FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Pancreatic Ductal Adenocarcinoma Detection in CT Images using Deep Learning

Rita Catarina Moreira Alves

M.SC. DISSERTATION

MASTER IN BIOMEDICAL ENGINEERING

Supervisor: Tânia Pereira, PhD

Second Supervisor: Hélder Filipe Oliveira, PhD

September 11, 2022

© Rita Catarina Moreira Alves, 2022

Pancreatic Ductal Adenocarcinoma Detection in CT Images using Deep Learning

Rita Catarina Moreira Alves

MASTER IN BIOMEDICAL ENGINEERING

September 11, 2022

Resumo

O cancro é um grupo de doenças caracterizado por uma divisão celular anormal e disseminação das células para outros tecidos. Esta condição é a principal causa de morte em todo o mundo, incluindo o cancro do pâncreas que é normalmente diagnosticado em estadios avançados. Sabe-se que a taxa de sobrevida aumentará se o diagnóstico for realizado precocemente. No entanto, o aumento de trabalho clínico, a fadiga, as distrações e as limitações do sistema visual humano contribuem para erros no diagnóstico. Comparando com outras modalidades de imagem, a Tomografia Computorizada é a mais utilizada para a avaliação inicial de suspeitas de cancro pancreático. No entanto, a avaliação utilizando imagens de Tomografia Computorizada é uma tarefa difícil devido à complexidade das imagens. O pâncreas apresenta grande variabilidade na forma, pouca distinção em relação aos outros órgãos abdominais, ocupando apenas uma área muito pequena de uma tomografia computorizada abdominal, aumentando a complexidade do problema. Recentemente, devido ao progresso notável do Deep Learning e ao aumento das necessidades médicas, os sistemas de diagnóstico assistido por computador tornaram-se um novo foco de pesquisa. É importante mencionar que apesar das estatísticas devastadoras causadas pelo cancro do pâncreas, não existem programas de rastreio para verificar precocemente a possibilidade da presença da patologia.

O progresso do Deep Learning contribui para o desenvolvimento de algoritmos precisos para distinguir entre o cancro pancreático e o tecido benigno. O principal objetivo com esta dissertação é desenvolver um modelo preciso baseado em deep learning para identificação do cancro pancreático, respetivamente, o Adenocarcinoma Ductal do Pâncreas. Para realizar esta tarefa este projeto foi dividido em duas etapas, tendo sido utilizado em ambas a técnica de aprendizagem por transferência. Na primeira etapa pretende-se identificar a área de interesse que contém o pâncreas, de modo a remover a informação não relevante. Nesta etapa foi utilizado o modelo Faster R-CNN ResNet-50 FPN. Na segunda etapa pretende-se identificar a presença de adenocarcinoma ductal do pâncreas e para isso é utilizado o modelo AlexNet, em que cada imagem é classificada como normal ou que contém o cancro. Neste projeto foi utilizada a framework Pytorch, encontrando-se os modelos e respetivos pesos utilizados disponíveis no módulo TorchVision. Foram utilizadas as bases de dados do Instituto Nacional de Saúde — Tomografia Computorizada do Pâncreas e Consórcio de Análise de Tumores Proteómicos Clínicos do Instituto Nacional do Cancro — Adenocarcinoma Ductal Pancreático. A tarefa de delimitação da região de interesse mostra um coeficiente de Dice médio de 92%, com uma especificidade de 94%, sensibilidade de 89% e exatidão de 92%.A segunda etapa alcança um desempenho semelhante em comparação com os métodos do estado da arte, alcançando uma precisão de 89%, especificidade de 93%, sensibilidade de 90%, exatidão de 92% com uma área abaixo da curva de 0,97 (com 95% de intervalo de confiança).

Abstract

Cancer is a group of diseases characterized by an abnormal cell division and spread of the cells for other tissues. This condition is the leading cause of death worldwide, including the pancreatic cancer which is generally diagnosed in advanced stages. It is known, that the survival rate increase if the diagnosis is performed early. However, the increases of the clinics workload, fatigue, distractions and the limitations of the human visual system contributes to misdiagnosis. Comparing with others imaging modalities the Computed Tomography is the most commonly used for the initial evaluation of suspected pancreatic cancer. However, the evaluation using computed tomography images is a difficult task due to the complexity of the images. Pancreas presents high variability in shape, poor distinction comparing with the other abdominal organs, occupying only a very small area of the entire abdominal computed tomography scans, increasing the complexity of the problem. Recently, due to the remarkable progress in Deep Learning and increasing of medical needs, the computer-aided-diagnosis systems have become a new research focus. It is important to mention that despite the devastating statistics caused by pancreatic cancer, there are no screening programs to check earlier the possibility of presence of the pathology.

The progress of Deep Learning contributes to development of accurate algorithms to distinguish between the pancreatic cancer and benign tissue. The main goal with this dissertation is to develop an accurate model based on deep learning for the identification of pancreatic cancer, respectively, the Pancreatic Ductal Adenocarcinoma. To accomplish this task, this project was divided into two stages, having been used the transfer learning technique.The first step, intend to identify an area of interest cointaining the pancreas, to remove the non-relevant information. In this step, was used the Faster R-CNN ResNet-50 FPN model. In the second stage,intend to identify the presence of the pancreatic ductal adenocarcinoma and for this was used the AlexNet model, which each image is classified as normal or containing cancer. In this project was used the Pytorch framework and the models and respective weights are available in the TorchVision module. It is used the databases of National Institutes of Health - Pancreas CT and National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium - Pancreatic Ductal Adenocarcinoma. The delimitation of the region of interest task in the images show an Average Dice Coefficient of 92% with a specificity of 94% , a sensitivity of 89% and an accuracy of 92%. The second step achieves similar performance in comparison with state-of-the-art methods, achieving a precision of 89%, a specificity of 93%, a sensitivity of 90% , an accuracy of 92% with a area under the curve of 0,97 (with 95% of confidence interval).

Acknowledgements

This research work would not be completed if I did not express my gratitude to those who collaborated, supported and accompanied me on this journey.

First of all, I would like to thank both of my supervisors, Professor Hélder Oliveira and Tânia Pereira for the constant support, guidance and advices in seeking research directions.

To my family, specially my mother and my brother for their constant love, patience and for being an example of humility and professionalism and for always being there for me, no matter the choices I make. I hope to have the time to return you everything, one day.

To my co-workers from the Emergency Department at University Hospital Center of São João, EPE, that helping me to attend the most classes: a sincere thanks.

Last, but not least I feel incredibly fortunate to be surrounded of my friends, for their patience, wonderful moments and psychological support. To Patrícia Sousa, Raquel Oliveira, Sónia Campos, Filipa Magalhães, Maria da Graça Silva, Emanuel Silva, Tiago Silva, Rita Silva and Carlos Silva, this is it, the end of a cycle.

Rita Alves

"There are no secrets to success.
It is the result of preparation, hard work, and learning from failure."

Collin Powell

Contents

1 Introduction **1**
1.1 Motivation . . . 1
1.2 Objectives . . . 2
1.3 Contributions . . . 3
1.4 Document Organization . . . 3

2 Fundamentals of Pancreatic Cancer **5**
2.1 Anatomical Context . . . 5
2.1.1 Gross Anatomy . . . 5
2.1.2 Pancreatic Divisions and Anatomic Relationships . . . 5
2.1.3 Pancreatic Ducts . . . 7
2.1.4 Arterial and Venous Blood Supply . . . 8
2.1.5 Lymphatic System of the Pancreas . . . 8
2.1.6 Functions of the Pancreas . . . 8
2.2 Pancreatic Cancer . . . 8
2.2.1 Mortality Rate and Significance of Early Detection . . . 9
2.2.2 Challenges of Early Detection . . . 9
2.2.3 Imaging Techniques . . . 10
2.2.4 Pancreatic Cancer Staging . . . 12

3 Literature Review **15**
3.1 Computer-Aided Diagnosis in Pancreatic Cancer . . . 15
3.1.1 Pancreatic Cancer Detection and Segmentation . . . 15
3.1.2 Pancreatic Cancer Classification . . . 17
3.2 State-of-the-Art Results . . . 19
3.3 Available Datasets . . . 20
3.3.1 NIH - Pancreas CT Dataset . . . 20
3.3.2 TCIA CPTAC-PDA Dataset . . . 20
3.4 Summary . . . 21

4 Background **22**
4.1 Biomedical Image Processing and Analysis . . . 22
4.2 Pre-processing Techniques . . . 23
4.3 Machine Learning . . . 24
4.3.1 Supervised Learning . . . 24
4.3.2 Unsupervised Learning . . . 27
4.3.3 Semi-supervised Learning . . . 27
4.4 Deep Learning . . . 28

4.4.1 Structure of an Artificial Neural Network 28
4.4.2 Convolution Neural Networks 29
4.4.3 Overfitting and Techniques to Avoid It 31
4.5 Evaluation Metrics 32

5 Towards an Automatic Pancreatic Ductal Adenocarcinoma Detection System in Computed Tomography 34
5.1 Method Overview 34
5.2 First Stage: Region of Interest Detection 35
5.2.1 Data Preprocessing 36
5.2.2 ROI Pancreas Detection 37
5.3 Second Stage: Pancreatic Ductal Adenocarcinoma Classification 45
5.3.1 Introduction 45
5.3.2 Data Preprocessing 45
5.3.3 Pancreatic Ductal Adenocarcinoma Classification 46
5.4 Conclusion Remarks 50

6 Conclusions and Future Work 52
6.1 Overview 52
6.2 Future Work 52

References 54

List of Figures

2.1 The pancreas. 6
2.2 The pancreatic ducts. 7
2.3 Hounsfield unit values for different types of tissue. 11
2.4 Pancreatic Adenocarcinoma. a) Axial image without contrast injection and three-phase contrast enhancement CT of the upper abdominal acquired in arterial phase (b), pancreatic phase (c) and portal venous phase (d). 11
2.5 Anatomy of a coronal slice and axial view in CT. 12
2.6 Tumor grading of the T category of the TNM scale (a) and figure depicting the stages of PDAC (b) 14

