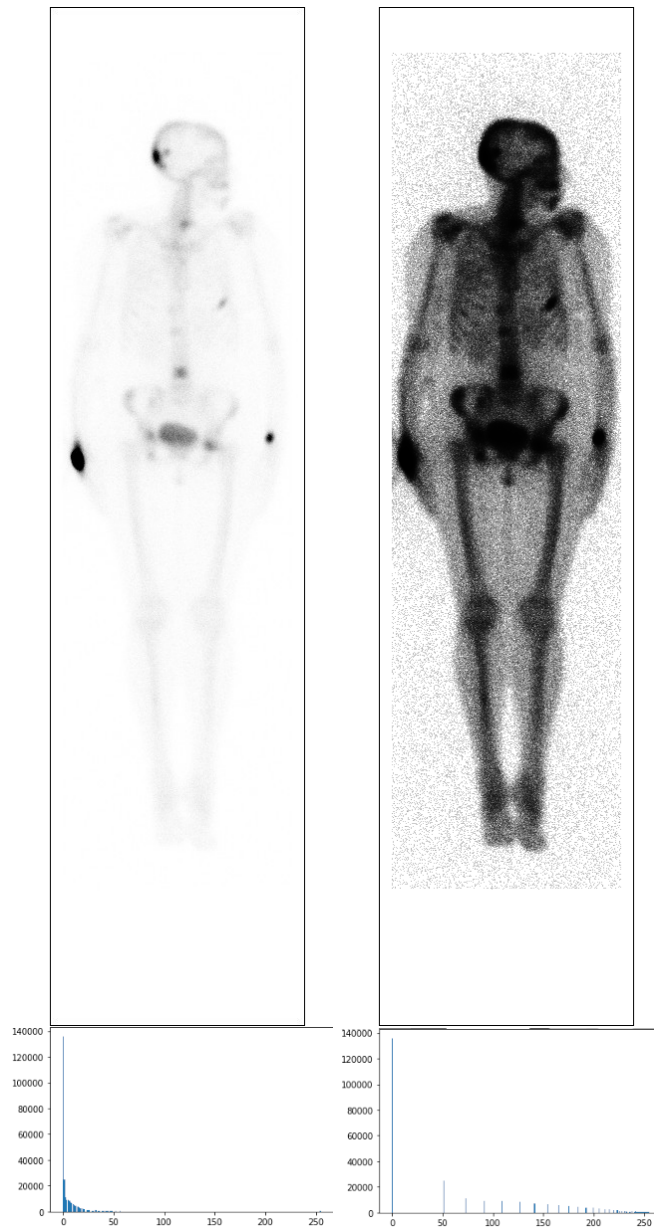


Figure 25 – Original x Equalization images and their histograms.

#### 6.10.5 CLAHE feature

In this dissertation, the implementation used to the CLAHE algorithm was the one provided by the Python library OpenCv. It is configurable by the following parameters: `tileGridSize` and `clipLimit`. The first one, `tileGridSize`, indicates which is the number of tiles in the row and column, that is used while the image is divided into tiles for applying CLAHE. The second, `clipLimit`, is the parameter of the threshold for contrast limiting, as mentioned in section 4.2.2.

Figure 26 shows on the right image the CLAHE result applied on the left image, and the histogram distribution obtained by the application of this technique with `tileGridSize` equals 2 and `tileGridSize` of 8x8 pixels, over the 8-bit image.

As with the global equalization, the result of CLAHE operation is to redefine for each pixel of the output image new gray level values. This new gray level for each pixel is considered as feature in the idea presented in this implementation. That is the CLAHE feature of each body pixel of the resulting image, that are not background in the original image, is the CLAHE feature of it. Then, the obtained two dimensional result was flattened and appended as a feature to the not null pixels dataset of the original exams, that is the original background pixels of the exams.

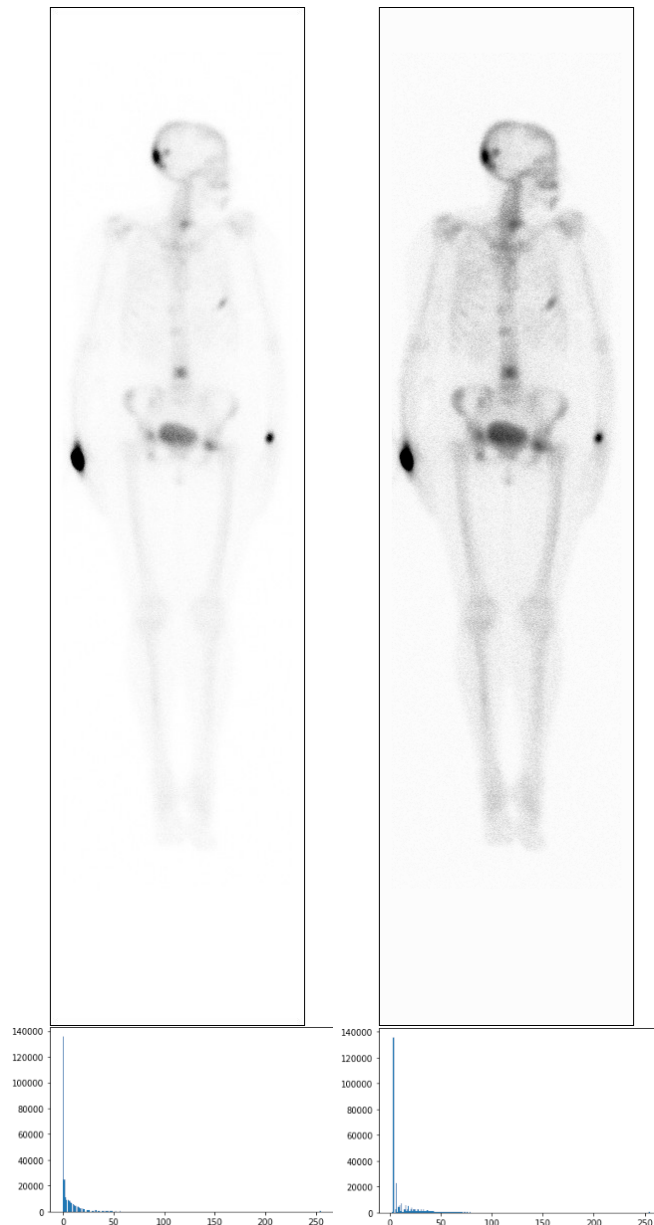


Figure 26 – Original x CLAHE images and their histograms.

It is important to note that even if the image resulting from the CLAHE operation in the original 8-bit image of the exam does not have pixels with an intensity value equal to 0. Only pixels whose values are not null in the original image will be part of the

learning dataset.

#### 6.10.6 Otsu feature

By applying the Otsu threshold technique to the bone scan images they are automatically transformed into a set of binary images. The implementation of this algorithm used in this dissertation was the one available in the Python library OpenCv. Comparing such results with the ground truth, it is possible to state, at least visually, that the Otsu threshold effectively segments some of the most visible high absorption regions present in the exam (therefore possible metastases regions) as shown in Figure 27 (In this Figure, the intensity value of 1 was remapped to 255 in order to become visible).

Otsu feature for each pixel is then a tag with 0 or 1, accordingly the results of Otsu threshold to the bone scan images. The flatten of the image results from this filter provides a binary feature to our problem, assigning these values (0 or 1) to each pixels.

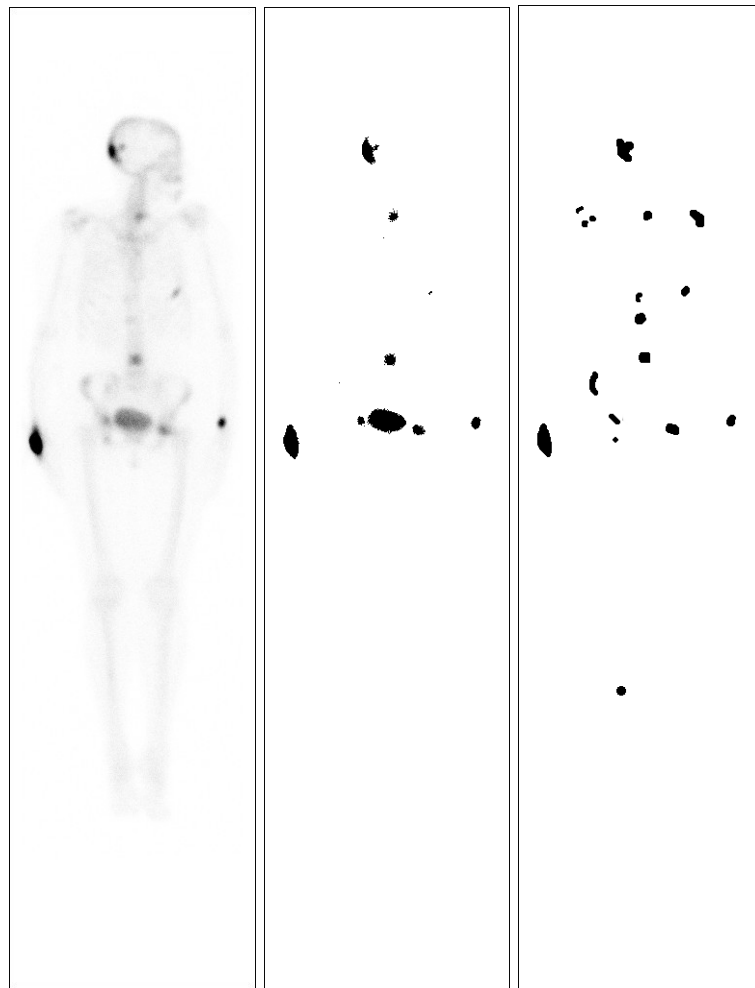


Figure 27 – Original, Otsu result and ground truth images of the same exam.

### 6.10.7 Gaussian features

The implementation of this feature is provided by the function `gaussian_filter` of the Python library Scipy. The most important parameter available for definition of the result is the standard deviation,  $\sigma$ , of the Gaussian equation as commented in section 4.2.4.

To compute the features using this filtering technique, firstly the filtering process of Gaussian filter over the 8-bit original bone scan images are performed. These are applied with the standard deviation parameters of ( $\sigma =$ ) 3, 5 and 7. Values greater than 7 for the standard deviation were tested, but they resulted in a heavy blurring effect that was considered undesirable. Examples of these filtering results can be seen in Figure 28.

After the image filtering, the new integer values of each pixel of the image results are assigned as features to the original bone scan foreground pixels(i.e. for pixels do not belonging to the background).

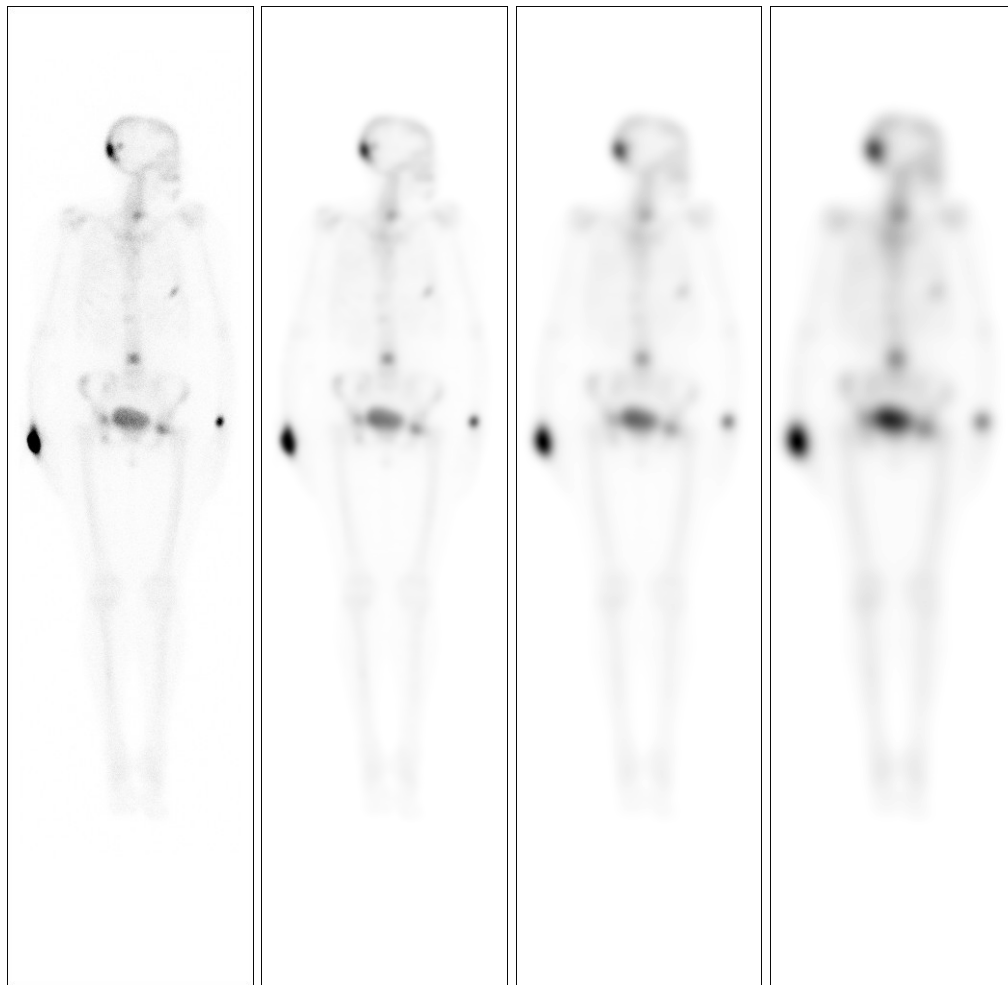


Figure 28 – Original and Gaussian filtering with standard deviation 3, 5 and 7.