3.1 Illustration of the HNN architecture and candidate bounding box region generation. 16
3.2 Illustration of the deep U-net model. 17
3.3 Framework of the approach of the authors. 17
3.4 Results of the authors. Correspondence in tumor location between model prediction (right panels) and radiologist's labeling (left panels). 18
3.5 The architecture of the pancreatic tumor detection network 18
3.6 Axial volume scan (**left image**) and the same scan with the respective label (**right image**). 20
3.7 Axial scan (**left image**), multiplanar reconstruction in sagittal plane (**middle image**) and multiplanar reconstruction in coronal plane (**right image**) from TCIA CPTAC-PDA dataset. 21

4.1 A typical Machine Learning pipeline. 24
4.2 Random Forest algorithm 26
4.3 Support Vector Machine algorithm 26
4.4 K-means example (K = 2). Clusters are distinguished by color. 27
4.5 Artificial Neural Network architecture.In this case, the input layer is composed by 4 inputs that are multiplied by a weight matrix (W) and passed to the next layer. This happens in each layer of the neural network, and the weights are adjusted during the training phase. This representation contains two hidden layers (with 5 nodes in first layer and 7 nodes in second layer). Finally, the last layer, called output layer with 3 nodes. 29
4.6 A typical CNN architecture. 30
4.7 Data augmentation technique. This technique can be used to artificially expand the size of a training dataset. 31
4.8 Area under the ROC curve. The figure shows the ROC curves for two classifiers (classifier A and classifier B). Note that the classifier A has a larger AUC value than classifier B. 33

5.1 Study selection flowchart of the TCIA CPTAC-PDA dataset. From 98 patients in the initial database, the final number included in the study is 59 patients. 35
5.2 The process flow of the raw data. Raw volume data were applied an orientation correction and then was applied an abdomen window. Window width (WW) and Window level (WL). 37
5.3 The architecture of Faster R-CNN containing a ResNet-50 and FPN as backbone. 38
5.4 Model losses for Region Proposal Network layer of the Faster R-CNN over the training epochs for the training and validation set. The **left** image shows the plot of objecteness loss, which used a binary cross entropy loss. The **right** image shows the plot of box loss, which use the Smooth L1 loss. 40
5.5 Model losses for ROI pooling layer of the Faster R-CNN over the training epochs for the training and validation set. The **left** image shows the plot of box regression loss, which used the Smooth L1 loss. The **right** image shows the plot of classifier loss, which use the cross entropy loss. 41
5.6 Total loss plot of the Faster R-CNN over the training epochs for the training and validation set. The total loss is the result of the sum of the previous loss functions. 41
5.7 The visualized results in the experiment, which corresponds to the correct results, that is, with the highest IoU. The green bounding boxes corresponds to the predictions and the red bounding boxes corresponds to the ground-truth. 43
5.8 The visualized results in the experiment, which corresponds to the incorrect results, that is, with the lowest IoU (IoU < 0.5).The green bounding boxes corresponds to the predictions and the red bounding boxes corresponds to the ground-truth. 43
5.9 The confusion matrix of the system. 44
5.10 The ground truth were cropped for training data generation. 45
5.11 The AlexNet architecture.Image from [1]. 47
5.12 Model loss and accuracy plot for training and validation set. 47
5.13 Confusion matrix of the system. 48
5.14 Receiver operating characteristic for the automatic system. 49
5.15 The visualized results in the experiment, which corresponds to the incorrect predictions belonging to unhealthy class. The value between parenthesis corresponds to the Jaccard index. 49
5.16 The visualized results in the experiment, which corresponds to the correct results belonging to unhealthy class.The value between parenthesis corresponds to the Jaccard index. 49
5.17 The visualized results in the experiment, which corresponds to the incorrect predictions belonging to normal class.The value between parenthesis corresponds to the Jaccard index. 50
5.18 The visualized results in the experiment, which corresponds to the correct results belonging to normal class.The value between parenthesis corresponds to the Jaccard index. 50

List of Tables

3.1 Summary of the performance metrics of the studies. 19
3.2 Summary of the performance metrics of the studies (continuation). 19

5.1 Number of the patients after the split of the dataset. 36
5.2 Number of the images after the split of the mixed dataset. 36
5.3 Average results on the test set for Faster R-CNN ResNet-50 FPN model. 42
5.4 Detection performance of our model. 42
5.5 Number of the patients after the split of the dataset. 46
5.9 Performance metrics of the AlexNet model. 48

Abbreviations and Symbols

AI	Artificial Intelligence
AJCC	American Joint Committee on Cancer
ANN	Artificial Neural Network
AUC	Area Under the Curve
CAD	Computer-Aided Diagnosis
CAS	Computer-Aided Surgery
CNN	Convolution Neural Network
CPTAC	Clinical Proteomic Tumor Analysis Consortium
CPU	Central Processing Unit
CT	Computed Tomography
DICOM	Digital Imaging and Communications in Medicine
ERCP	Endoscopic Retrograde Cholangiopancreatography
GPU	Graphics Processing Unit
HNN	Holistically-nested Convolution Networks
HU	Hounsfield Units
IPMN	Intraductal Papillary Mucinous Neoplasms
MIP	Maximum Intensity Projection
MRI	Magnetic Resonance Imaging
MSD	Medical Segmentation Decathlon
NIfTI	Neuroimaging Informatics Technology Initiative
NIH	National Institutes of Health
PDAC	Pancreatic Ductal Adenocarcinoma
PET	Positron Emission Tomography
PNET	Pancreatic Neuroendocrine Tumor
ReLU	Rectified Linear Unit
ROC	Receiver Operating Characteristic Curve
SGD	Stochastic Gradient Descent
SVM	Support Vector Machine
TCIA	Cancer Imaging Archive
US	Ultrasonography
NCCN	National Comprehensive Cancer Network
DC	Dependencies Computation
BMI	Body Mass Index
VGG	Visual Geometry Group
ILSVRC	ImageNet Large Scale Visual Recognition Challenge
ROI	Region of Interest
YOLO	You Only Look Once
SSD	Single Shot Multibox Detector
FPN	Feature Pyramid Network
RPN	Region Proposal Network
COCO	Common Objects in Context

Chapter 1

Introduction

Human bodies are formed from trillions of cells that converge to form tissues and organs. The genetic material inside a cell regulates the activity of the cells. These genes are also responsible for the growth and division of cells.
Cancer cells are formed when there is a mutation in the gene of the healthy cells, giving rise to the abnormal growth of cells without any control. The gene mutations can be inherited or developed over the time or by engaging in some risk factors like smoking and alcohol. So, cancer is a group of diseases characterized by unregulated cell division, avoidance of cell death, tissue invasion and the ability to metastasize.
Known as a silent disease, and in particular, the pancreatic cancer is one of the most lethal types of cancer of the digestive system, detected frequently in advanced stages, making it with poor survival rates. The incidence of cancer varies greatly across the regions, which suggests that lifestyle factors play an important role in its etiology. Although the causes for the onset of the cancer are not fully understood, there are several factors that can increase the risk, including modifiable factors (e.g., obesity, alcohol, tobacco) and non-modifiable factors (e.g., age, gender, family history)[2]. An increase in the prevalence of these risk factors may lead to increases in incidence of cancer. Estimated number of annual cancer cases and mortality are expected to increase to 22 million of new cases and 11.2 million of deaths by 2025 [3].

1.1 Motivation

As the current incidence of pancreatic cancer is constantly increasing around the world, there is the need to reduce the mortality rate and to improve the survival rate of pancreatic cancer patients, along with diagnostic ability and cure technology. It is important to highlight that the disease may be present in the pancreas many years before patients become ill. There are no early detection tests for pancreatic cancer. The devastating outlook of this disease emphasizes the need for new strategies to aid in early diagnosis and staging. The pancreatic cancer is usually asymptomatic in the early stages and without the use of invasive procedures, the current screening methods are unable to achieve an effective early diagnosis [4]. When appear symptoms, generally, is associated to

the presentation in an advanced stage of the pancreatic cancer, turning the patient incurable. Radiologic high-resolution cross-sectional imaging by Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) are essential and represents the advanced techniques for early lesion detection, pre-therapeutic patient staging and therapy response evaluation [5, 6].
The Artificial Intelligence (AI) as tools for risk stratification have the potential to change the detection landscape. Innovation and research is this area improve understanding of this complex type of cancer, identify the right tools to achieve early detection and reverse the survival rates of patients [7]. Medical imaging analysis is an important requisite for many clinical applications, such as computer-aided diagnosis (CAD), computer-aided surgery (CAS) and therapies or treatments and play a role in modern healthcare diagnosis and procedures. Deep learning - based techniques has achieved remarkable progress in many domains, due to the development of intelligent medical systems for healthcare has gain great research attention. Towards medical machine intelligence there are some critical challenges:

- Detection of small targets, such as the pancreas from medical images due to their low resolution and poorly defined margins, which makes it easily confounded by the complex background.
- Deep learning models, generally, depends on large and well annotated datasets to achieve high performance. However, annotating medical images demands clinical experience and manual labor, which difficult to obtain high-quality datasets.

Medical imaging algorithms has been proposed for decades, but almost none have been integrated into clinical systems. Consequently, clinicians are frequently required to demarcate regions of interest manually for a diversity of applications, such as volumetric measurements of a tumor or treatment planning and endlessly others. In many cases when the medical data are publicly accessible, the labels are rarely available due to the lack of experts to annotation regions of interest, impairing the performance of the algorithms. It is clear that the significant progress will be the result of strategically designed collaboration among researchers and institutions from multidisciplinary backgrounds and support by committee funders.

1.2 Objectives

According to Lulu Zhang *et al.*[8] the poor survival rates of the pancreatic cancer have not changed significantly in nearly 40 years. Some characteristics important for the diagnosis of the pancreatic cancer are subtle or even go unnoticed by some experienced radiologists. The study of pancreas is a challenging task because this organ changes in size, shape and abdominal location between patients and the fat around the pancreas (visceral fat) can cause large variations in contrast along its boundaries. The main goal of this study is to contribute to a better knowledge in the field of pancreatic cancer, respectively of Pancreatic Ductal Adenocarcinoma. There is a need for an automatic method that is able to perform this assessment more easily by receive the medical images as input because despite the devastating statistics caused by pancreatic cancer, the Portuguese

National Cancer's Plan [9] not include the pancreatic cancer and there are no screening programs. To accomplish this, the main goals of this project are:

- Propose a method to reduce the region of interest in the abdominal CT images, decreasing the complexity of the background of the images and simplify the application of a deep neural network to obtain higher accuracy values.
- Investigation of deep learning approaches for Pancreatic Ductal Adenocarcinoma detection.

1.3 Contributions

This dissertation will mostly focus on the development of deep learning-based method for region of interest delimitation that includes the pancreas and the Pancreatic Ductal Adenocarcinoma classification.The contributions for these fields are:

- a study that shows that using the radiology knowledge allows improving the preprocessing process of the CT scans, using the parameters presented in DICOM format to improve the performance of the deep learning models.
- a deep learning-based system for automatic region of interest delimitation that cover the entire pancreas.
- a study that shows the possibility to identify the Pancreatic Ductal Adenocarcinoma.

1.4 Document Organization

In addition to the Introduction, this document is organized as follows:
First, Chapter 2 presents an explanation of the basic concepts needed to understand this study, related with anatomical context of the problem in study, some important aspects regarding the pancreatic cancer as well as the fundamentals of CT imaging.
Then, Chapter 3 describes the state-of-art methods for the detection, segmentation and classification of pancreatic cancer in CT scans. To end the chapter, is presented a summary of the datasets used in this project.
The Chapter 4 introduces the fundamental concepts of biomedical images processing and the methods and techniques most commonly used in the literature under the scope of this dissertation. This chapter describes some fundamental concepts related to Machine Learning, followed by the presentation of some aspects related to Deep Learning, such as basic CNN concepts and the most widely used CNN architectures relevant for this work. The chapter ends with the description of the performance metrics for evaluating the quality of the models.
The Chapter 5 presents the proposed PDAC classification algorithm. The method is composed by two steps, respectively, the ROI delimitation and PDAC classification. The results obtained in each step are also presented.

Finally, the Chapter 6 summarizes the main contributions of the performed research and suggests the future work to further improve the present method.

Chapter 2

Fundamentals of Pancreatic Cancer

This chapter aims to give an overview of the basic concepts needed to understand this work. In the following sections, the anatomical context of the pancreas and its relationship with other organs will be explored, along with the description of the functions of pancreas, the CT protocol and the staging scale for the study of the pancreatic cancer.

2.1 Anatomical Context

2.1.1 Gross Anatomy

The pancreas lies in an oblique position and is located behind the stomach. It is a retroperitoneal organ extended across the posterior abdominal wall from the duodenum on the right to the spleen on the left, positioned at the level of 12th thoracic vertebrae (T12) to 1st or 3rd lumbar vertebrae (L1 – L3) [10] . This organ is a large accessory gland of the digestive system covered with fine connective tissue but not have a true capsule such as, for example, the liver. The shape of this organ is compared to the letter "J" turned sideways, with the loop of "J" around the duodenal C-loop [11].

2.1.2 Pancreatic Divisions and Anatomic Relationships

The pancreas is differentiated into five regions: head, uncinate process, neck, body and tail.

- **Head**

  The head of the pancreas lies within the C-shaped duodenum curve, and it is the widest portion of the pancreas. Several major vascular structures and significant landmarks in vascular anatomy passes near to the pancreas. The inferior vena cava, right renal artery and both renal veins lies behind the head of the pancreas.

- **Uncinate process**

  The uncinate process is a projection arising from the lower part of the head of the pancreas (inferomedial extension). This structure lies between two major vascular vessels, namely, the superior mesenteric vessels and aorta.

- **Neck**

  The neck of the pancreas is a short structure that connects the head to the body. This portion of the pancreas is located in front of superior mesenteric vessels and behind passes the superior mesenteric and splenic veins, which form the portal vein.

- **Body**

  This part of the pancreas is elongated and extends from neck to the tail of the pancreas, located behind of omentum lesser sac and in front of aorta, left adrenal gland, left kidney and renal vessels.

- **Tail**

  The tail of the pancreas is located close to the hilum of the spleen and in front of the left kidney, and it is contained within the splenorenal ligament and splenic vessels. It is the only part of the pancreas that is intraperitoneal.

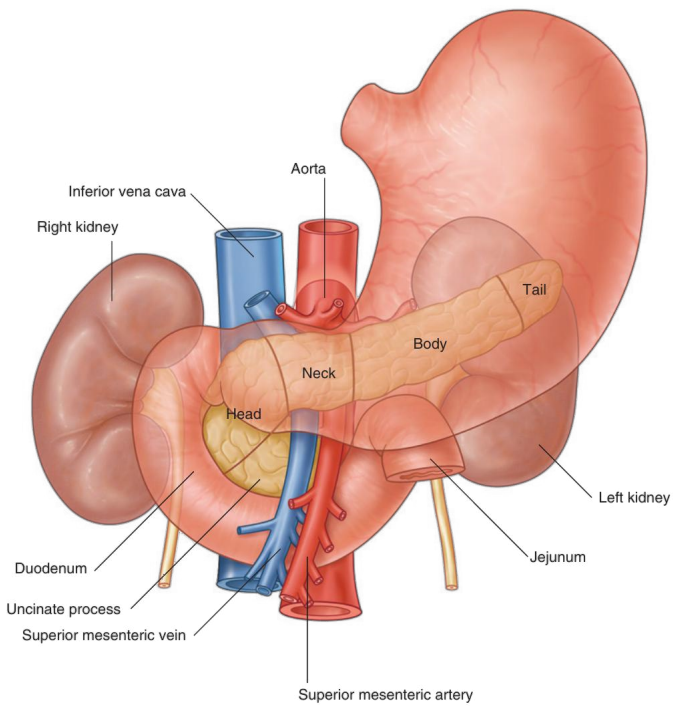

Figure 2.1: The pancreas.Image from [10].

2.1.3 Pancreatic Ducts

The pancreas has a soft, lobed appearance and is organized around a central pancreatic duct called the main pancreatic duct or Wirsung duct. This duct extends from the tail to the head of the pancreas. After entering the head of pancreas in the lower part of the head joins to the bile duct, formed the hepatopancreatic ampulla or ampulla of Vater, which enters to the descending part of the duodenum at the major duodenal papilla. Surrounding the ampulla, there are a collection of smooth muscles that regulates the passage of the digestive fluids, respectively the bile and pancreatic juices, and prevents the reflux of gastroduodenal contents called sphincter of the Oddi. The normal diameter of the main pancreatic duct in the head is 3-4 mm, in the body is 2-3 mm and in the tail is 1-2 mm. The diameter of the main pancreatic duct tends to increase with age and with presence of pancreatic disease [11].

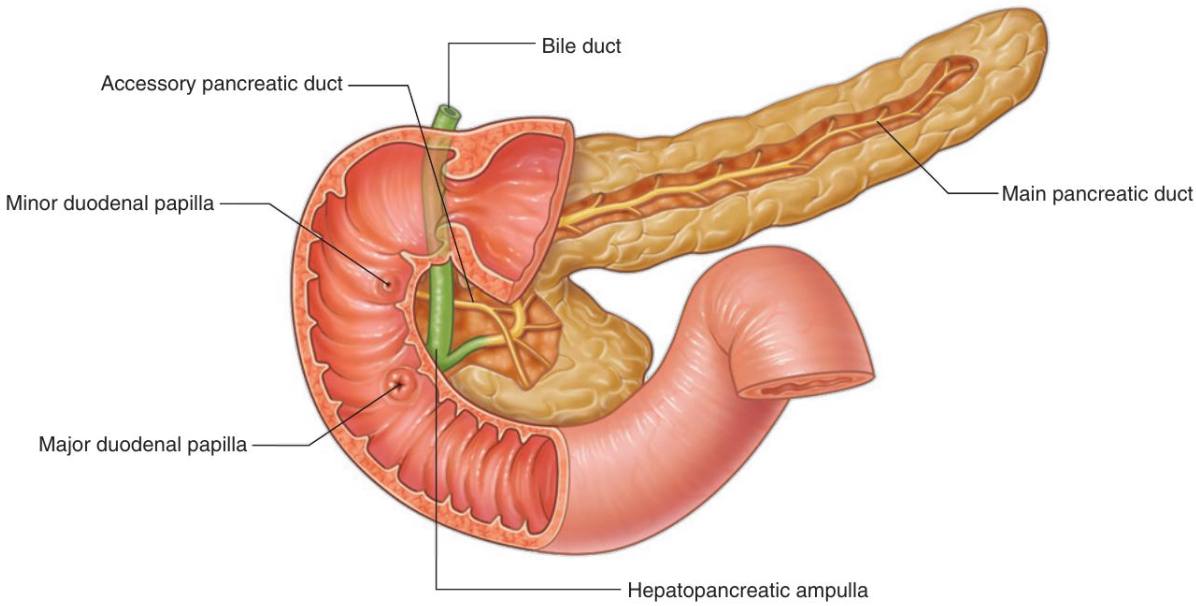

Figure 2.2: The pancreatic ducts.Image from [10].

There is another duct designated by accessory pancreatic duct or Santorini duct, which drains a small portion of the upper part of the head of the pancreas. It drains to the second part of the duodenum (descending part) and form the minor duodenal papilla. Most of the time, this duct is not identified in CT or MRI[10, 11]. Both ducts usually communicate with each other.

2.1.4 Arterial and Venous Blood Supply

Pancreas has a very rich arterial circulation and is supplied by two branches of aorta [11, 12] :

- Celiac artery
- Superior mesenteric artery

The head and the uncinate process of the pancreas are supplied by superior and inferior pancreaticoduodenal arteries, which are branches of the superior mesenteric artery. Each pancreaticoduodenal artery is divided into anterior and posterior branches that project along the pancreas. The body and tail of the pancreas are supplied by the pancreatic arteries, originated by the splenic artery, gastroduodenal artery and superior mesenteric artery. The venous drainage of the head of the pancreas create a network around the head and follows a branching pattern similar to the arterial supply. In the pancreatic venous system, it is important to call the attention to the junction of the superior mesenteric vein and the splenic vein, which form the portal vein. Their relationship with the pancreas are the most important for the long-term prognosis of pancreatic neoplasms and in pancreatic surgery. In the body and tail of the pancreas, the venous drainage is more variable, but consists of multiple branches draining into splenic vein [13].