### 6.10.8 Median features

Some filters are performed over the space domain by doing the convolution using a kernel of a previous defined size. Others filters like Median uses a previous defined window size to operate for other type of calculation. To select the window size for the median filter technique, some evaluations of the average size (in pixels) of the bounding box of some of the regions already known as metastatic, that are present in the available ground truth files are performed. The more frequent sizes of regions from different exams are 7x6, 7x7, 8x7, 10x8, and 10x9; leading to an approximate number of 8 pixels as average extent of these regions.

The median features in the vector of attributes are the integer results (from 0 to 255) of the median computation using the function *median\_filter*, available at Scipy Python library, on square regions of 3x3 and 5x5 size over the 8-bit original scintigraphy images of non background pixels. Samples of this results are presented as images on Figure 29.

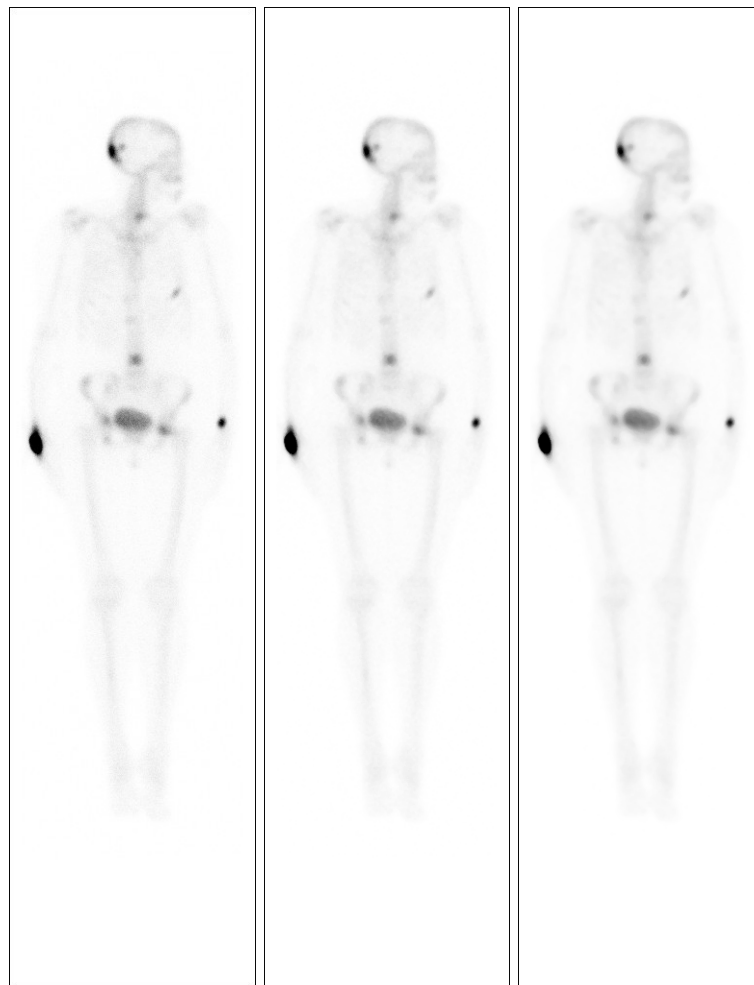


Figure 29 – Original and Median filtering with windows of 3x3 and 5x5.

### 6.10.9 Entropy feature

The Entropy feature used in this work is based on the function *entropy* available in the Python library Skimage. The configurable parameters of this function corresponds to the choice of the shape and size of the region where the entropy will be calculated. The image characteristics used for Entropy computation are collected by the application of this function on a circular region of radius 3 around the pixel that is operated by the Entropy filter.

The application of the entropy filter was performed directly on the original 8-bits images and it's resulting floating point values were added to the set of features for the originally not null pixels. A sample of this filter in a bone scan image is presented at Figure 30.

As this filter considerably reduces the scale of the intensities range, due to the log characteristics, in order to actually visualize the image result in this dissertation, To this particular feature, Figure 30 corresponds to the transformation result of the floating points values to the 0-255 range by applying the same equation described at 6.4 but with the denominator being the maximum value obtained to this feature in this particular bone scan image (i.e.  $L = 4.85$ ).

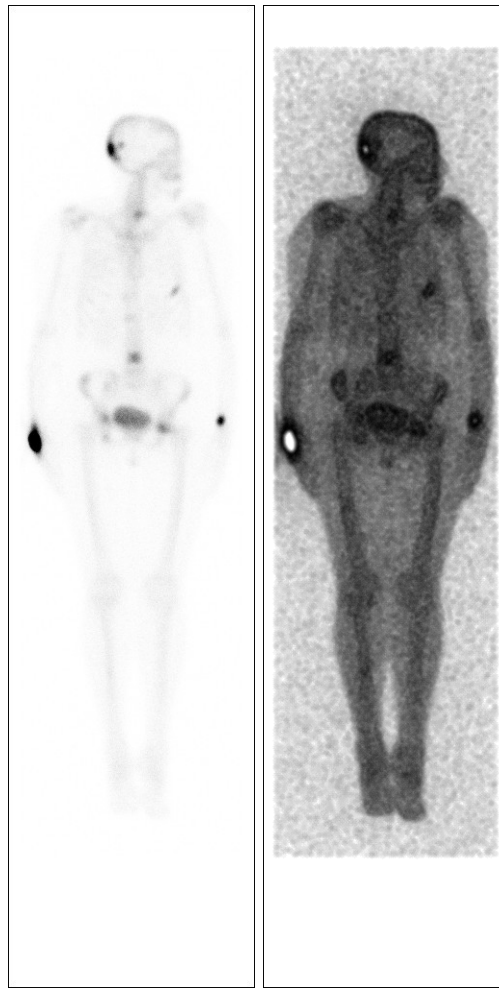


Figure 30 – Original image and it after Entropy filtering.

#### 6.10.10 Roberts feature

In this feature, the Robert's cross operator implementation provided by the Python library Skimage [4.20](#) was applied on the 8-bit original images. It results to images like the one presented at Figure [31](#). An observation of this feature is that the results obtained by this filter were usually low contrast images.

The integer values obtained at the result of this filter for each non background (in the original image) are the named Roberts feature element to be included in the future feature vector of each pixel.

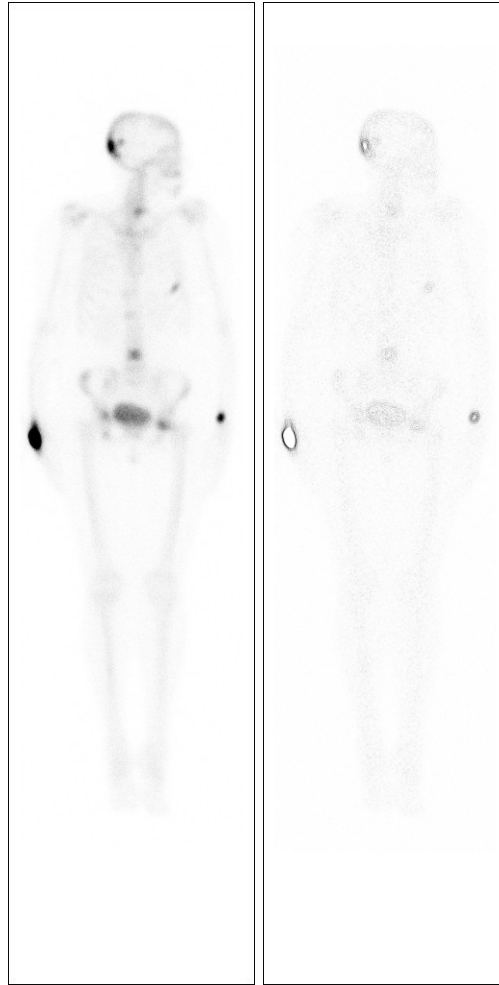


Figure 31 – Original and after Robert filter images.

#### 6.10.11 Sobel feature

For computation of the Sobel feature, the implementation of the filter available in the Skimage Python library is used. This operation is applied over the 8-bit original bone scan images resulting in a integer number for the pixel in the center of the 3x3 region operated by the Sobel masks of Equation 4.21. The results were directly defined as an element of the features vector. Figure 32 shows a sample of a result obtained by this filter operated on a bone scan image.

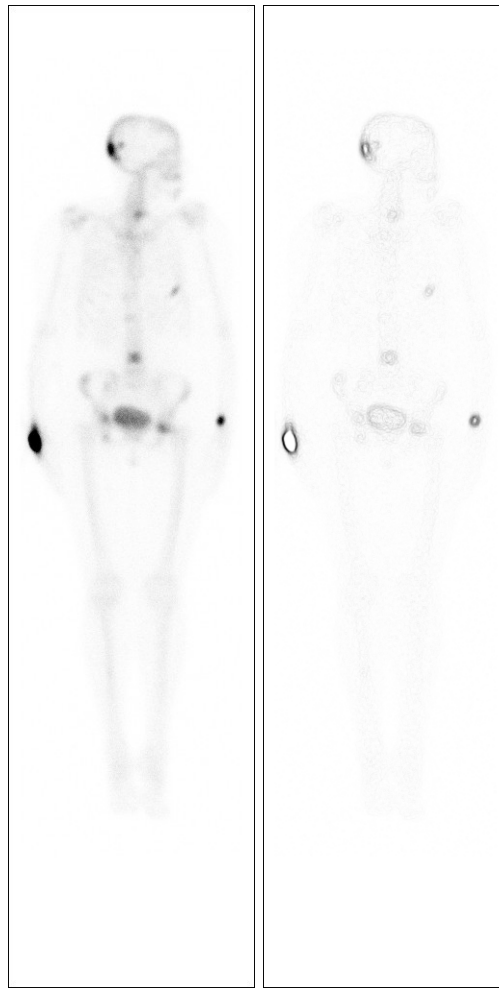


Figure 32 – Original and after Sobel filter images.

#### 6.10.12 Feature of Variance

The variance filtering was computed with the function `var` of the Numpy Python library. This function was executed with a sample window of size of 3x3 pixels. Figure 33 presents an example of an image and it after variance filtering. The contribution of this filter to the collection of elements of the feature vector corresponds to the integer number obtained as the result of applying the variance calculation to the 8-bit original bone scintigraphy images.

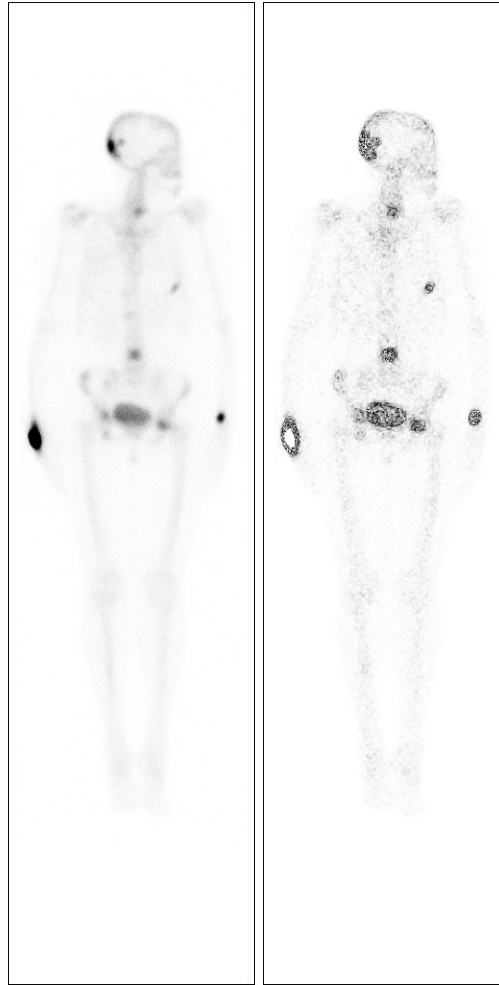


Figure 33 – Original and after Variance filter images.

#### 6.10.13 Gabor features

By varying the parameters present in the equation 4.23 it is possible to obtain different convolution functions capable of locating and selecting specific sets of information according to angle of orientation,  $\lambda$  wavelength of the harmonic or their frequency and desired distribution (defined by  $\sigma$ ). A common approach in the field of image analysis and also applied to the problem described in this dissertation is the use of a set of Gabor filters or a bank of filters. This set contains a considerable amount of kernels obtained by the combination of the adequate parameters of equation 4.23.

In this dissertation is used the implementation of this filter available at the OpenCv Python library, where the parameters of this function are:

- `ksize`: it specifies the edge size of the square kernel of the filter computed.
- `sigma`: it represents the programmer defined standard deviation of the Gaussian envelope to be used in the equation 4.23.

- `theta`: it represents the programmer desired value for  $\theta$  in Equation 4.26, that is the angle of orientation of the normal to the parallel stripes of the output Gabor function.
- `lambda`: it defines the wavelength  $\lambda$  of the sinusoidal factor of equation 4.23.
- `gamma`: it defines the spatial aspect ratio between  $\sigma_x$  and  $\sigma_y$  in Equation 4.23.
- `psi`: last parameter, it defines the phase offset in the equation. This parameter is automatically multiplied by  $\pi$  by the tool, meaning that the value of 1 for this parameter results in an angle of  $180^\circ$  4.23.

First, a study was made for definition of an adequate range for Gabor parameters, that aid to identify those that somehow improve the visual contrast in the images or highlight possible hotspot regions. The gamma parameter, referring to the aspect ratio of the Gaussian kernel to be used (see Equation 4.2.4), was tested with 2 values 0.01 (almost rectilinear) and 1 (circular or  $\sigma_x = \sigma_y$ ). The study has observed the results achieved with 3 angles settled to the  $\theta$  parameter: 0,  $\frac{\pi}{4}$  and  $\frac{\pi}{2}$ . The standard deviation parameter,  $\sigma$  were also varied to be equal to 1 or 5. The kernel size used was tested by using 2 values: 3 and 7 and the phase offset parameter  $\psi$  with values 0 and 1.