2.1.5 Lymphatic System of the Pancreas

This organ has an extensive network of lymphatic channels that interconnect and draining to the surrounding lymph nodes. This fact explains why patients with pancreatic neoplasms frequently have positive lymph nodes at the time of the initial diagnosis and have a high incidence of local or distal metastasis. The location of the cancer in the pancreas indicates the spread and the lymph nodes involved.

2.1.6 Functions of the Pancreas

The pancreas is an organ with exocrine and endocrine functions. The main function of the exocrine pancreas is to aid the digestion of the nutrients from the diet, because of most of the pancreas produces a juice composed of enzymes of the food ingested. There are a small group of cells in the pancreas, designed by pancreatic islets. These islets produce hormones, such as glucagon and insulin, and other digestive enzymes that maintain the proper level of sugar in the blood and the adequate metabolism of the human body [12, 14].

2.2 Pancreatic Cancer

The term “pancreatic cancer” contains distinct types of cancer that are frequently confused, which can appear from exocrine cells or endocrine cells. Tumors of exocrine cells represents more than 95 % of all pancreatic cancers and is composed by different histological subtypes, for example the adenocarcinoma. Adenocarcinoma is the most common type of pancreatic cancer. Endocrine

tumors could be malignant or benign and can be classified in functional, which produce hormones, and nonfunctional that is producing non hormones. This type of tumors have better prognosis than exocrine tumors [14].

2.2.1 Mortality Rate and Significance of Early Detection

Pancreatic cancer remains a devastating disease across the world. According to Shallu Midha *et al.* [3] estimated, "annual cancer cases and mortality are expected to increase to 22 million of new cases and 11,2 million cancer-related deaths by 2025". Among all cancers, the pancreatic cancer is associated with a poor prognosis and a high case-mortality rate.
The American Cancer Society estimate for 2022 about 62,210 new cases of pancreatic cancer, of which 49,830 of the cases will be related to deaths in the United States of America [15]. Statistics have estimated in Portugal, that this disease in 2020 was responsible for 1792 new cases and almost the same number corresponds to deaths (1770 deaths) [16].

Despite all the progress achieved by the various medical disciplines who deal with this pathology, the general 5-year survival rate of pancreatic cancer is between 3% to 9% [7]. A 5-year survival rate means the percentage of people who will be alive for at least five years after the detection of cancer. These survival rates are established on many factors in combination with the distinct stage of disease when it is detected. The prognosis for the patients has been maintained similar in the last 40 years. If cancer is diagnosed in the early stages, then surgical eradication of the tumor is feasible and the 5-year rate would increase to 37%. If cancer has advanced to neighboring organs or tissues then the 5-year survival rate will be reduced to 12%. For those patients who are diagnosed when cancer has already advanced to different parts of the body the 5-year survival rate is 3% [17].

2.2.2 Challenges of Early Detection

Pancreatic cancer is challenging for diagnosing because there are no particular cost-efficient tests that can conveniently and accurately find early stage cancer with patients having negligible symptoms. Since the pancreas is located deeply in the abdomen, the presence of a tumor cannot be felt during a physical examination. The tests that analyze pancreatic cancer sometimes fail to detect pre-cancers, small lesions and early-stage cancer. So, there are no proven biomarkers that can detect the tumor early and consequently the pancreatic cancer is mostly not found until the last stages when it cannot be removed by surgery or has been metastasized. The initial tests for pancreatic cancer include knowing the medical and family history of the patient, risk factors associated with it and some common symptoms of pancreatic cancer like jaundice, swelling up of liver or gallbladder and gastrointestinal problems [11]. The jaundice is a medical condition in which the skin and the whites of the eyes turn yellowish hue due to the accumulation of excess of bilirubin, that occurs by the blockage of the bile duct by the tumor [2]. If these symptoms are positive, then various imaging exams will be done to examine for symptoms of cancer recurrence, areas that

might be suspicious for cancer, how far it is has spread and to check if treatment is successful or not.

2.2.3 Imaging Techniques

Therefore, remains a huge need for earlier diagnosis for this type of cancer, which can be result of unhealthy habits, for example smoking and alcohol consumption. For the diagnosis of this pathology, the pancreatic imaging is an essential tool. Frequently, requires the combination of different imaging modalities for the evaluation of the pancreatic parenchyma and the surrounding tissues. For pancreatic cancer screening and preliminary diagnosis is used the Ultrasonography (US), Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET) and Endoscopic Retrograde Cholangiopancreatography (ERCP). Each of these imaging modalities have advantages and disadvantages, and their clinical use is justified by the anatomic and functional characteristics of the pathology.

This dissertation focus precisely on the abdominal CT for the study of the pancreatic cancer. The CT is the modality of choice for the initial diagnosis and for evaluation of inflammatory and neoplastic condition of the pancreas, which have a sensitivity of 86% to 97% for PDAC but is less effective in diagnosing small lesions (inferior to 2 centimeters), especially the isodense lesions and lesions that do not produce a contour deformity or pancreatic duct dilation [18]. In pancreatic neoplasms, this imaging modality accurately depicts the tumor morphology, ducts anatomy and their relationships to surrounding organs and vascular structures. In inflammatory disease, this technique provides an excellent visualization of abnormalities in the pancreatic parenchyma and the spread of extrapancreatic disease. The CT images are produced by radiating selected portions of the body with a narrow beam of x-rays, oriented to a patient and quickly rotated around the body, which produce signals that are processed by a computer to generate cross-sectional images taken from different angles around the body. The 2-D pixels in each image correspond to 3-D voxels in the patient. These slices have different intensities based on the attenuation value of the tissues. The grayscale in CT images (see Figure 2.3) is a quantitatively meaningful value, called Hounsfield units (HU), in honor of Sir Godfrey Hounsfield, who was one of the principal inventor of CT technology in the early days [19].

According to the most recent 2021 National Comprehensive Cancer Network (NCCN) guidelines for patients with a clinical suspicion of pancreatic cancer or evidence of a dilated pancreatic and/or bile duct, a pancreatic protocol CT is the primary imaging modality of choice [20]. The dilation of both the pancreatic and common bile duct (double duct sign) is highly suggestive of an obstructing pancreatic head tumor, and it is present in about 62% to 77% of the patients with a cancer in the pancreatic head. An abrupt change in the caliber of both ducts can be a sign of early pancreatic cancer [21].

A pancreas CT protocol (see Figure 2.4) can vary between institutions but normally include an intravenous contrast injection and is acquired in three phases of contrast (arterial phase, pancreatic phase and portal venous phase), with pre-contrast imaging and with thin-section imaging and multiplanar reconstructions. The arterial phase is acquired 20 seconds after the contrast injection

and is used, for example, to the detection of vascular lesions. The pancreatic phase is acquired normally 35 to 50 seconds after contrast injection for optimal evaluation of pancreatic parenchyma and for assessment of peripancreatic arterial anatomic structures. The portal venous phase is acquired 60 to 90 seconds for optimum assessment of the venous anatomical structures and their involvement, along with the detection of hepatic and distant metastases [22, 18, 23, 21] .

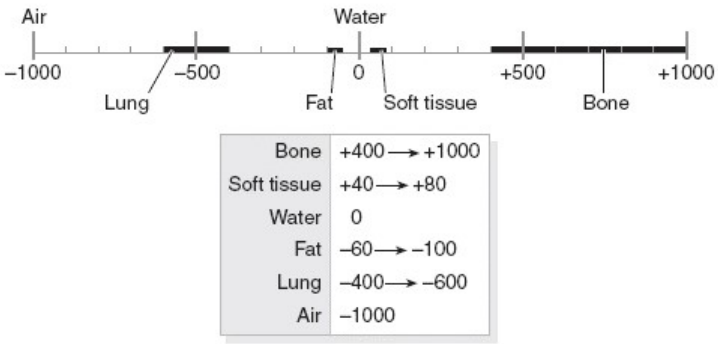

Figure 2.3: Hounsfield unit values for different types of tissue, from [24].

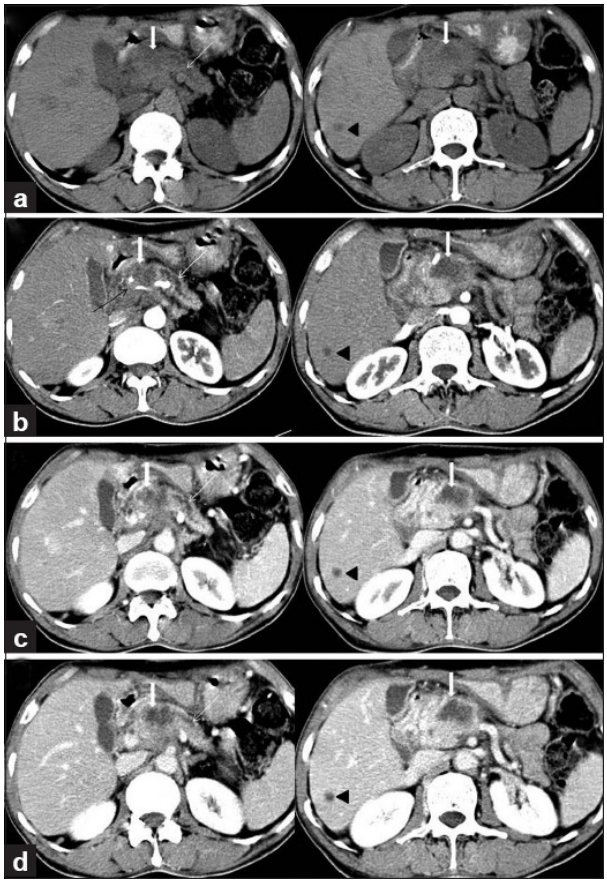

Figure 2.4: Pancreatic Adenocarcinoma. a) Axial image without contrast injection and three-phase contrast enhancement CT of the upper abdominal acquired in arterial phase (b), pancreatic phase (c) and portal venous phase (d).Image from [22].

Special attention should be given, because the pancreatic adenocarcinomas tend to enhance less than the normal pancreatic tissue and contour abnormality are also seen. The accuracy of the diagnosis is improved if the CT is performed before the biliary stenting, which provoke artifacts and consequently difficult the identification of a tumor in the pancreatic head. Most of pancreatic cancers are located in the pancreatic head, followed by the body and tail. One study of 163 pancreatic cancers found 6,1% of tumors in the uncinate process, 98,7% in the pancreatic head, 15,3% in the body and 9,8% in the tail [25]. The thin sections in a CT protocol are important for image post-processing, including multiplanar reformatted images (coronal and sagittal anatomical planes), maximum intensity projection (MIP) images and volume rendering. The volume rendering images help to assess vessels in the optimal plane for evaluation of variant anatomic structures. The CT is the modality preferred for guidance method for percutaneous needle biopsy of pancreatic masses and for aspiration or percutaneous drainage of a pancreatic collection.

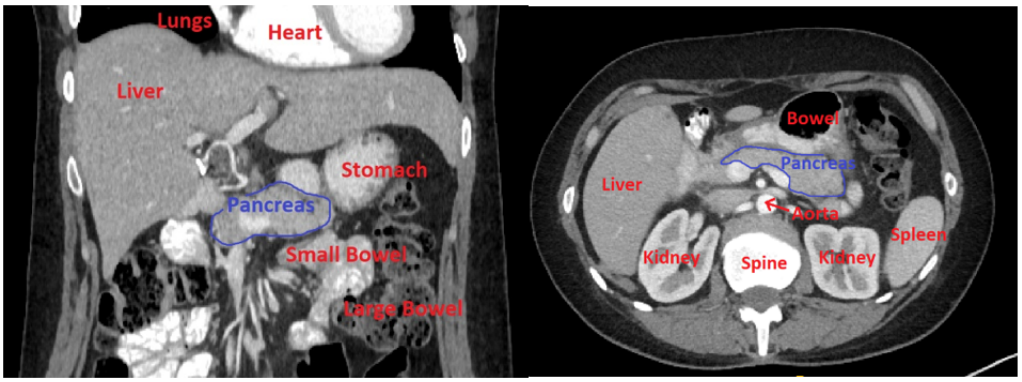

Figure 2.5: Anatomy of a coronal slice and axial view in CT. Image from [26].

In a standardized CT procedure, the equipment is calibrated, so the Hounsfield scale remains true, making it easier to segment different types of tissues and analyze different structures and details of interest for the characterization of the disease.

2.2.4 Pancreatic Cancer Staging

The staging is the method of describing where the cancer is located, where it has spread and whether it is affecting other parts of the body. Knowing the stage helps the clinicians recommend the best treatment and help to predict the patient's prognosis and identify clinical trials that may be treatment options.Normally, after the realization of a CT, the clinicians make the staging of the pancreatic cancer.

The NCCN clinical practice guidelines recommend a multidisciplinary approach when reviewing imaging and evaluating patients with pancreatic cancer for resectability. Radiology plays an important role in the multidisciplinary management of the patients with pancreatic cancer and in the assessment of treatment options. Preoperative imaging is used to characterize the pancreatic cancer on the following four categories[21]:

- **Resectable** - this type of pancreatic cancer can be surgically removed because the tumor may be located only in the pancreas or extend beyond it without invasion of blood vessels.
- **Borderline resectable** - describes a tumor that may be difficult or not possible to remove surgically when it is first diagnosed. If the chemotherapy or radiation therapy shrink the tumor first, the surgery may be possible to remove the lesion later.
- **Locally advanced** - this type of pancreatic cancer is located only in the area around the pancreas, but it cannot be surgically removed because it has growth into or close to nearby blood vessels or organs and the risk of damaging this anatomical structures is too high.
- **Metastatic disease** - The pancreatic cancer has spread beyond the area of the pancreas and to other organs, so cannot be removed completely with surgery (unresectable).

Surgical resection remains the main curative option, but most patients present with unresectable disease. The staging system of the American Joint Committee on Cancer (AJCC) is the commonly used for the staging of the pancreatic tumor, designed by TNM scale [23] (see Figure 2.6 a).
The T category corresponds to the tumor size and is assessed by measuring the largest diameter of the tumor in axial plane. Should be reported the relationship between the tumor and vessels and with the adjacent anatomical structures. The main goal is delineating the tumor extension to provide useful information for a potential pancreatic tumor resection. For that, the multiplanar reconstruction in CT is very helpful to visualize the circumferential relationship of the tumor with the vascular structures [23, 27] .
The N category refers to the degree of spread to the lymph nodes and is a challenge for clinicians because it is difficult to analyze microscopic metastatic disease in lymph nodes in all imaging modalities. The criteria for lymph node tumor infiltration suspicion include that the minor axis is greater than 1 centimeter, round morphology, central necrosis and heterogeneity [27].
The M category corresponds to the metastatic disease, and the evaluation of possible metastases is best done in contrast-enhanced CT. The most common sites of metastases are the liver and peritoneum [23, 27].
Then, are combined the T, N, M classifications to determine the stage of the cancer, which describe the cancer and allow planning the best treatment. The stages I and II are classified as resectable, with the absence of tumor contact with the adjacent vessels (celiac trunk, hepatic artery, superior mesenteric artery, superior mesenteric vein and portal vein). The stage III defined as a localized tumor with major vessels involvement and is subdivided in two categories, respectively, locally advanced unresectable and borderline resectable. Finally, the stage IV is defined by the presence of distant metastatic disease and precludes resection because patients will have no benefit from the surgery (see Figure 2.6 b).

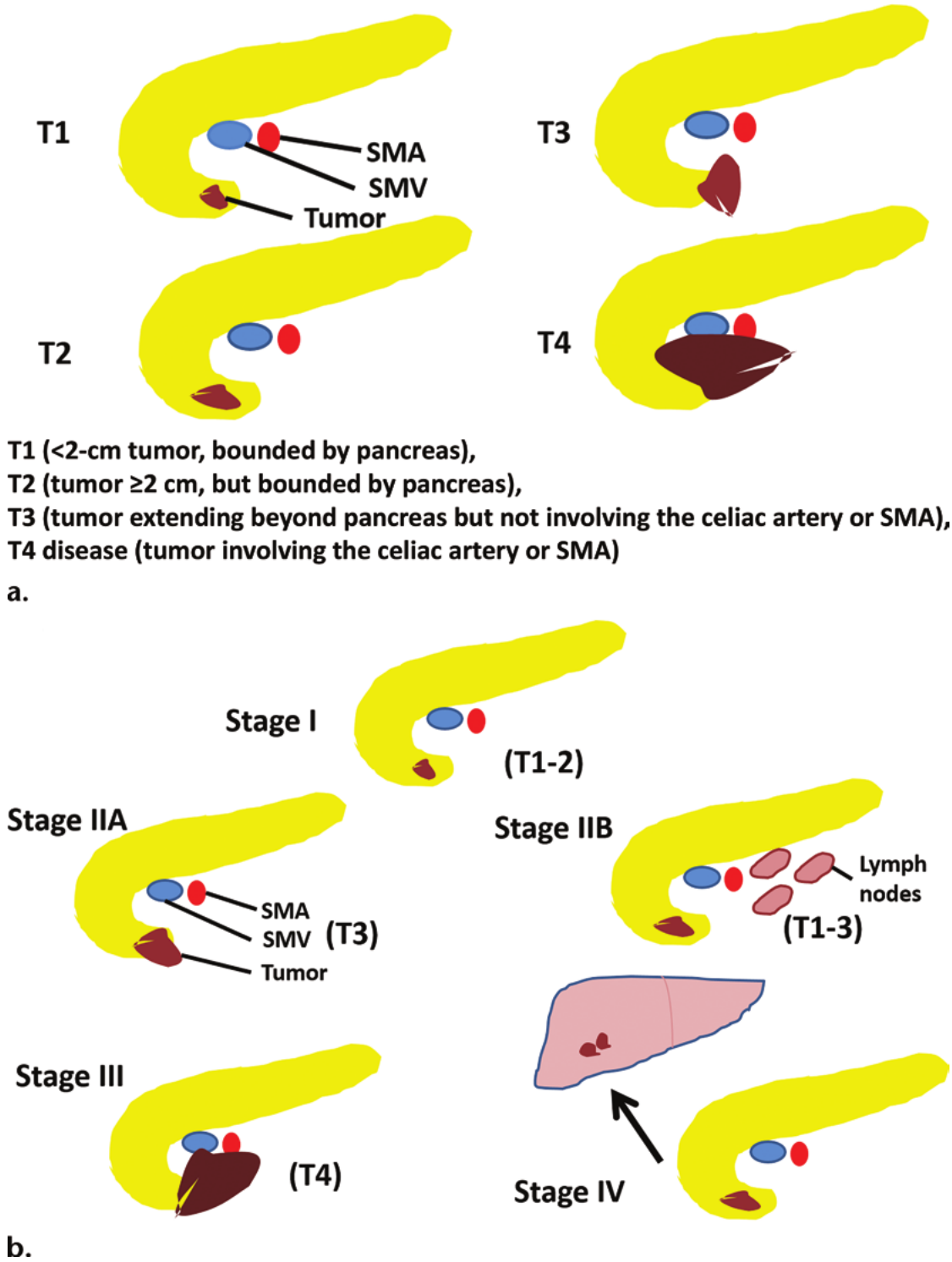

Figure 2.6: Tumor grading of the T category of the TNM scale (a) and figure depicting the stages of PDAC (b). Image from [23].