As for the  $\lambda$  parameter, the tests indicated that lower values than  $\frac{\pi}{4}$  in the Gabor executions end up overlooking small features at the image. Thus, this parameter was fixed in this value for all filter executions.

Figure 34 shows results of Gabor filtering with kernels 7x7 and parameters:  $\theta = 0^\circ$ ,  $\lambda = \pi/4$ ,  $\sigma = 1$ ,  $\gamma = 0.01$ ,  $\psi = 1$ , and  $\theta = 90^\circ$ ,  $\lambda = \pi/4$ ,  $\sigma = 5$ ,  $\gamma = 1$ ,  $\psi = 1$ .

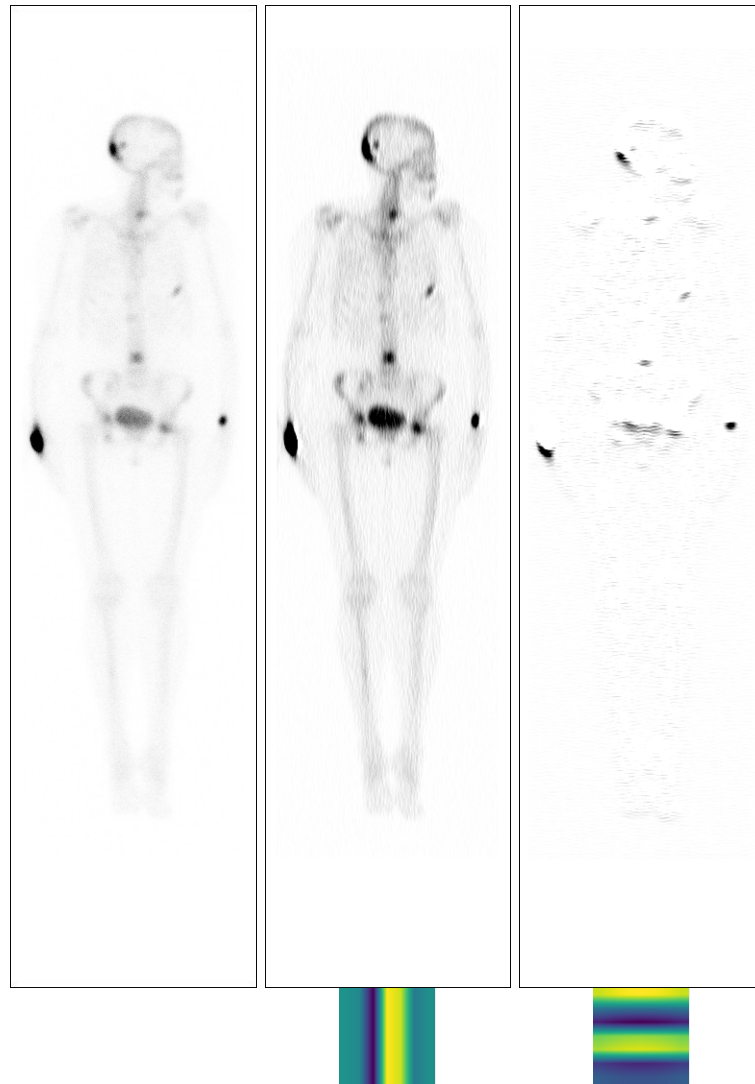


Figure 34 – Original and Gabor filter result with angles of 0 and 90 degrees.

The Gabor functions with these parameters produced 48 kernels. They were operated by convolution with the 8-bit original exam images and the result of them composed the Gabor features of each non back ground pixel. At total, 48 Gabor features were included in the set of each foreground pixel entry.

This set of 48 integer numbers were the last entries for the composed features vector, thus, begins the process of building the machine learning models which will be described at the next sections.

## 6.11 Dataset processing and number of records

For each of the exams received, the operations of reading, converting and normalizing the DICOM file and its metadata were executed. After the processing, this exams went to the computation of the features vector of each bone scan exam through

the operations described in previous section 6.10.

As already mentioned many times, each originally non-null pixel (belonging to the foreground) presents in the exam has its feature vector computed. At the end of the processing task of all bone scan exams provided, the informations of the dataset are transformed to an array formed by 2,512,386 lines and 65 columns. From which 64 of the columns are independent attributes (the elements of the feature vector of each pixel) and 1 is a dependent attribute. More specifically, these 65 elements are:

- 1 integer corresponding to intensity of absorption value on this pixel;
- 2 integers referring to vertical and horizontal pixel position in the image;
- 1 binary corresponding to be or not in the pelvic region of the patient;
- 3 integers from Gaussian filters with  $\sigma = 3, 5$  and  $7$ ;
- 2 integers from median of a  $3 \times 3$  and  $5 \times 5$  neighborhood;
- 1 integer from global histogram equalization;
- 1 integer from local adaptive histogram equalization;
- 1 binary from Otsu thresholding;
- 1 integer from Roberts filter;
- 1 integer from Sobel filter;
- 1 integer from entropy filter;
- 1 integer from variance filter;
- 48 integers from Gabor filters;
- 1 binary attribute representing metastasis (1) or do not (0).

The position 65 of the feature vector is a tag. It is included in order to simplify the notation. Using it, the possible diagnoses were mapped to binary class, assuming the value 1 for a metastasis diagnosis and 0 for non-metastatic. Regarding the class distribution of the dataset, 2,457,703 entries belong to the healthy bone class and 54,683 belong to the metastasis class. This unbalanced class distribution is expected since there are very few cases where the affected area in the bone scintigraphy is greater than the normal bone pixels count. Figure 35 shows a chart of the classes distribution.

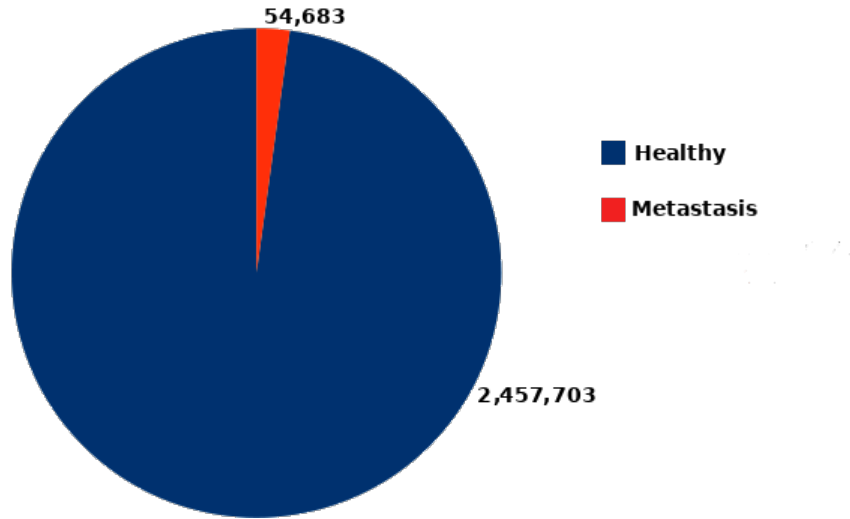


Figure 35 – Classes distribution of the generated dataset.

#### 6.11.1 Data normalization

After the definition of all vectors, the process moves to the ML phase. However, before using these data to train the models, a normalization of the data is carried out, so that they are transformed to floating point values in range [0, 1]. The goal of this normalization is to put all of them in a common scale, in order do not introduce bias due to possible great differences in the ranges of their values.

In order to perform this operation, the maximum and minimum possible values of each type of features (each column of the database array) were re-scaled according to the normalization equation 6.4, this equation is applied across the entire dataset (now represented as 2d array and subsequently as a 1d):

$$el_{norm} = \frac{(el - el_{min})}{(el_{max} - el_{min})} \quad (6.4)$$

where:  $el_{norm}$  is the new normalized element of a given feature type (or column),  $el$  is its computed value,  $el_{min}$  and  $el_{max}$  are the range of the values of this feature type (or column).

To normalize the first feature, that is the intensity of absorption, the minimum and maximum values were considered to be 0 and 4,000 respectively.

To normalize the pixel position features, image resolution is considered the maximum (1023 to height and 255 to width) and the minimum value used was 0.

For other features with unknown ranges, this normalization parameters were collected directly from the generated dataset by going through the list of possible values on each column and taking their maximum and minimum values.

### 6.11.2 Dataset reduction

Once the set was normalized, the PCA technique was used to reduce the number of features. The PCA algorithm applied was the one available in the Python library Sklearn. The most important 25 components of the feature vector are the number of components that remains.

In order to find this number of principal components to retain, the PCA technique was first fit with the all the 64 initial independent variables count, then, the results in terms of **Explained variance** were observed. The values of **Explained variance** against the **number of components** on the feature vector are plotted in the graph as showed in Figure 36. By analyzing this plot, it is possible to note that after 25 features, very little **Explained variance** is gained by including additional features, making this number a good choice to the number of components to be retained. After this analysis, the PCA algorithm was applied to the obtained dataset, resulting in a new dataset composed of 2,512,386 records, 25 features and 1 target attribute.

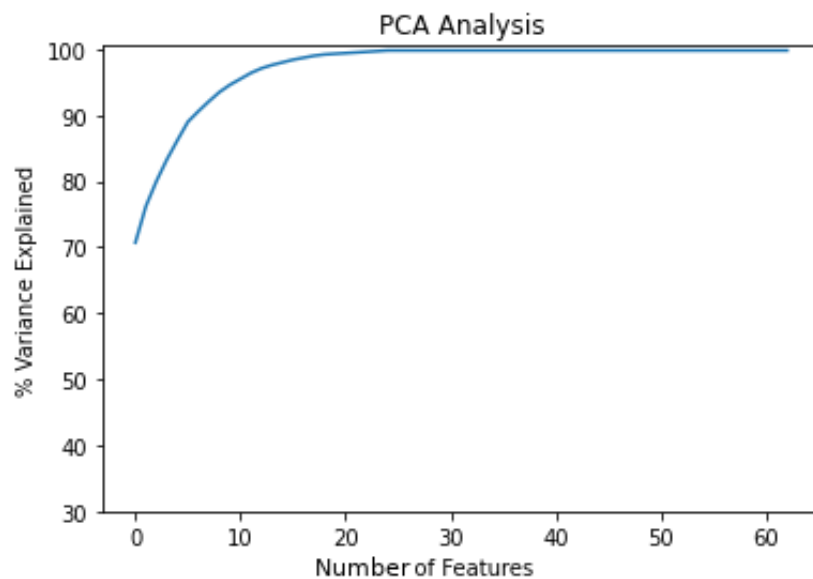


Figure 36 – PCA analysis plot.

## 6.12 Training results

After obtaining the final database, the next step consists in actually implementing, training and evaluating the obtained results from each classifier. At the end of this comparison, the technique with best overall results and performance will be applied as the CADx classifier.

In order to verify the performance of the selected supervised learning techniques, the train-test split strategy used was K-Fold Cross Validation with a parameter of  $k = 5$ .

Thus, 5 sets of size 502,477 were generated, each one of these alternating between training and testing roles.

The implementation used for the the Gaussian Naive Bayes (NB) classifier was the one provided by the function `GaussianNB` of Sklearn Python library. There were no parameters to be adjusted in this function.

To the implementation the Random Forest (RF) algorithm, was used the implementation available at the Python library Sklearn. The adjusted parameter related to the technique applied to dividing and building the trees was the *Gini Index*. The numbers of estimators parameter was varied between 10, 20, 30, 45 and 50 voting trees. After 45 estimators was noticed the beginning of a decay in the precision of the algorithm.

For SVM training, this work applied the implementation provided by the Sklearn Python library. No iteration limits were used in the executions in order to enable a longer learning period and without premature interruptions. The *Kernel trick function* used was RBF, defined by the equation:

$$RBF(V) = e^{(-\gamma_p(\|V-V'\|^2)} \quad (6.5)$$

where  $V$  and  $V'$  are entries from a subset of the training data and  $\gamma_p$  is a parameter which must be a positive value. For executions, the value of  $\gamma_p$  was set using the following function:

$$\gamma_p = 1/(25 \times \sigma^2) \quad (6.6)$$

where square sigma ( $\sigma^2$ ) represents the variance of a random subset of the input set and 25 is the number of attributes (found in the last subsection).

To evaluate the generalization capacities of the Multi Layer Perceptron (MLP) technique in this problem, it was tested with 10 and 100 hidden layers. In this dissertation was used the implementation provided by the Sklearn Python library and the selected activation function was ReLU.

Table 3 presents the prediction results obtained by cross-validation for Recall, Precision and AUC for all different executions of each algorithm on the generated database.

Table 3 – Cross-Validation training results

Algorithm	Recall (%)	Precision (%)	AUC
Gaussian NB	90.04	61.22	0.97
RF 10	74.35	83.59	0.93
RF 20	75.97	84.09	0.95
RF 30	76.47	83.63	0.96
RF 45	77.34	83.54	0.96
RF 50	77.38	83.03	0.96
SVM	75.03	83.48	0.97
MLP 10	78.85	84.07	0.98
MLP 100	76.45	83.26	0.98

### 6.12.1 Discussion about decisions made

After the execution and evaluation of the described algorithms, the task of comparing the results obtained begins. This has the objective of determining which of the techniques will be made available for use by the professionals in the implemented solution.

The Naive Bayes algorithm presented good Recall results and the shortest training time among the solutions. However, the low precision of 61% makes it unfeasible when compared to other solutions.

The other techniques showed promising results in relation to the analyzed parameters (that is Recall, Precision and AUC): The random forest (RF) and multilevel neural network (MLP) algorithms obtained Recalls above 75% of those from Support Vector Machine (SVM). The SVM obtained a precision competitive in relation to those two techniques.

Comparing the overall performance, the technique selected for the application was the Random Forest with 45 estimators. Because It is simplest, has a more transparent learning then the others (i.e. this method provides a complete read on it's classification and what decisions led to the class). Although the MLP model with 10 hidden layers had better numerical results, the recall and the precision between this and the best execution of Random Forest can be considered almost a tie, since without the use of a rounding criteria, their differences were 0.51% in recall and 0.53% in precision. Moreover , the results provides by the RF complete exam classification were also more similar to the ground truth images provided.

It is noteworthy to say, that the other techniques were not discarded from the project. Because it is in a constant training process through the variation of new parameters or entry of new samples into the training dataset. They continues in the implementation

code, but bypassed in this version, being easily included again and tested about their possibilities of obtaining better results when a larger training base became available or more hyper parameters could be considered.

### 6.13 Computation of the metastatic area

After classifying each foreground pixel present in the bone scan exam image with the selected classifier, three quantitative elements are calculated. These elements are: the amount of **total area** classified in the exam in  $mm^2$  (this being the area occupied by patient body in the image); the **amount of area affected by metastases**, also in  $mm^2$  and the **percentage** of area affected in relation to the total area classified.

The total and affected area are obtained respectively by counting the non-zero pixels present in the scan and by counting the pixels whose classification result was metastatic. To convert these counts to  $mm^2$ , the results are multiplied by the `pixelSpacing` value obtained from the DICOM metadata of the exam, as previously mentioned in section 6.6. The percentage between affected area and total area analyzed is obtained by simply dividing these two values followed by the multiplication of this result by 100.

### 6.14 Implementation output

At the end of the image processing tasks the metastatic areas detected by the application are presented to the user in various ways. The implemented solution is also responsible for producing a report containing the quantitative evaluations of the findings (described in last section).

Figure 37 shows the system screen where users can visualize the results of the processed exam or export the report (that appears in the bottom). The user can also alternate (using the radio button on the right of the screen) between white foreground and black background or a black foreground and white background. These possible combinations for the application screen are shown on Figure 37 and 38, respectively.

This page presents to the users the following combination of images: original image; edge (i.e. binary image that shows only the borders of detected metastasis); segmented or binary image classification result (i.e. the result of the automated classification of each pixel in the exam, in this image pixels classified as metastasis are show as foreground); overlay (i.e. the original image with the bone metastasis detection in red color). The user can alternate among the visualization options by check-in one of the three squares on the right menu.

Home

Logout

Exam: Anterior1

Back

Segmentation & Classification

Classifier probability threshold

Lower probability allows...

0 1

Visualization contrast

☐
☒

Visualizations:

☐ Original + segmented + overlay
 ☐ Original + segmented + edged
 ☒ Original + edged + overlay

Report

Jhon doe - Paciente4

Images projections obtained after intravenous administration of the radiotracer after 5 hours show focal areas of increased radiotracer fixation:

Result image: [https://api-him.dig.ay.gov.br/api/v1/segmentation/513/4420\\_58900416/visualize/170.png](#)

Edge image: [https://api-him.dig.ay.gov.br/api/v1/segmentation/513/4420\\_58900416/visualize/170.png](#)

Overlay image: [https://api-him.dig.ay.gov.br/api/v1/segmentation/513/4420\\_58900416/visualize/170.png](#)

SUMMARY:

Total classified area: 642750.60 mm<sup>2</sup>

Total affected area: 10695.31 mm<sup>2</sup>

Affected area ratio: 2%

Export as txt

Figure 37 – Processing results with dark background.

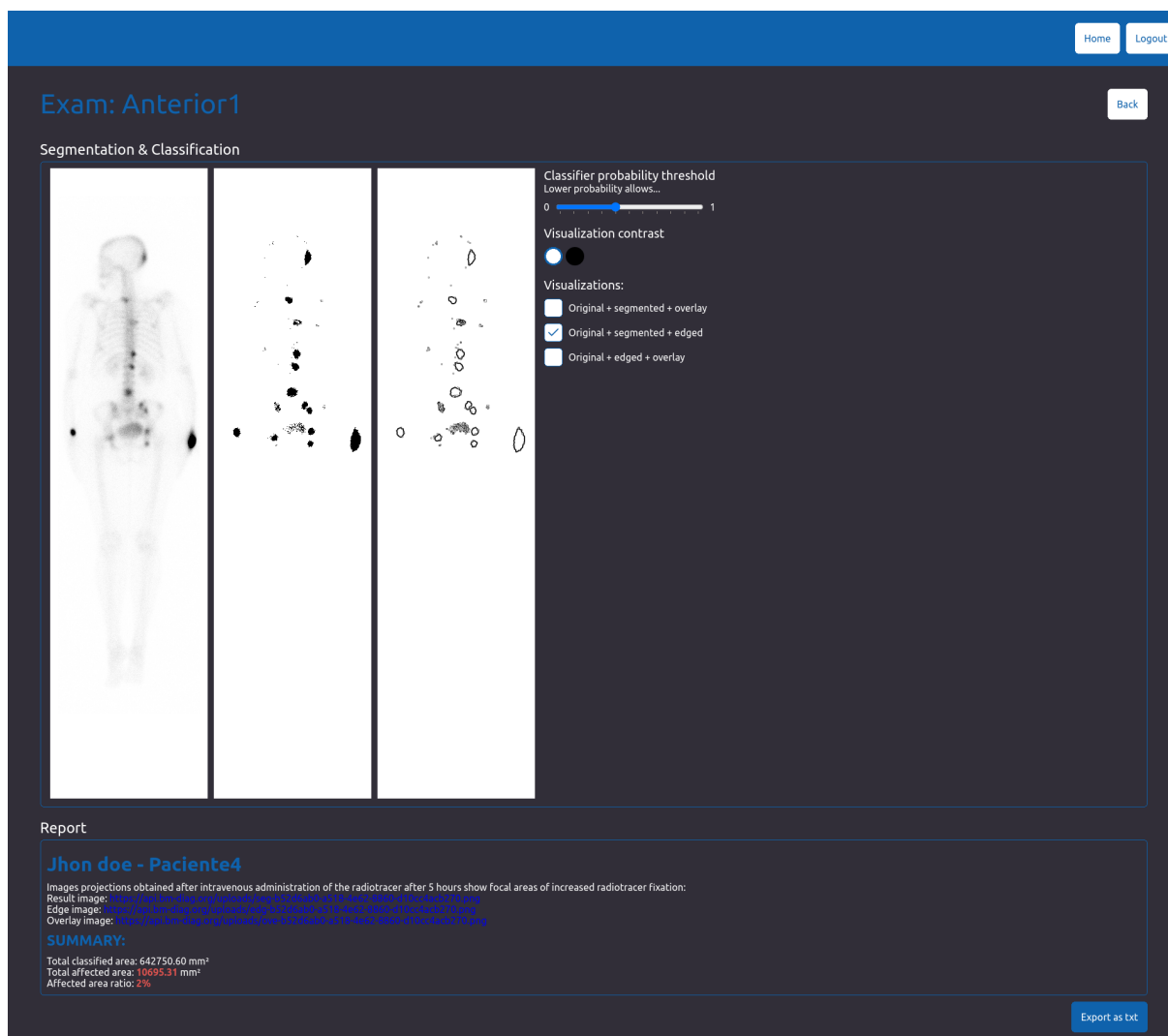


Figure 38 – Processing results with white background.

Another functionality presented to users in this page deals with the possibility of adjusting the desired confidence level for the classification. This parameter defines the degree of certainty required by the classifier for a pixel to be considered a metastasis. This value is adjusted through a sliding input and can obtain a maximum value of 1 and a minimum of 0. Any adjustment by the user of this input automatically triggers an action responsible for classifying the exam pixels again, but with this new parameter adjusted.

Figure 39 presents an example of a original bone scan image; the application output image with the pixels classified as metastases presented as foreground pixels and the ground truth image provided by the professionals for this bone scan.

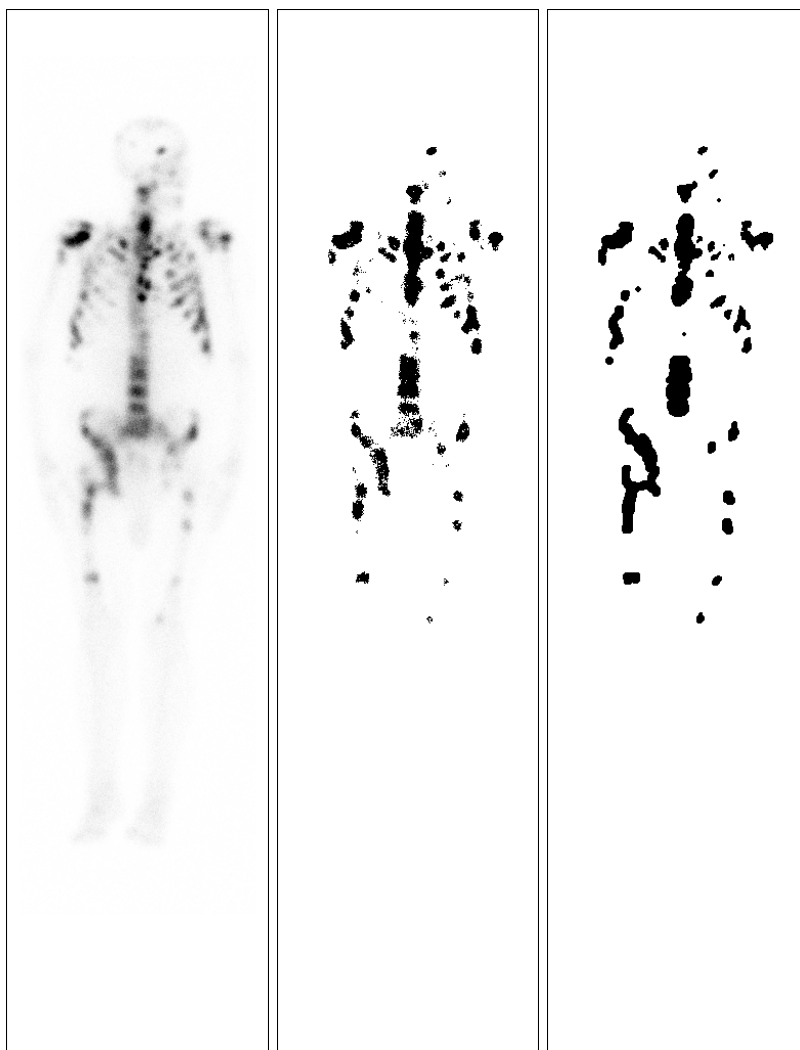


Figure 39 – Original, automatic classification output and ground truth.

To enable the user to print and share the application report, the solution produces and makes available for download a text file. This file is organized according to standards established by the medical community. After the details about the patient there are 2 sections: RESULTS and SUMMARY. The RESULTS presents details of the exams and the public uniform resource locators (URLs) to the segmented, edge images and false colors obtained by the classification process. The SUMMARY presents a quantitative evaluations of the findings. Figure 40 shows a sample of a report text file.

<p>Jhon Doe - Paciente4</p> <p>RESULTS:</p> <p>Whole-body images in the anterior and posterior projections obtained 3 hours after intravenous administration of the radiotracer show focal areas of increased radiotracer fixation:</p> <p>* Result image: <a href="https://api.bm-diag.org/uploads/seg-1e1314b8-92cb-4a97-a37b-ee7e0808bbd8.png">https://api.bm-diag.org/uploads/seg-1e1314b8-92cb-4a97-a37b-ee7e0808bbd8.png</a> * Edge image: <a href="https://api.bm-diag.org/uploads/edg-1e1314b8-92cb-4a97-a37b-ee7e0808bbd8.png">https://api.bm-diag.org/uploads/edg-1e1314b8-92cb-4a97-a37b-ee7e0808bbd8.png</a> * Overlay image: <a href="https://api.bm-diag.org/uploads/ove-1e1314b8-92cb-4a97-a37b-ee7e0808bbd8.png">https://api.bm-diag.org/uploads/ove-1e1314b8-92cb-4a97-a37b-ee7e0808bbd8.png</a></p> <p>SUMMARY:</p> <p>* Affected area ratio: 1.31% * Total affected area: 8361.14 mm<sup>2</sup> * Total classified area: 642750.6 mm<sup>2</sup></p>
--

Figure 40 – Example of an exported report.

After all comments about the implementation of the CADe and CADx it is important to discuss the suitability and relevance of the developed solution and draw up concrete plans to overcome weaknesses or improve existing strengths and functionalities. These are the subject of next chapter.

## 7 Results and conclusions

This chapter presents comparison with results obtained by other works on bone metastasis, the opinion of the final users about the implemented software, final conclusions and future works.

### 7.1 Comparison with related works and previous versions

Although the work developed in this dissertation seems to be original about pixel level classification of bone metastasis in scintigraphy exams, it is worth providing an overall comparison with related works in order to understand the relevance and impact obtained by this realization.

Among related works discussed in Chapter 3, there is a previous version of this dissertation presented in an international conference that is summarized in Appendix A. Table 4 compares for each work: the number of patients used, primary goals, and results obtained in terms of accuracy, precision, recall, ROC, etc.