Chapter 3

Literature Review

This chapter reviews the current state of the art related to the detection, classification and segmentation of the pancreatic cancer as well as the main methods used to assess it. The present literature review intend to provide knowledge of what has been already studied in this field.
As stated in Chapter 1, in the pancreatic cancer is notorious the late detection and the high mortality rate. In medical images the differentiation in pancreas between benign disease (e.g.chronic pancreatitis) or malignant disease (e.g.pancreatic ductal adenocarcinoma (PDAC), intraductal papillary mucinous neoplasms (IPMN) and others) is challenging. However, it can be present in radiologic imaging, but the humans can not distinguish with the naked eye. Medical imaging analysis have three main tasks: disease diagnosis, lesions detection and lesions and organ segmentation. Each patient produces easily more than 300 images and the radiologist must discern the lesions from background and make reports to help another physicians. Sometimes, the fatigue, distraction and limitations of the Human Visual System can conduct to misdiagnosis.
So, some CAD systems based mainly in Deep Learning models have been created to help radiologists and physicians in making diagnosis from medical images. These systems have become one of the most research subjects in area of medical imaging, diagnostic radiology and Gastroenterology field. It is important to notice, that there are few papers on the pancreas comparing with other anatomical structures (e.g. lungs, liver and others). The reasons for it are probably due to the significance of the research on other anatomical structures is greater than that of the pancreas and due to the difficulty of detecting, classifying and segmenting the pancreas is much higher than in the liver and lungs. Next sections present the most representative works for these tasks and will be described to better contextualize the work presented in this dissertation. The results of these studies are presented in Tables 3.1 and 3.2.

3.1 Computer-Aided Diagnosis in Pancreatic Cancer

3.1.1 Pancreatic Cancer Detection and Segmentation

There are several recent studies in pancreatic cancer detection due to the most of the pancreatic cancer patients have been diagnosed with an advanced stage of the disease. The segmentation

of pancreas is a challenging task too because of the anatomical inter-patient variability of the pancreas in shape and volume. The study of Wazir Muhammad *et al.*[28] assessed the predictions of pancreatic cancer based on individual level incorporating 18 features in an Artificial Neural Network.

The Holger Roth *et al.* [29] study proposed a method to fill this gap based on two stage cascaded approach.

This method consist in two steps. The first step consists in a fully deep learning approach based on the application of Holistically-nested Convolution Networks (HNN) on three orthogonal views (axial, sagittal and coronal) to localize the pancreas from the whole image in CT scan. In the second step is used the bounding-boxes computed previously and integrates semantic mid-level cues of deeply learned organ.

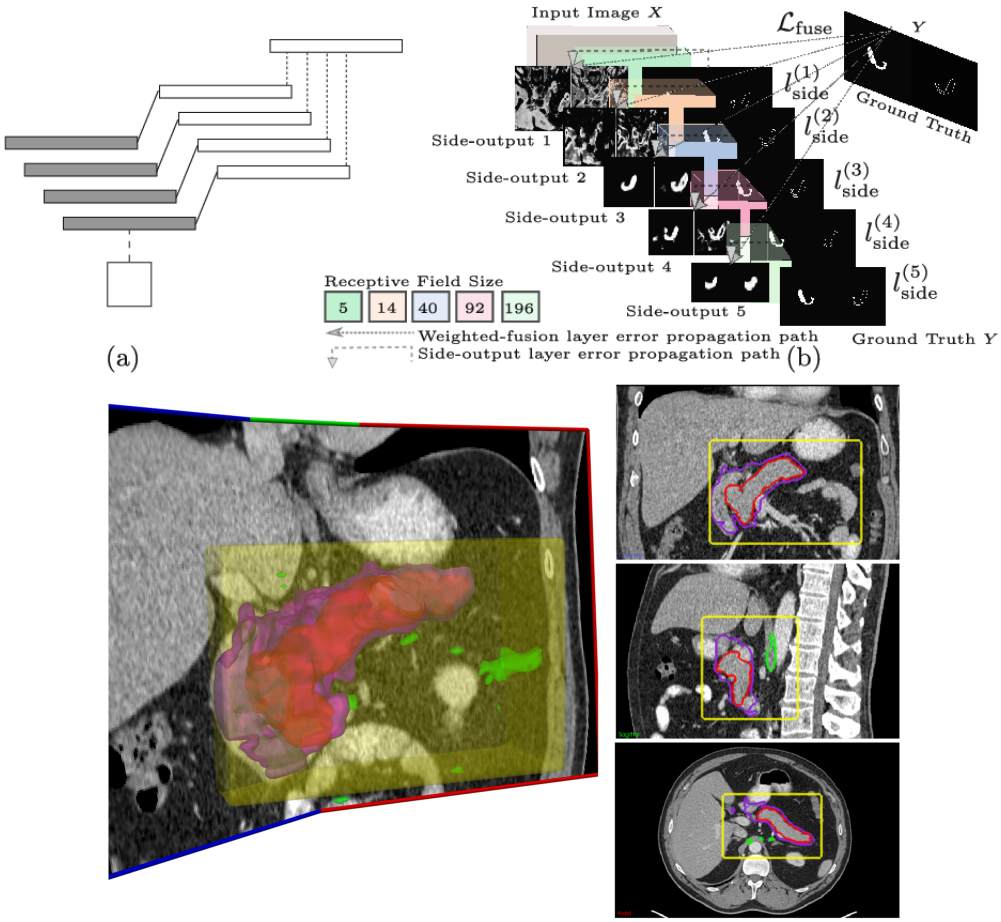

Figure 3.1: Illustration of the HNN architecture and candidate bounding box region generation.Image from [29].

Deep Learning approaches are the most commonly used to assess these tasks. Another study Mizuho Nishio *et al.* [30] combines data augmentation with two deep learning architectures.The authors proposed a baseline U-net and a Deep U-net, that is, the number of feature maps in the convolution layer was doubled after each downsampling.

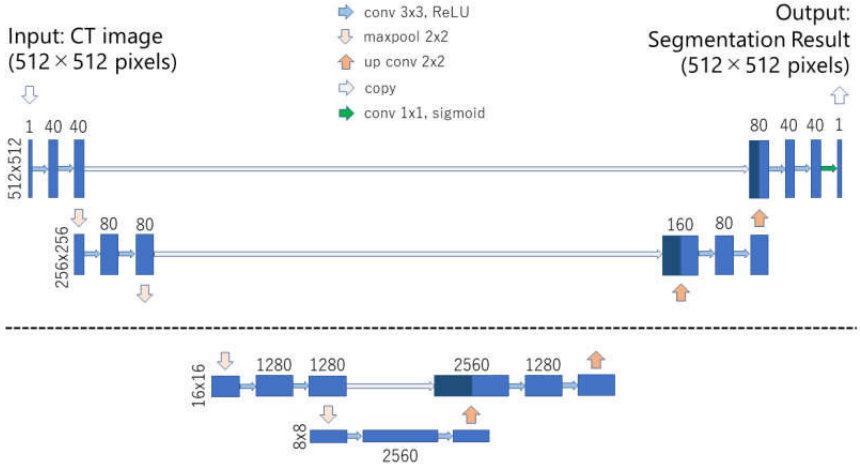

Figure 3.2: Illustration of the deep U-net model. **Abbreviations:** convolution layer (conv), maxpooling layer (maxpool), upsampling and convolution layer (up conv), rectified linear unit (ReLU).Image from [30].

3.1.2 Pancreatic Cancer Classification

Pancreatic cancers can be divided, as stated in Chapter 2 into endocrine or exocrine tumors. Each category has several cancer types, justified the large amount of studies in that area. The study of Fengze Liu *et al.* [31], in a 2018 paper, have the main goal to detect PDAC in abdominal CT scans and proposed a two-stage framework, which first step is to train an encoder-decoder network for voxel-wise pancreas segmentation in CT scans (segmentation model). The second step is to use the joint shape representation and classification network to predict if the patients suffers from PDAC (mask model).

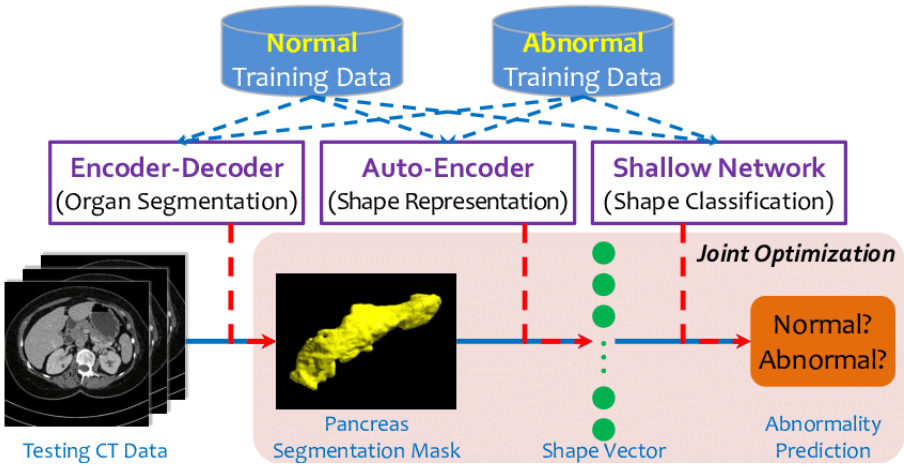

Figure 3.3: Framework of the approach of the authors, from [31].

Another study Po-Ting Chen *et al.* [32] in this field and with the same objective, distinguish between pancreatic cancer and noncancerous pancreas in CT images. This study used three datasets,

respectively, the first dataset with 436 PDAC patients and 479 healthy controls for training and testing the model, the second is composed by 100 patients with small PDAC and 100 healthy controls patients and the third dataset contains 182 PDAC and 82 healthy controls patients.The second and third dataset was used only to test the model. The authors processed the images into patches and applied a XGBoost model to classify the patches.Then, was performed a patch-based and patient-based analysis.