Table 4 – Comparing related works

Work	Year	Patients	Primary goal	Results
EXINI v1[4]	2006	200	Metastasis present or absent	sensitivity=90%; specificity=74%
EXINI v2[5]	2009	810	Metastasis present or absent	sensitivity=90%; specificity=89%
BONENAVI v1[6]	2012	904	Metastasis present or absent	AUC=0.91
BONENAVI v2[7]	2013	1532	Metastasis present or absent	AUC=0.94; sensitivity=90%; specificity=84%
CADBOSS[94]	2016	60	Metastasis present or absent	accuracy=92%; sensitivity=94%; specificity=87%
Dang [8]	2016	2164	Metastatic / non-metastatic hotspots classification	sensitivity=98%; specificity=65%
Belcher [9]	2017	2164	Metastatic / non-metastatic hotspots classification	AUC=0.94
Larisa group [10]	2020	507	Metastasis present or absent	recall=98%; accuracy=97%; precision=95%
BS: Previous version [78]	2021	10	Metastatic / non-metastatic hotspots classification	accuracy=86%; precision=77%; recall=87%; AUC=0.87
BS: Actual version	2022	21	Metastatic / non-metastatic pixel classification	precision=77%; recall=84%; AUC=0.96

Over this 16 years, 8 methodologies and applications were implemented to deal with the difficulties of diagnosing bone metastasis on scintigraphic exams. Although the goal of classification and quantification form may vary between the different approaches, the results obtained by this dissertation presents great relevance taking care the Receiver Operation Characteristic, specially by the considerably lower number of patients and samples used in comparison with the other related works.

When comparing the results obtained by the work carried out in this dissertation and the first versions of this work, the new methodology contributed to a more solid result as the number of samples available for training and testing increased from 77 hotspots to 2,512,386 pixels.

## 7.2 User Survey

As the developed system aims to offer an intuitive interface for physicians and specialized professionals to perform analysis of bone metastasis exams in DICOM format. A survey was carried out in order to evaluate how satisfactory the developed solution is from the perspective of users (professionals who will actually use the system) in their daily routine.

The survey is composed by a questionnaire developed with the Google Forms technology and had the answers of 3 physicians invited to test the implemented application.

There are 12 questions in the questionnaire. These were elaborated in a form of 8 of them requiring mandatory answers, while the rest of the questions were optional. These free questions aim to identify errors found in the system and possible improvements from the reported suggestions.

Throughout the next 3 subsections, the questions of the survey and the general result obtained at the end of the survey analysis from collected answers will be presented.

### 7.2.1 Evaluation of patient management of functionalities

The objective of this subsection is to verify the satisfaction about the expectations of the users on manage the exams. These "managements" referred to the operations of: registration, deletion, edition and search of patients. Thus, statements about these actions are provided, along with the exclusive options: "I agree", "I don't know" and "I disagree". Table 5 presents the answers of questions proposed in the survey analyzed in percentage.

Table 5 – Results of questions about patients management

Complete these sentences? "I can easily..."	Agree(%)	Don't Know(%)	Disagree(%)
insert a new patient into the system.	33	33	33
find patients that I have already added.	66	0	33
edit patients that I have already added.	66	0	33
access the exams of a patient.	66	0	33
remove patients that I have already added.	66	0	33

### 7.2.2 Evaluation of exam management functionalities

This step of the survey seeks to evaluate the degree of satisfaction with the performance of operations such as registration, search, processing and presentation of the exams of a patient. Thus, statements with agreement options were presented regarding the functionalities of the system. Table 6 displays the obtained results.

Table 6 – Results of questions about exams management

Complete these sentences? "I can easily..."	Agree(%)	Don't Know(%)	Disagree(%)
enter a new exam into the system.	66	0	33
find exams that I have previously added.	66	0	33
access the processing page of a added exam.	66	0	33
reprocess exams that I have previously added.	66	0	33

### 7.2.3 Evaluation of exam processing tasks and learning

This section of the survey aims to measure how satisfactory are the findings obtained from the processing and bone metastasis detection tasks as well as the completeness of the results and issuance of the automated report. This section of the survey comprises 8 questions, of which six are mandatory and two are free-text and optional.

In a first step, the user's satisfaction with the interface and experience of the learning and report preparation processes on the platform is evaluated. The results obtained are presented in the table 7.

Table 7 – Results of questions about exams processing

Complete these sentences?"I can easily "	Agree(%)	Don't Know(%)	Disagree(%)
process an exam in the system.	66	33	0
change the parameters used by the processing.	66	0	33
visualize the results obtained by the task.	66	0	33
export the results obtained to text document format.	66	33	0

Regarding the overall completeness of the results provided by the software, the users were asked to rate some aspects through the ratings bad, regular, good and excellent. The results obtained by this question are presented as a chart in Figure 41.

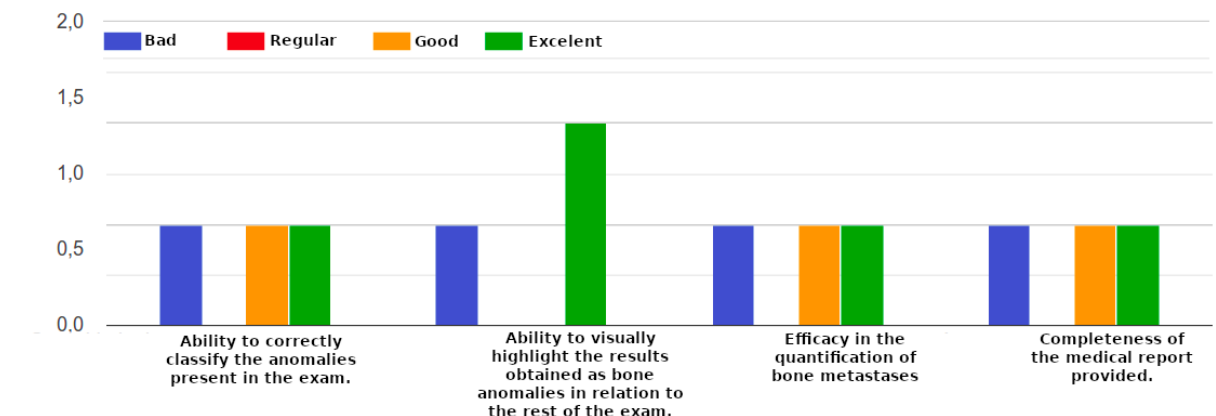


Figure 41 – Evaluation of the completeness of the results obtained by the application.

Some of the mandatory questions evaluate the user's agreement with a set of sentences and the satisfaction with the results by means of one-dimensional Likert scale. This scale, applied on multiple occasions in this questionnaire, is a method used with the objective of evaluating the opinion and attitudes of a group of people.

To quantify the user satisfaction with the described functionalities, a unipolar and increasing *Likert* scale with options from 1 to 5 was offered, where 1 symbolizes bad and 5 means an optimal result at the evaluated point.

Three questions were asked about the accuracy of the results obtained by the classifier, in this questions the user was asked to evaluate the following aspects using a *Likert* scale: As for the exam pixels classified as bone metastases, evaluate the

degree of true positives found; regarding the exam pixels classified as bone metastases, evaluate the degree of false positives found and as for the classified exam pixels, evaluate the degree of false negatives found. The results obtained for each one of this evaluations are presented in charts at figures 42, 43 and 44, respectively.

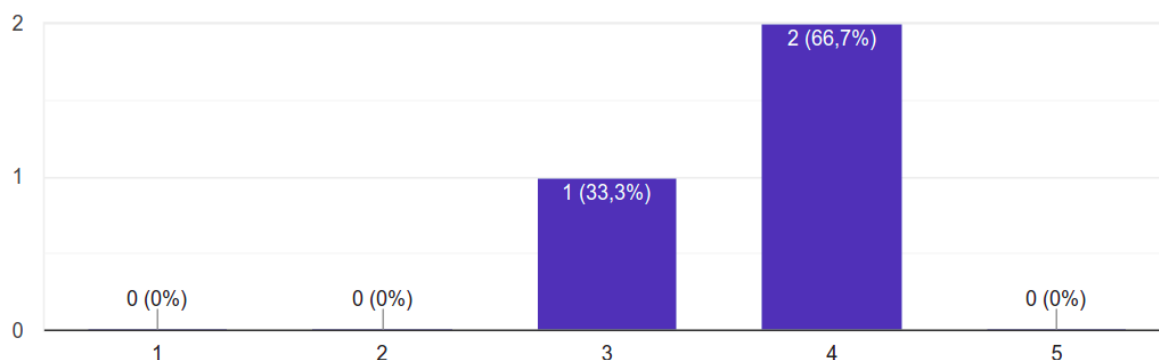


Figure 42 – Likert scale evaluating the classification of true positives.

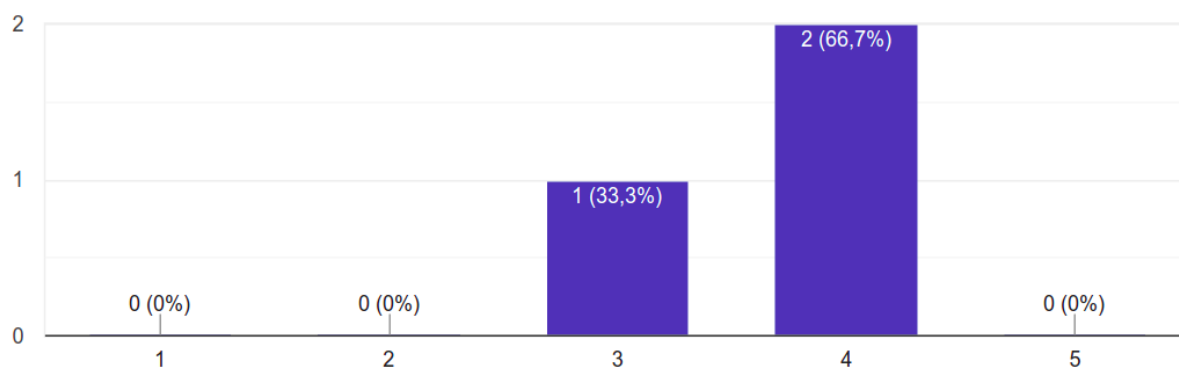


Figure 43 – Likert scale evaluating the classification of false positives.

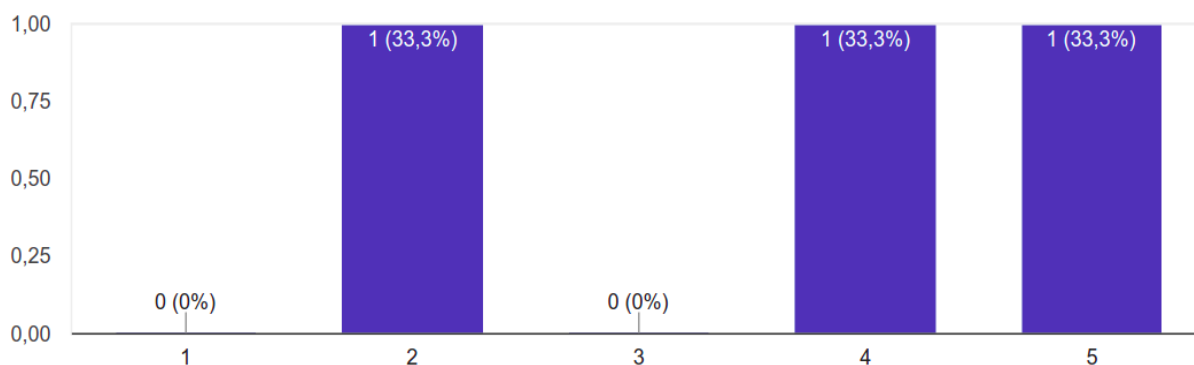


Figure 44 – Likert scale evaluating the classification of false negatives.

The last evaluated aspect of the application in this survey was a simple yes or no question regarding the accordance of the text report emitted by the software with the current standards used by the medical community. In this question, 100% of the interviewees answered that: "yes, the report is in agreement".

## 7.3 Conclusions

The most important concrete aspect of this work is the developed application that allow HUAP's colleagues to analyze (and count) easily the disease evolution by scintigraphy exams and judge if the bone metastasis are growing, stable or healing. That is the implementation (or implementations including Appendix A) design and programming.

The extensive revision of related works and their reorganization done in chapter 3 are an other relevant aspect in a broad sense, as well as the interdisciplinary group of study created and performed in such a subject during the time of this development.

### 7.3.1 Contributions of this dissertation

Specifically considering computer science aspects there are other contributions as well. Nowadays, a number of works can be found applying artificial neural networks (ANN) and convolutional neural networks (CNN) or deep learning (DL) to find and count bone metastasis by scintigraphy exam [8], [9], [3]. Although, specially in this application such a number is not so huge as in the other medical aspects, where it is almost impossible to find recent works (in height impact journal or quality conferences) that do not use one of these techniques of Artificial Intelligence (AI) or Machine Learning (ML), even in Bone scans, the use of CNN, ANN or DL to automatically compute features from images is very intense (as mentioned in Chapter 3). However, in Computer Science as in the other aspects of the life, *diversity* is welcome.

In this direction, the here presented and developed solution allows AI tasks (from computing feature to the final decision) depend only of simpler image processing and decision techniques to detect and measure such bone metastasis. Moreover, all the presented developments, even being traditional techniques have here a new programming approach. The main aspect of this is the way the features are computed from the images and used in the decision process. This will be commented in next lines.

### 7.3.2 About the originality of the work

There are 2 original ideas in the work here presented, and a lot of implementation and experimental job, as well. The first focus in the way normally enhancement or the

also named pre process step is considered in Digital Signal and Image Processing. The second is related to the image area from where the feature are extracted in the analysis done. Next subsections explains better each one.

#### 7.3.2.1 Using innate pre-process technique as features

For most of the computer practical applications of Signal and Image Processing there are always an enhancement initial phase and a final coding or understanding one [45], [29], [28].

In the enhancement phase, data are processed by traditional techniques (as the name ensure) only for improve theirs appearance to human viewers or to increase systems performances (in the subsequent phases of adopted strategies) [49]. However, the common use of any technique even concentrated and firstly designed for one process can be tried in a diverse manner. Moreover, such deviation can promote a "nice cut" in the "trip" conducting to a successfully end [50].

In the here related implementation a number of development normally involved only to image pre processing are used as features. This is the case of the techniques of Histogram Equalization, Contrast Limited Adaptive Histogram Equalization (CLAHE), Otsu thresholding, Median filtering, Gaussian filtering, Entropy filtering, Roberts edge detection, Sobel edge detection, Variance filtering and Gabor filter banks. Traditionally, all of them are used to transform an image (or a part of an image) in a enhance one related to some concepts. However, they are now converted in an numeric element of the feature vector (or the feature array) commented in section 6.10. The knowledge of their numbers after normalization (as explained in section 5.4) in each pixel related to the patient body are proposed to ML techniques. The importance and independence of each one are then consideration for the PCA used to decide when each pixels is related to a metastatic point or does not (for a specific patient on focus).

#### 7.3.2.2 Maximizing the samples to be analyzed for each exam

The second original aspect of this work is related to the area used for computation of the feature vectors to be used to detect anomalies (CAdE) and decision about it as with or without metastatic cancer (due the amount of gamma ray emission). This can be see as related to number region of interest (ROI) or segmented area on analysis in the traditional pattern recognition implementations.

As mentioned in section 6.10 all not background pixels in the frame on analysis is considered as a ROI candidate do be analysed by the developed implementation. For each one, all features are computed and the feature vector analyzed. This allow to have the biggest possible number of cases for learning with the very reduced set of exams

and patients that today are in the used dataset. As far as we know such consideration of minimum possible area from analysis or maximum possible set of feature vectors for analysis was never presented and reported before.

### 7.3.3 Final considerations

Of course, the incipient number of patients that are available to us for this development has directed to the used techniques commented in last subsections. However, it is important to say that the existence of so few cases to use can be the case present in many other circumstances and also represents important peculiarities that must be consider in other cases and ages.

Maybe, it is paramount to realize that there will be always new issue to study and possible aspect that present few samples for learning solutions. In these cases a possible solution for the problem is better than no solution at all. New diseases could appear and spread quicker then the capacity of the medical community to organize annotated datasets to be used in the machine learning supervised approaches and to allow CNN and ANN uses [48],[46],[47].

The knowledge of all possible techniques is important not only for historical and academic point of view but mainly to allow the development of new solutions in any circumstance. This and their uses are very important for real cases in unpredictable future and environments.

## 7.4 Future works

Successful development of new and real applications involves proper uses of both traditional and artificial intelligence concepts. Complete knowledge of their basic characteristics is needed to decide when each one is more relevant to a specific aspect of the problems on focus. Considering this, there is a obvious hope to be able to overcome in a near future the challenge of having more data. Then with more annotated dataset and with techniques such as data augmentation and transfer of learning [54] [55], the techniques of ML, IA and data mining can be used with more efficiency in this so important aspect of nuclear medicine [56].

Specifically addressing the steps of training of the learning algorithms, some improvements can be made regarding the experimentation of promising modern classifiers techniques such as XGBoost and the use of an auto-learning strategy, in order to delegate the task of selecting the ML algorithm that best fits into the problem to an automated solution. Regarding the pre-processing stages of training, strategies such as data augmentation and under sampling of the true negatives class can also contribute

to the improvement of the results obtained.

Train-test split strategies other than K-Fold cross validation can also be tested as future works, as an example, one common approach in medical ML problems is the use of a technique called leave one patient out. This strategy works analogously to cross-validation, however, the selected test set at a time corresponds to a subset of data belonging to a single patient. This technique helps to isolate from the training set information with a very close level of similarity (same patient), which in some cases can bias the obtained results.

The concrete goal, or at least one of the more important aims of this study is to help medical doctors to analyze the disease evolution by scintigraphy exams and judge if the bone metastasis are growing, stable or healing. However, for the here presented dissertation, only direct measurements is discussed. So, a way to automatically evaluate such modifications could be compare the area of pixels associate with the metastasis body points after *Registering* exams of same patient in two dates [51]. (In image analysis, *Register* is a technique that, basically, take two images and transforms one image into the other, allowing match objects common in both, so that similar points in world space are also similar in image coordinate space [51].) After this match patient evolution can be fairly compared using same enhancement techniques [53] [52]. The disease progress evaluation then can be properly made by subtract the images. This is the first simple and direct improvement suggested [52].

The Bone Scan Index (BSI) can also be added in the future as another quantitative measure output in the application [91]. In order to achieve such measure, the already available functionality of calculating the ratio of metastasis over foreground pixels must be extended in order to be calculated only over pixels that belongs to the patient skeleton. Segmentation of the skeleton in a bone scan image is a task that requires a set of well defined IA and IP operations and ground truth images in order to evaluate the accuracy of the segmentation. Some initial studies were performed in this dissertation to verify how difficult is this segmentation task in comparison with the study developed in this work. Although there are no ground truth for skeleton available for realization of this study, good visual results were obtained by combination of a denoise filter and histogram based segmentation.

Another important incoming feature for the application is the support to process bone scan exam from small regions and not only whole-body exams. With this, physicians will be able to use all the already available functionalities (to whole-body) also to process localized and small regions scintigraphy exams such as scans containing only cranium, pelvic and others.

In order to do this small regions bone scans classification, some tests are being

made, including training new classifiers with some adjusted features. Although the majority of the implemented features applied at full body scans can be easily used in small regions bone scans, some features like pixel position and pixel on pelvic region needs to be adapted.

At this moment, support to cranium exams is already at development stage and the physicians can test the potential of this new functionality in the application as an *alpha* mode. The available implementation is a simple adaptation of the classifier available to full-body bone with the adjustments of the pixel position feature. Unfortunately, the number of current annotated exams containing skulls and having hotspots is considerably less than the number of exams for the whole body (there are 2 exams of this type available). Thus, one of the exams was added to the training dataset to the classifier and the other was used to test the obtained model. Figure 45 shows the obtained result for this incipient test.

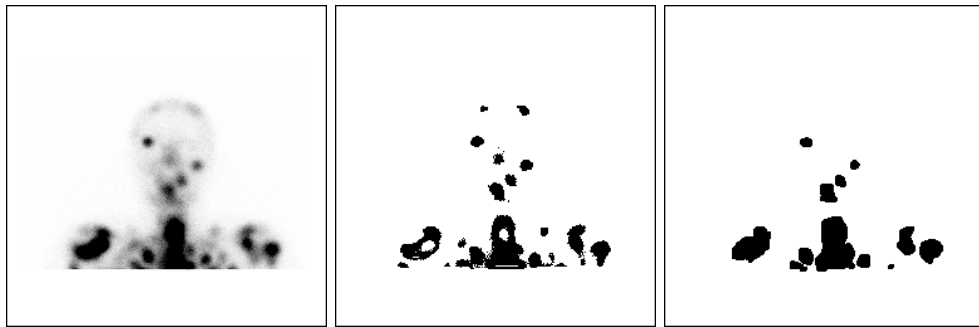


Figure 45 – Original, results and ground truth of cranium.

Finally, it is important to highlight that as a real application case, develop new features and improve the already available functionalities will be always a work in progress for this study.

## References

- [1] OSEAS, A. Metodos para sistemas CADe e CADx de nódulos pulmonares baseados em tomografia computadorizada usando análise de forma e textura. UFMA. 2016. 20
- [2] BRAZ, J. et al. A mass classification using spatial diversity approaches in mammography images for false positive reduction. Expert Systems with Application. 2013. 20
- [3] NAKAJIMA, K; EDENBRANDT L; MIZOKAMI, A. Bone scan index: A new biomarker of bone metastasis in patients with prostate cancer. Int J Urol. 2017. 33, 104
- [4] SADIK, M. JAKOBSSON, D.; OLOFSSON, F.; OHLSSON, M.; SUURKULA, M.; EDENBRANDT, L. A new computer-based decision-support system for the interpretation of bone scans. Nuclear medicine communications. 2006. 32, 33, 99
- [5] SADIK, M. HAMADEH, I.; NORDBLOM, P.; SUURKULA, M.; HOGLUND, P.; OGLSSON, M.; EDENBRANDT, L. Computer-Assisted Interpretation of Planar Whole-Body Bone Scans. J Nucl Med. 2008. 32, 33, 99
- [6] HORIKOSHI, H.; KIKUCHI A.; ONOBUCHI M. Computeraided diagnosis system for bone scintigrams from Japanese patients: importance of training database. Ann Nuclear Medicine. 2012. 33, 99
- [7] NAKAJIMA, K.; NAKAJIMA, Y.; HORIKOSHI, H. Enhanced diagnostic accuracy for quantitative bone scan using an artificial neural network system: a Japanese multi-center database project. EJNMMI Res. 2013. 33, 34, 99
- [8] DANG, J. Classification in bone scintigraphy images using convolutional neural networks. Master's Theses in Mathematical Sciences. 2016. 10, 34, 35, 99, 104
- [9] BELCHER, L. Convolutional neural networks for classification of prostate cancer metastases using bone scan images. Lund University. 2017. 35, 99, 104
- [10] PAPANDRIANOS N.; PAPAGEORGIOU E.; ANAGNOSTIS A.; PAPAGEORGIOU K. Bone metastasis classification using whole body images from prostate cancer patients based on convolutional neural networks application. PLoS ONE. 2020. 10, 20, 35, 36, 99
- [11] MACEDO, F.; LADEIRA, K.; PINHO F.; SARAIVA, N.; BONITO, N.; PINTO, L.; GONCALVES, F. Bone Metastases: An Overview. Oncology reviews. 2017. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5444408>. Accessed at: 06 oct. 2020. 23, 24

- [12] STUMP-SUTLIFF K.; CUNNINGHAM L., GERSTEN T. Bone Metastases: When Cancer Spreads to the Bones. Rochester. 2021. Available at: <https://www.urmc.rochester.edu/encyclopedia/content.aspx?ContentTypeID=34&ContentID=25952-1>. Accessed at: 04 set. 2021 20, 23
- [13] WYNGAERT, T.; STROBEL, K.; KAMPEN, W.; KUWERT, T.; BRUGGEN, W.; MOHAN, H.; GNANASEGARAN, G.; DELGADO-BOLTON, R.; WEBER, W.; BEHESHTI, W.; LANGSTEGGER, M.; GIAMMARILE, W.; MOTTAGHY, F.; PAYCHA, F. The EANM practice guidelines for bone scintigraphy. European Journal of Nuclear Medicine and Molecular Imaging. 2016. 29
- [14] WYANT, T. Bone metastases. Cancer.org. 2021 Available at: <https://www.cancer.org/treatment/understanding-your-diagnosis/advanced-cancer/bone-metastases.html>. Accessed: at 28 aug. 2021 23
- [15] Westmead BCI Institute. Bone Metastasis. Westmead BCI Institute. 2017. Available at: <https://www.bci.org.au/breast-cancer-information/fact-sheets/bone-metastases/>. Accessed at: 20 aug. 2021. 24, 25, 27, 28
- [16] Rogel U-M Cancer Center. Bone Metastasis. Rogel U-M Cancer Center. 2021, Available at: <https://www.rogelcancercenter.org/bone-metastasis>. Accessed at: 04 aug. 2021. 23
- [17] RIEDEN, K. Conventional imaging and computerized tomography in diagnosis of skeletal metastases Radiologe. Der Radiologe. 1995. 29
- [18] DONOHOE K.; BROWN M.; COLLIER, D.; CARRETTA, R.; HENKIN, R.; O'MARA, R.; ROYAL, H. Society of Nuclear Medicine Procedure Guideline for Bone Scintigraphy. 2023. 30
- [19] Hopkins Medicine. Bone Scan. Hopkins Medicine. 2021. Available at: <https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/bone-scan?amp=true>. Accessed at: 10 nov. 2020 29, 30, 31
- [20] DICOM standard. About DICOM: Overview. DICOM standard. 2021. Available at: <https://www.dicomstandard.org/about>. Accessed at: 18 dez, 2020 31
- [21] DICOM nema. DICOM PS3.1 2021c - Introduction and Overview. DICOM nema. 2021. Available at: <http://dicom.nema.org/medical/dicom/current/output/pdf/part01.pdf>. Accessed at: 18 dez. 2020. 31
- [22] MCCORMICK, M. DICOM Rescale Intercept / Rescale Slope and ITK. Blog Kitware. 2014. Available at: <https://blog.kitware.com/>