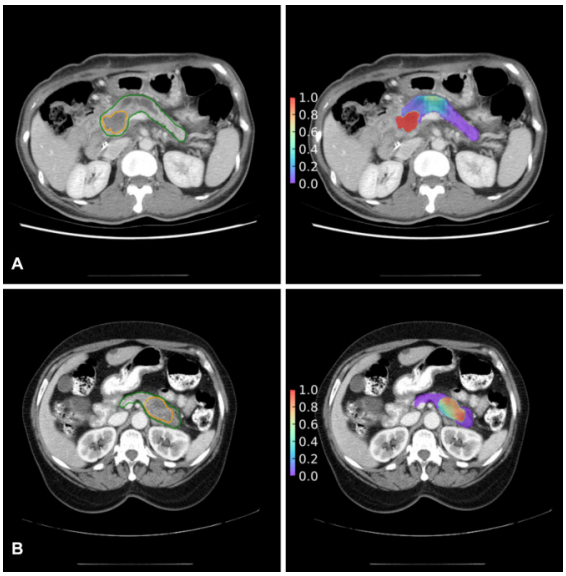

Figure 3.4: Results of the authors. Correspondence in tumor location between model prediction (**right panels**) and radiologist's labeling (**left panels**), from [32].

The Zhengdong Zhang *et al.* [33] study proposed an advanced framework to evaluate pancreatic cancer in CT imaging. The method proposed is composed by three components: 1) Augmented Feature Pyramid networks, 2) Self-Adaptive Feature Fusion and 3) Dependencies Computation (DC) Module. The first method is used to fully extract and propagate low-level accurate location information, the second method can encode the context information at multiple scales and the third method capture the iteration information between predictions and surrounding tissues.

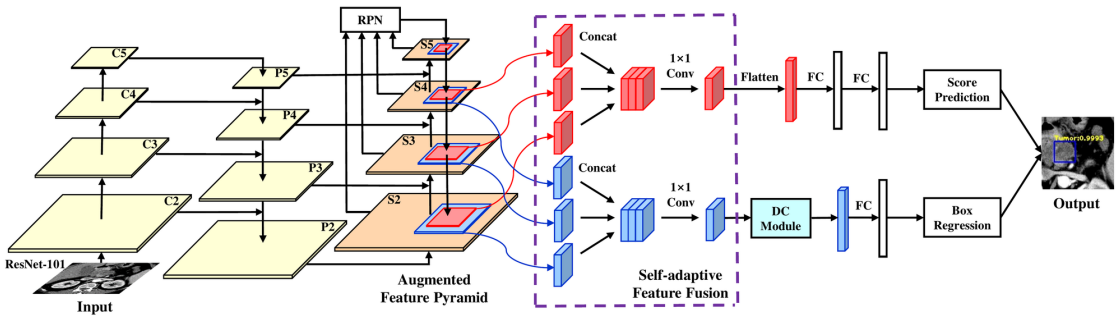

Figure 3.5: The architecture of the pancreatic tumor detection network, from [33].

The Han Ma *et al.* [34] propose a method based on the construction of a CNN to identify pancreatic cancer in CT. For that, it was used a dataset with 222 patients with pathologically confirmed pancreatic cancer and 190 patients with normal pancreas. The approach was evaluated in terms of a binary classification (i.e., cancer or not) and ternary classification (i.e., no cancer, cancer at tail/body, cancer at head/neck of the pancreas).
The Kao-Lang Liu *et al.* [35] propose a CNN-based method to distinguish individuals with and without pancreatic cancer on CT, compared with radiologist interpretation. The proposed method use contrast-enhanced CT images of 370 patients with pancreatic cancer and 320 patients with normal pancreas. The CT images were preprocessed into patches and a CNN was trained to classify the patches as cancerous or non-cancerous.

3.2 State-of-the-Art Results

Table 3.1: Summary of the performance metrics of the studies.

Reference	Dataset	Acc.	Sens.	Spec.
[28]	Private	-	80.7%	80.7%
[29]	Public (NIH-82)	-	-	-
[30]	Public (NIH-82)	-	-	-
[31]	Private	-	80.2%	90.2%
[32]	Private	DT1 + DT2: 95% DT3: 86.5%	DT1 + DT2: 94.7%) DT3: 80.6%	DT1 + DT2: 95.4%) DT3: 100%
[33]	Private	90.2%	83.8%	91.8%
[34]	Private	95.2%	92.3%	97.9%
[35]	Private	83.2%	79.0%	97.6%

Table 3.2: Summary of the performance metrics of the studies (continuation).

Reference	Dataset	AUC	DSC Coeff.
[28]	Private	0.85	-
[29]	Public (NIH-82)	-	81.27%
[30]	Public (NIH-82)	-	Deep U-net: 70.3% - 78.9% Baseline U-net: 68.6% - 74.8%
[31]	Private	-	-
[32]	Private	DT1 + DT2: 0.98 DT3: 0.91	-
[33]	Private	0.95	-
[34]	Private	0.96	-
[35]	Private	0.92	-

Abbreviations: *NIH - National Institute Hospital, Acc.- Accuracy, Sens.-Sensitivity,Spec.- Specificity, AUC - Area Under the Curve, DT1 - Dataset 1, DT2 - Dataset 2, DT3 - Dataset 3, DSC Coeff.- Dice similarity coefficient or Sørensen–Dice coefficient*

3.3 Available Datasets

It is important to mention that a limited factor in the study of pancreas and development of strategies is the lack of publicly available datasets with the corresponding labeled images. The evaluation and validation of CAD systems requires a large amount of data. This section describes the abdominal CT datasets used for the development of this dissertation, which present some quality gaps [36]

3.3.1 NIH - Pancreas CT Dataset

This dataset [37] consists of 82 portal-venous phase abdominal contrast-enhanced CT scans done at the National Institutes of Health Clinical Center (NIH). Subjects' have a mean age of 46.8 ± 16.7 years (range from 18 to 76 years). All the patients have morphologically normal pancreas. Seventeen of the subjects are healthy kidney donors, who had been scanned prior to nephrectomy. The remaining 65 patients were selected by a radiologist from patients who neither had major abdominal pathologies nor pancreatic cancer lesions. The dataset also includes pancreas segmentations that were verified/modified by an experienced radiologist. The time period of acquisition of the CTs and basic patient information are not specified. All the CT images are in DICOM format, but the provided segmentations are in NIfTI format. Evaluating the dataset, the image quality of the CTs are optimal, but some parts of pancreas have been excluded in the provided segmentations.

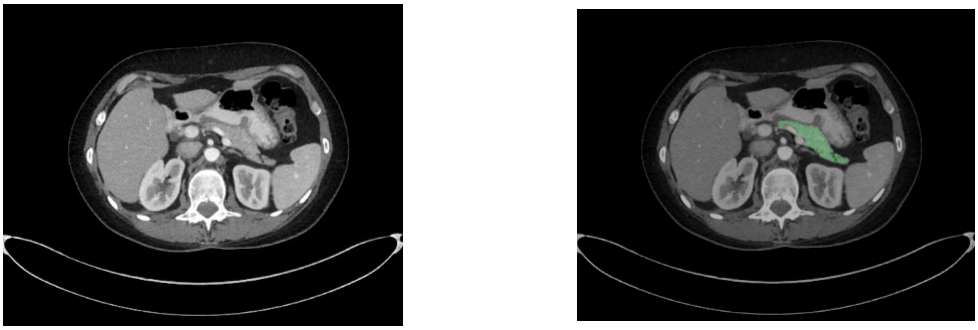

Figure 3.6: Axial volume scan (**left image**) and the same scan with the respective label (**right image**).

3.3.2 TCIA CPTAC-PDA Dataset

The TCIA CPTAC-PDA dataset [38] contains 98 patients from the National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium Pancreatic Ductal Adenocarcinoma (CPTAC-PDA) cohort. This dataset contains radiology and pathology images.
This radiology data is collected from standard-of-care imaging performed on patients immediately before the pathological diagnosis and from follow-up scans. For this reason, the radiology imaging data is heterogeneous in terms of imaging modalities and acquisition protocols.
It is available a supporting document with the metadata: demographics (age, gender, height,

weight, BMI, race, ethnicity), tumor pathology (tumor site, histologic type, grade, staging of the tumor, margin status and tumor regression in response of neoadjuvant therapy after being applied surgical resection) and other variables (alcohol consumption, tobacco smoking, date of last follow-up, survival duration and cause of death). However, information is only available on all patients on a few variables.All the images in the dataset are in DICOM format.

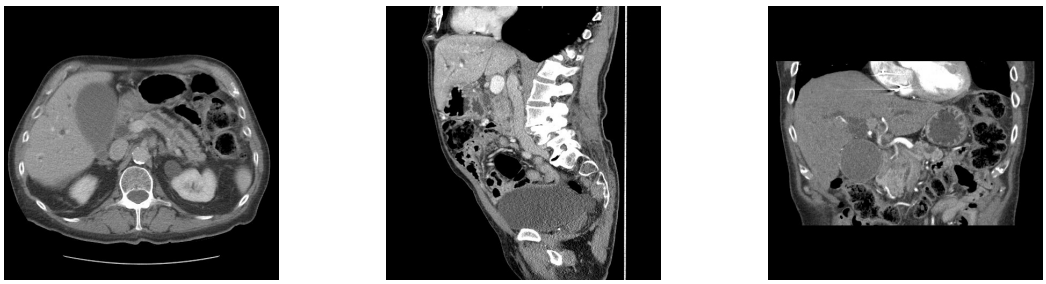

Figure 3.7: Axial scan (**left image**), multiplanar reconstruction in sagittal plane (**middle image**) and multiplanar reconstruction in coronal plane (**right image)** from TCIA CPTAC-PDA dataset.

3.4 Summary

After analysis of the state-of-the-art it is possible to infer that the majority of the literature uses private datasets and the researchers focuses mainly on deep learning-based methods. Focusing on data availability, there are few public datasets which have many quality gaps [36]. This difficult to other researchers to test their algorithms on openly available data. There are also a lack of deep learning-based studies in pancreatic cancer classification, detection and segmentation.

Chapter 4

Background

This chapter focuses on the fundamental concepts that are the essential building blocks of any Machine Learning system. First, the chapter starts with a summary of some pre-processing techniques, an overview of the Machine Learning and then an overview of the Deep Learning. The chapter ends with Convolutional Neural Networks and some common architectures relevant to this work and finally is presented the evaluation metrics.