- [dicom-rescale-intercept-rescale-slope-and-itk/](#). Accessed at: 18 set. 2021. 65, 66
- [23] FISHER, R.; PERKINS, S.; WALKER, A; WOLFART, E. Gaussian Smoothing. University of Edinburgh. 2003. Available at: <https://homepages.inf.ed.ac.uk/rbf/HIPR2/gsmooth.htm>. Accessed at: 09 oct. 2021. 41, 42
- [24] THUM, C. Measurement of the Entropy of an Image with Application to Image Focusing. *Optica Acta: International Journal of Optics*. 1984. 45
- [25] Scikit image. Entropy. Scikit image. 2021. Available at: [https://scikit-image.org/docs/dev/auto\\_examples/filters/plot\\_entropy.html](https://scikit-image.org/docs/dev/auto_examples/filters/plot_entropy.html). Accessed at: 20 set. 2021 45
- [26] LS, D. A survey of edge detection techniques. *Computer Graphics and Image Processing*, vol 4. 1975 45
- [27] CASTRO, J.; FERNANDES, M.; ALMEIDA, R.; *Estatística e Probabilidade*. 2015 41
- [28] GONZALEZ. R.; WOODS, R. *Digital image processing*, vol 3. Pearson. 2002. 37, 38, 39, 70, 105
- [29] SONKA, M.; HLAVAC, V.; BOYLE, R.; *Image processing, analysis and machine vision*. Cengage Learning. 2014. 37, 39, 44, 105
- [30] ALLIBONE, T.; Dennis Gabor. 5 June 1900 – 9 February 1979. *Biographical Memoirs of Fellows of the Royal Society*. 1980.
- [31] FOGEL, I.; SAGI, D. Gabor filters as texture discriminator. *Biological cybernetics*, v. 61. 1999. 46
- [32] HAYKIN, S. *Redes Neurais- Princípios e Práticas*. BOOKMAN. 2001. 57, 58
- [33] IRINA, R. (2001). An Empirical Study of the Naïve Bayes Classifier. *IJCAI 2001 Work Empir Methods Artif Intell*. 3. 55
- [34] LANGLEY, P.; IBA, W.; THOMPSON, K. An analysis of Bayesian classifiers. In *Proceedings of the Tenth National Conference on Artificial Intelligence*, pages 399–406, San Jose, CA, 1992. 55
- [35] SCIKIT. Support Vector Machines. 2017. Available at: <https://scikit-learn.org/stable/modules/svm.html>. Accessed at: 21 set. 2021 56
- [36] HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Springer. 2008. 56

- [37] BOSER, B.; GUYON, I.; VAPNIK, V. A training algorithm for optimal margin classifiers. Proceedings of the fifth annual workshop on Computational learning theory. 1992. 57
- [38] VANDERPLAS, J. Python Data Science Handbook. O'Reilly Media. 2017 50, 52, 55
- [39] ROSEBROCK, A. OpenCV Template Matching. Pyimagesearch. 2021 Available at: <https://www.pyimagesearch.com/2021/03/22/opencv-template-matching-cv2-matchtemplate>. Accessed at: 20 oct. 2021 48
- [40] LAMPERT, C.; BLASCHKO, M.; HOFMANN, T. Beyond sliding windows: Object localization by efficient subwindow search. IEEE conference on computer vision and pattern recognition. 2008. 48, 72
- [41] RAO, Y.; PRATHAPANI, N.; NAGABHOOSHANAM, E. Application of normalized cross correlation to image registration. International Journal of Research in Engineering and Technology. 2014. 48, 49
- [42] REZA, A. Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for Real-Time Image Enhancement. The Journal of VLSI Signal Processing-Systems for Signal, Image, and Video Technology. 2004. 40, 41
- [43] YOUSEFI, J. Image binarization using Otsu thresholding algorithm. University of Guelph. 2011. 43
- [44] OTSU, N. A threshold selection method from gray level histograms. IEEE Trans. systems. Man. and Cybernetics. 1979. 43
- [45] LIM, J. Two-Dimensional Signal and Image Processing. Prentice-Hall. 1990. 37, 105
- [46] BARROS, B.; LACERDA P.; ALBUQUERQUE C.; CONCI, A. Pulmonary COVID-19: Learning Spatiotemporal Features Combining CNN and LSTM Networks for Lung Ultrasound Video Classification. Special Issue Image Sensing and Processing with Convolutional Neural Networks. 2021 106
- [47] BARROS, B.; LACERDA P.; ALBUQUERQUE C.; CONCI, A. Hyperparameter Optimization for COVID-19 Pneumonia Diagnosis Based on Chest CT, Biosensors. Sensors. 2021. 106
- [48] GONCALVES, L.; ARAGAO, D., OLIVEIRA, E., BEZERRA, A.; SANTOS, D.; GONCALVES, A.; GADELHA, I.; PISCITELLI, P., MIANI, A.; DISTANTE, C.; SALAS, J.;

- CONCI, A. Multivariate Data Driven Prediction of COVID-19 Dynamics: Towards New Results with Temperature, Humidity and Air Quality Data. *Journal of Environmental Research*. 2022. 106
- [49] RESMINI, R.; SILVA, L.; MEDEIROS, P.; ARAUJO, A.; MUCHALUAT-SAADE, D.; CONCI, A.; Recent Trends in Cognitive Computing for Healthcare, A Hybrid Methodology for Breast Screening and Cancer Diagnosis Using Thermography. *Computers in Biology and Medicine*, Volume 135. 2021. 105
- [50] RODRIGUES, E; CONCI, A.; LIATSI, P. ELEMENT: Multi-Modal Retinal Vessel Segmentation Based on a Coupled Region Growing and Machine Learning Approach. *IEEE Journal of Biomedical and Health Informatics*. 2020. 105
- [51] SAYAN, N., Image Registration Techniques: A Survey. *CoRR*, vol abs/1712.07540, 2017. 107
- [52] RUBLEE, E.. RABAU, V., KONOLIGE, K.; BRADSKI, G. ORB: An efficient alternative to SIFT or SURF. *ICCV* . 2011. 107
- [53] FISCHLER, M.; BOLLES, R. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*. 1981. 107
- [54] MORAN, M.; FARIA, M.; GIRALDI, G.; BASTOS, L.; CONCI, A. Using super-resolution generative adversarial network models and transfer learning to obtain high resolution digital periapical radiographs. *Computers in Biology and Medicine*. 2020. 106
- [55] FALCONI L.; PEREZ, M.; AGUILAR, W.; CONCI, A.; Transfer Learning and Fine Tuning in breast mammogram abnormalities classification on CBIS-DDSM database. *Advances in Science, Technology and Engineering Systems Journal*, volume 5. 2020. 106
- [56] TINOCO, C.; SANTOS W.; CUNHA, J. *Cardiologia Nuclear em 2020 - Perspectivas da Nova Diretriz da Sociedade Brasileira de Cardiologia*. *Arquivos Brasileiros de Cardiologia*. 2020. 106
- [57] MATTEN, L.; POSTMA E. Dimensionality Reduction: A Comparative Review. *J Mach Learn Res*. 2009. 51
- [58] FUKUNAGA, K. *Introduction to Statistical Pattern Recognition*. Elsevier, 2013. 51
- [59] JIMENEZ, L.; LANDGREBE, D. Supervised classification in high-dimensional space: geometrical, statistical, and asymptotical properties of multivariate data. *IEEE Transactions on Systems, Man and Cybernetics*. 1997. 51

- [60] ANGUIA, D.; GUIO, A.; RIDELLA, S.; STERPI, D. K-Fold Cross Validation for Error Rate Estimate in Support Vector Machines, DMIN. 2009. 53
- [61] ANGUIA, D.; RIDELLA, S.; RIVIECCIO, F. K-Fold Generalization Capability Assessment for Support Vector Classifiers, Proc. of the IEEE Int. Joint Conf. on Neural Networks, IJCNN . 2005. 53
- [62] QUINLAN, J. Simplifying decision trees. International Journal of Man-Machine Studies. 1987. 56
- [63] BREIMAN, L. Machine learning: Random Forests. Springer. 2001. 56
- [64] BISHOP, C. Pattern recognition and machine learning. Springer. 2006. 49
- [65] BELLMAN, R. Dynamic programming. Princeton University. 1957. 51
- [66] BELLMAN R. Dynamic Programming. Courier Dover Publications. 2003. 51
- [67] BREMS, M. A One-Stop Shop for Principal Component Analysis. Towards Data Science. 2017. Available at: <https://towardsdatascience.com/a-one-stop-shop-for-principal-component-analysis-5582fb7e0a9c>. Accessed at: 10 oct. 2021. 52, 53
- [68] YIU, T. Understanding PCA (Principal Components Analysis). Towards Data Science. 2019. Available at: <https://towardsdatascience.com/understanding-pca-fae3e243731d>. Accessed at: 08 oct. 2021. 52
- [69] CHAKRAVORTY, P. What is a Signal? [Lecture Notes]. IEEE Signal Processing Magazine. 2018. 37
- [70] HERNANDEZ, P. Fiftieth Anniversary of First Digital Image Marked. NIST. 2007.
- [71] AZRIEL, R. Picture Processing by Computer. New York: Academic Press. 1969.
- [72] KASHYAPA, R. Image Acquisition Components. Machine Vision Blogs. 2020. Available at: <https://qualitastech.com/image-acquisition-components>. Accessed at: 13 dec. 2021. 37
- [73] HALDAR, S. Photogeology, Remote Sensing and Geographic Information System in Mineral Exploration .Elsevier. 2013. 38
- [74] SHAPIRO, L.; STOCKMAN, G. Computer Vision. pp 279–325. Prentice-Hall, ISBN 0-13-030796-3. 2001. 37
- [75] RAJESHWAR, D., PRIYANKA; SWAPNA, D. Image Segmentation Techniques. DCR University of Science and Technology. 2012. 39

- [76] SINGH, K.; SINGH, A. A Study of Image Segmentation Algorithms for Different Types of Images. *International Journal of Computer Science Issues*, Vol. 7. 2010.
- [77] SUTTN, E. Histograms and the Zone System. *Illustrated Photography*. 2015. 39
- [78] MORISTA, J.; BOTELHO, H.; FERNANDES, F.; TINOCO, C; CONCI, A. Presenting a system to aid on the scintigraphy bone metastasis analysis using DICOM files. *International Conference on Systems, Signals and Image Processing*. 2021. 38, 69, 99, 118
- [79] SUZUKI, S. Topological structural analysis of digitized binary images by border following. *Computer Vision, Graphics, and Image Processing*, vol. 30, pages 32–46. Elsevier. 1985. 119
- [80] MURZOVA, A. Otsu's Thresholding with OpenCV. *Learn Opencv*. 2020. Available at: <https://learnopencv.com/otsu-thresholding-with-opencv/>. Accessed at: 15 dec. 2021. 43
- [81] WEISSTEIN, E. Gaussian Function. *MathWorld—A Wolfram Web Resource*. Available at: <https://mathworld.wolfram.com/GaussianFunction.html>. Accessed at: 15 dec. 2021. 41
- [82] WASSERMAN, L. *All of Statistics: a concise course in statistical inference*. Springer texts in statistics. p. 51. ISBN 9781441923226. 2005. 41
- [83] PETKOVA, G. The Gold Standard – The Key to Information Extraction and Data Quality Control. *OntoText*. 2021. Available at: <https://www.ontotext.com/blog/gold-standard-key-to-information-extration-data-quality-control/>. Accessed at: 27 dec. 2021. 51
- [84] KOZYRKOV, C. What is “Ground Truth” in AI? (A warning.). *Towards Data Science*. 2020. Available at: <https://towardsdatascience.com/in-ai-the-objective-is-subjective-4614795d179b>. Accessed at: 27 dec. 2021. 51
- [85] APEER, An intuitive annotation tool for your machine learning needs. *APEER*. Available at: <https://www.apeer.com/annotate>. Accessed at: 30 dec. 2021. 68
- [86] ŁUKASZEWSKI, B.; NAZAR, J.; GOCH, M.; ŁUKASZEWSKA, M.; STĘPIŃSKI, A.; JURCZYK, M. Diagnostic methods for detection of bone metastases. *Contemporary oncology (Poznan, Poland)*, 21(2), 98–103. <https://doi.org/10.5114/wo.2017.68617>. 2017. 25, 27, 28

- [87] Cancer.org. Tests for Bone Cancer. Cancer.org. 2021. Available at: <https://www.cancer.org/cancer/bone-cancer/detection-diagnosis-staging/how-diagnosed.html>. Accessed: at 07 jan. 2022. 24
- [88] NIBIB. Computed Tomography (CT). National Institute of Biomedical Imaging and Bioengineering. <https://www.nibib.nih.gov/science-education/science-topics/computed-tomography-ct>. Accessed at 07 jan. 2022. 26
- [89] NIBIB. Magnetic Resonance Imaging (MRI). National Institute of Biomedical Imaging and Bioengineering. Available at: <https://www.nibib.nih.gov/science-education/science-topics/magnetic-resonance-imaging-mri>. Accessed at 08 jan. 2022. 27
- [90] Hopkins Medicine. Positron Emission Tomography (PET). Hopkins Medicine. Available at: <https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/positron-emission-tomography-pet>. Accessed at 08. jan. 2022. 28
- [91] ULMERT, D.; KABOTEH, R.; FOX JJ. A novel automated platform for quantifying the extent of skeletal tumour involvement in prostate cancer patients using the Bone Scan Index. *Eur Urol.* 62(1):78-84.PubMedGoogle ScholarCrossref. 2012. 32, 107
- [92] SABBATINI, P.; LARSON, SM.; KREMER, A. Prognostic significance of extent of disease in bone in patients with androgen-independent prostate cancer. *J Clin Oncol.* 17(3):948-957. 1999. 32
- [93] ERDI, YE.; HUMM, JL.; IMBRIACO, M.; YEUNG, H.; LARSON, SM. Quantitative bone metastases analysis based on image segmentation. *J. Nucl. Med.* 38: 1401–6. 1997. 32
- [94] ASLANTAS, A.; DANDIL, E.; SAGLAM, S.; ÇAKIROGLU, M. CADBOSS: A computer-aided diagnosis system for whole-body bone scintigraphy scans. *Journal of cancer research and therapeutics*, 12(2), 787. 2016. 34, 99
- [95] HEINDEL, W.; GÜBITZ, R.; VIETH, V.; WECKESSER, M.; SCHÖBER O.; SCHÄFFERS M. The diagnostic imaging of bone metastases. *Dtsch Arztebl Int.*;111(44):741-747. doi:10.3238/arztebl.2014.0741. 2014 8, 25, 27, 28
- [96] COOK G. PET and PET/CT imaging of skeletal metastases. *Cancer Imaging*;10(1):1-8. doi:10.1102/1470-7330.2010.0022. 2010. 8, 29

# Appendix

## APPENDIX A – Previous developments

For the challenge of develop this CADx-CADe some approaches were tempted. The first one was to perform a standard sequence of image processing operations on the bone scan images, starting from image enhancement operations like: median filter to denoise; histogram equalization to contrast improvement; ending up with automatic segmentation using the Otsu technique. Although it seems to be a good strategy, after a few bone scans processing, it was clear that although some of the selected techniques and parameters perform well for some files, for others, they lead to poor segmentation and denoising results. The inability to find a good combination of operations and the related parameters that works well for the majority of images in the dataset leads think on other approach.

A second approach was released to the physicians as the first version of this solution [78]. In this version, a semi-automated implementation was applied for bone metastasis detection and classification. Figure 46 shows the system screen of this version related to the segmentation step. Note that, at that time, the unique language available for communication in the interface was (Brazilian) Portuguese.

Firstly, in this version, a linear transformation of the values of the absorption matrix extracted from the DICOM file was performed as described at section 6.6. After the conversion of the bone scan DICOM file to gray scale images, the output image was presented to the physicians so they could be guided through a series of image processing operations, choosing operations to be applied and their parameters at each steps. The first one was image denoising and the available operations presented to the physician were the median, Gaussian and bilateral filters. After the image treatment is carried out, the segmentation phase begins. In this phase, a threshold related to the intensity of the gray tones of the pixels present in the image is defined.

The default segmentation method was Otsu combined with the unsupervised learning clustering technique K-Means. However, if the user was not satisfied with the result or did not wish to use the suggested algorithm, a manual thresholding could also be applied with a Random Walker algorithm. The process of applying the algorithms described in this phase could be repeated as wanted until the segmentation was considered satisfactory by the user.

The next operation considered is edge detection. The default method of this application version was the Sobel filter. This standard choice was motivated by the good performance and low computational cost of this method on some performed tests. However, other options for edge enhancement were available such as Prewitt, Roberts

and Scharr. Once an image with well-defined edges and regions was obtained, the process of generating the feature vectors that will be used in subsequent supervised learning tasks began. For this purpose, Suzuki following border algorithm was used [79] (it is available in the OpenCV Python library). After executing this technique, a list of entries is obtained, each one referring as a sub region containing a closed contour of the image. From this map of region it was possible to extract geometric information such as area, enclosing rectangle and others.

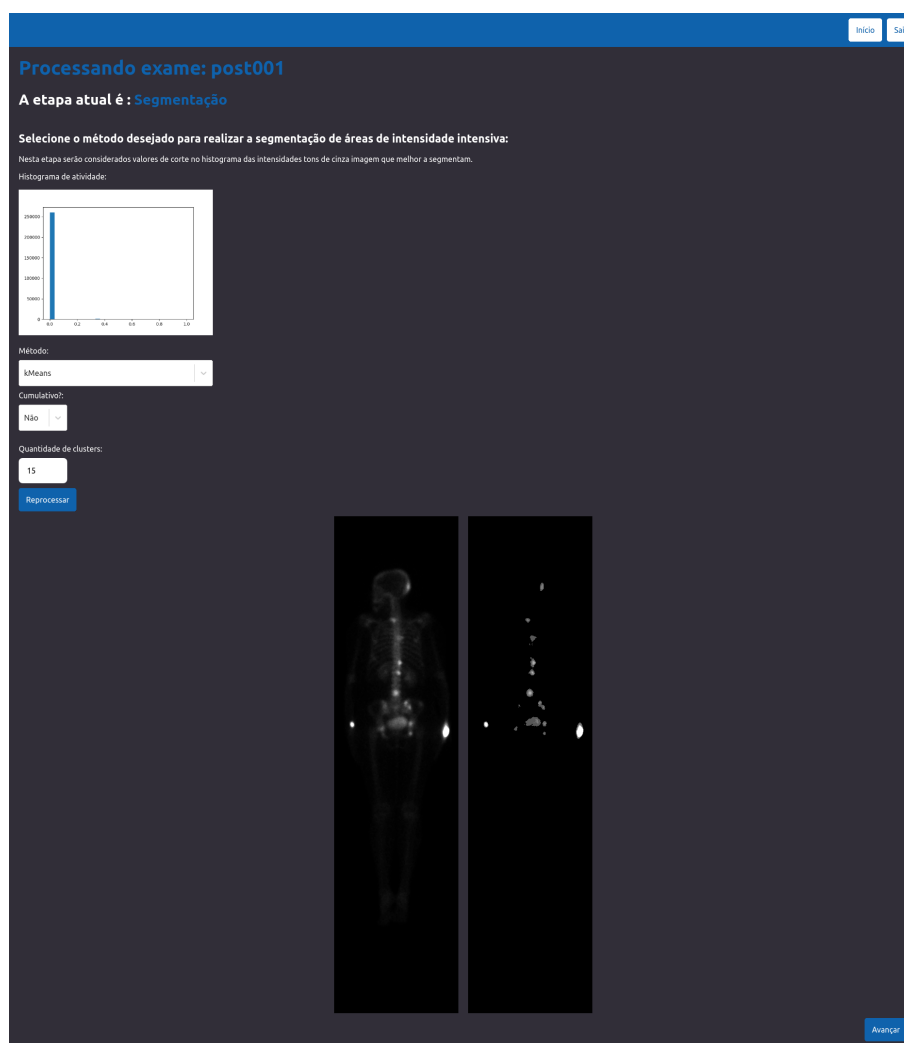


Figure 46 – Threshold operation of first version.

Ten (10) features were computed and used to compose the feature vector to describe each region mapped as described in last paragraph. The features for each possible metastatic area were: Area, Perimeter, Position of the centroid, Aspect ratio of the bounding box (BB) of the region, ratio of the areas, equivalent diameter, average intensity of the area, orientation of the adjusted ellipse, and eccentricity of this ellipse.

The feature named *Area* means the area of the region. It is computed by the

number of pixels present inside each boundary converted to square millimeters using the metadata available in the DICOM exam file as Section 6.6 has commented.

The feature named *Perimeter* is the perimeter of each region is obtained by counting the border pixels. This value was converted to millimeters (*mm*) using the DICOM metadata available 6.6.

The third and fourth represent the *Horizontal* and *Vertical* location of the centroid of each possible metastasis on the exam image. These values are calculated considering the point of origin of the coordinate system at the top left of the image and obtained through the first geometric moments of the region divided by the area (or the zero order moment of each region).

The fifth feature of the vector symbolizes the area proportion (or aspect ratio). For this, initially, the bounding box is calculated for each one, and then its horizontal (width) and vertical (height) size are evaluated. The ratio between width and height of this bounding box represents this feature.

The feature referring to the sixth position of the vector represents the ratio between the area of the region (calculated previously and described by the first position of this vector of features) and the area of the bounding box computed to represent the fifth element of this feature vector.

Position seven of the feature vector considers an attribute called *equivalent diameter*. This value represents the diameter of the circle of the same area of the grouped region (square root of the area divided by the irrational number  $\pi=\pi$ ).

The eighth position shows the value of the *average intensity* of the pixels present inside the region.

Position number nine symbolizes the *orientation* of the region. Defined as the orientation of the ellipse that has the same second moments of the region, ranging from  $-\pi/2$  to  $\pi/2$  counterclockwise.

The tenth position of the vector refers to the *eccentricity* of the region. This is calculated by the distance between adjusted ellipse focal points. Adjusted ellipse definition referees to one with the same second moments of the region considered at position one. The eccentricity is the proportion of the focal length (distance between the focal points) over the length of the main axis. The value must be in the range  $[0, 1]$ . When 0 means that the ellipse becomes a circle.

At the end of this processing step, a set of feature vectors is obtained, each one will be then verified if really is referring to a region of metastases in the image after the required learning processing in new operations of exam analysis.

The green boxes on Figure 47 outlines some findings of this approach. The

rectangles enclose areas of the original image that received the label `metastasis` as result. The numbers presented (over each one) are the area computed as metastasis in the region. As already explained they are computed by the number of pixels in the metastasis group transformed in equivalent square millimeters (approximated up to one digit after the integer point position, that is  $\frac{1}{10} mm$ ).

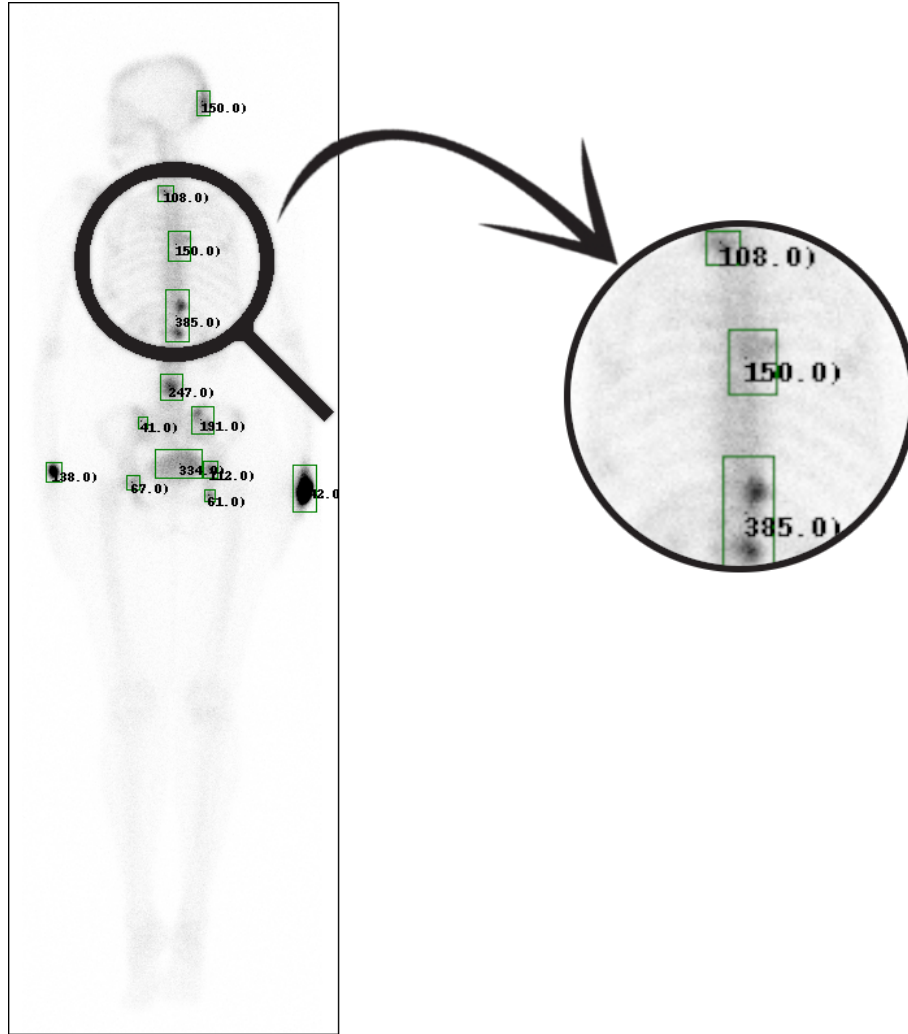


Figure 47 – Detected regions and metastasis areas in  $mm^2$ .

After processing the exams of the patients (there were only 10 exams available), a total of 77 entries was obtained. For each one of this entries, in addition to the 10 features related to the geometry and intensities of the regions, there was one more related to the patient's medical data informed by the physicians. This was 1 or 0 symbolizing true or false. For the train test split, a simple split approach was adopted where 70% of the samples contained in the database were used for training while the remaining 30% were used to quantify the accuracy of the results.

The detections dataset generated by this described steps was used for trai-

ning four classifiers: K-Nearest neighbors (K-NN), Decision Tree (DT), Support Vector Machine (SVM) and Multilayer Perceptron (MLP).

The KNN algorithm used a neighbors equals to 1. This was selected because during the executions, this parameter presented the highest accuracy value. For the MLP, 100 hidden layers, Adam optimizer and logistic activation function were taken as hyper-parameters. SVM parameters were radial basis function kernel with a gamma of 0.08 and no limit of iterations.

Accuracy, recall, precision and the area under the Receiver Operation Characteristic Curve (ROC) were computed with the obtained results. The results of the training and test sets are presented (relatively to the total of cases in percentage) in tables 8 and 9.

Table 8 – Results for training set.

Classifier	Accuracy (%)	Precision (%)	Recall (%)	ROC
K-Nearest Neighbor (KNN)	100	100	100	1.00
Multilayer Perceptron (MLP)	81	84	57	0.76
Decision Tree (DT)	100	100	100	1.00
Support Vector Machine (SVM)	94	100	84	0.92

Table 9 – Results for test set.

Classifier	Accuracy (%)	Precision (%)	Recall (%)	ROC
K-Nearest Neighbor (KNN)	82	70	87	0.83
Multilayer Perceptron (MLP)	73	62	62	0.71
Decision Tree (DT)	86	77	87	0.87
Support Vector Machine (SVM)	78	66	75	0.77

Although the results with DT were really promising at that time, the patients dataset was still too small to effectively improve the generalization capacity of the implementation. The low results obtained by the SVM and MLP indicated the need of a larger set for training.

Another inefficiency with this version was the need of users interaction to approve and coordinate the segmentation steps. The user had to understand something of image processing and learning for adequate selection of tuning parameters. That is, the process was not completely autonomous. Unfortunately, this semi automated steps were leading to a high learning curve for the physicians to make effective use of the implemented set of operations and their sequences for achieve to a good solution that works well for the bone scintigraphies.

These obstacles (small dataset, difficult to select parameters for operations and user interaction needed) leaded to the a decision for finding the methodology applied in this dissertation commented in Section 6.9.