4.1 Biomedical Image Processing and Analysis

Medical imaging is an area that play a central role in Medicine for clinical diagnosis and treatment of several diseases and for support of medical interventions. The X-ray radiography became the first diagnostic modality. Since then, many imaging modalities were invented, such as CT, MRI, US, PET/CT and more imaging procedures. From that, the clinicians can extract many image information's, which have an essential role in decision-making in several ways, respectively, in detection, characterization, staging, guiding interventional procedures, treatment response and in patient care process.

The number of images to be interpreted per patient and the workload for radiologists and other physicians increased, making difficult the clinical decisions to patient care process. So, the development of methods and techniques in scientific areas, such as Machine Learning and Deep Learning for medical image analysis and the development of CAD systems for second opinion to assist radiologists in images interpretation is essential.

For an accurate CAD system it is important to apply some image processing techniques, because sometimes the medical images have low resolution, noise or degradation from image acquisition. The image processing can preprocess the images and convert it in acceptable results for further analysis. Nowadays, we can perform algorithms to recognize patterns in data and make predictions, however the challenge in AI is how to get this informal knowledge into a computer.

4.2 Pre-processing Techniques

There are some image transforms that can be applied to the medical images, for example point operators, geometric transformations and global optimization methods.
The point operators are the simplest method of image processing transforms. In this kind of operators, each output pixel's value depends on only the corresponding input pixel value. Some examples of these operators include the brightness and contrast adjustments and color correction. However, can be applied geometric transformations to the images which "modify the spatial arrangement of pixel in an image" [39] and includes rotations, translations and scaling operations.
All image acquisition systems have an inherent noise. The presence of noise makes the images unclear and decrease the images quality, influencing posteriorly an adequate diagnosis. Denoising methods is an important and essential step in pre-processing techniques for further analysis [40].
The smoothing spatial filters reduce the sharp transitions in intensity and can be classified in two types, respectively, the spatial domain filters and transform domain filters. The first group includes, for example, the median filters, which performs well in several images when have high-noise levels. In this filter is selected the median value from each pixel neighborhood. Another example is the mean filtering, but unlike the previous filter it blurs the edges of the images because it is uniformly distributed the noise in the images.
The transform domain techniques are able to distinguish between noise and signal. One example is the wavelet transform, that are used in Computer Vision applications. These filters are multiresolution representations of signals and images and delimit a signal in space and frequency [41]. The sharp transitions in images are preserved, important in image filtering.
For medical image enhancement, which aims to highlight important features from medical images to increases the visual perception, there are two categories of enhancement methods: spatial domain methods and frequency domain methods. The spatial domain methods refer to the manipulation of values of pixels in the images to achieve an enhancement. Some examples of operations that are included in this category are equalization histogram, power-law transformations, images negatives and endlessly others. In frequency domain methods, the image is transformed in frequency domain by the Fourier Transform.
The image processing also comprises techniques for image segmentation. Image segmentation refers to the process of the image partition in a set of disjoint regions. In medical image field corresponds to the separation between anatomical structures and background. The methods for image segmentation can be grouped into feature domain techniques and image domain techniques.
In feature domain techniques includes the clustering and thresholding methods. Thresholding method is the simplest method that from a grayscale image can create binary images using a threshold value, which setting the pixels to white when the value is above a certain threshold value and to black when the value is below the respective threshold value [42]. The clustering methods identifies similar groups of data in a dataset, for example the K-means algorithm (see section 4.3.2.1). The image domain techniques consist in region-based techniques and boundary-based methods.

- **Region-based techniques:** includes some methods, such as region growing, split and merge, and others. The region growing method starts growing from an initial seed point and from these grow regions appends new pixels/voxels based on predefined criteria to growth. The split and merge method divide the image into a set of disjoint parts and then merge or split the regions to satisfy the conditions of segmentation [39].

- **Boundary-based techniques:** includes for example edge-based methods and deformable models. The edge-based methods allow detecting edges in images. Edges occur at boundaries between regions of different colors, intensity or texture. Some examples of edge-based methods are Sobel and Canny operators [43].
  The deformable models [44] are curves or surfaces defined by the influence of internal forces (coming from the model) and external forces (computed from data). These two forces are defined to perform an object boundary. In the literature, there are two models: the parametric deformable models (e.g., snakes) and geometric deformable models (e.g., level-sets method).

4.3 Machine Learning

The Machine Learning has evolved as a subfield of AI and consists in self-learning algorithms to obtain knowledge from data and make predictions in the future. A machine learning model is composed by several steps, namely, data preparation, data pre-processing, feature extraction, feature selection, learning phase and validation and finally the testing phase.

Figure 4.1: A typical Machine Learning pipeline.

In the machine learning problems, is no way of knowing a priori which feature extraction algorithm is best for a given problem. The features are considered good to a problem analyzing their correlation or based on experimental approaches, which requires a long time to optimize the best features. The process of using domain knowledge and trying to derive useful features and combine some of them is called Feature engineering and matters for performance of the model. Then, is necessary to select the most relevant features to the problem, and this process is known as Feature selection. Both steps are correlated because the model only be able to learn if the training data contains enough features and data are not irrelevant. There are three main types of Machine Leaning, respectively: supervised learning, unsupervised learning and semi-supervised learning [45, 46].

4.3.1 Supervised Learning

The supervised learning learn a model from labeled training data that allow us to make predictions about unseen data, that is, from the features extracted from data (input) the algorithm learn a

function (f) and make predictions (output).

$$output = f\left(features\ of\ the\ input\right) \tag{4.1}$$

The problems in supervised learning can be grouped into:

- **Classification:** predict categorical class labels for new data based on past observations. The classifications problems can be binary classification task or multiclass, according to the kind of the problem to solve.
- **Regression:** the output variable is a real value. The main objective is to plot the best fit line or a curve between the data and finding correlation between the variables.

There are some popular examples of supervised learning algorithms, for example, the Linear Regression for regression problems, Random Forest for classification or regression problems and the Support Vector Machine for classification problem. These algorithms are described below.

4.3.1.1 Linear Regression

The linear regression is a mathematical modelling approach that can be used to understanding the relationship between input and output variables [47]. The representation is very simple and combines a set of input values (x) to the predicted outputs (y), by fitting a linear equation to observed data.

$$y = a + b.x \tag{4.2}$$

The linear equation 4.2 assigns a scale factor to the input known as slope (b) and is additioned to the intercept (a).

4.3.1.2 Random Forest

The Random Forest (figure 4.2) is a type of ensemble machine learning algorithm, which used the Bootstrap Aggregation that reduce the variance in algorithms with high variance, for example the decision trees [48]. The decision tree [49] are algorithms sensitive to a specific data on which are trained and are the building blocks of a Random Forest algorithm, with a structure similar to a tree. It is composed by three elements, respectively, the decision nodes, leaf nodes and root node [50].

In a decision tree is divided the training data into branches and then into another branches until a leaf node is attained. The majority of the decisions trees becomes the final output of the Random Forest system.

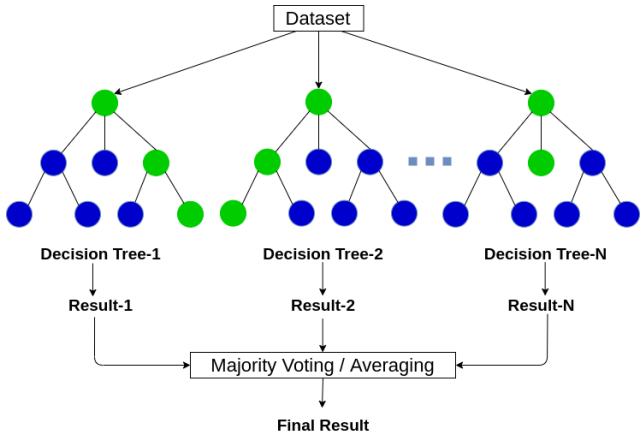

Figure 4.2: Random Forest algorithm. Image from [51].

4.3.1.3 Support Vector Machine

The Support Vector Machine (SVM) works with the definition of Maximal-margin classifier [48]. For example, consider a training set with examples that belong to two classes and the examples are linearly separable. Because of this linear separability, there are innumerous hyperplanes. A hyperplane is a line that splits the input variable space. The goal of this algorithm is to find a hyperplane that best distinctly classifies the data and maximize the margins between the hyperplane and the support vectors (the closest values to the classification margin). In this example, the SVM is a linear classifier (see figure 4.3).

For examples that not linearly separable this algorithm can be used, by mapping the training points to a higher-dimensional space that become linearly separable with the contribution of the kernel functions. There are some available kernel functions and can be used according to the problem to solve [52, 53].

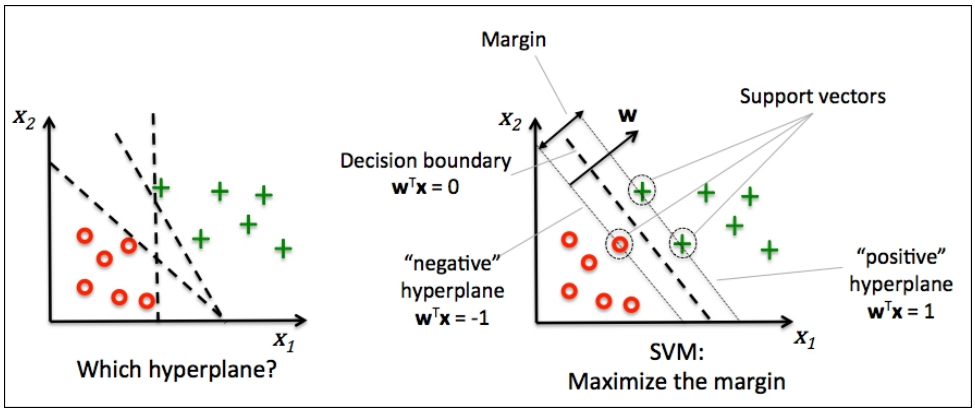

Figure 4.3: Support Vector Machine algorithm, from [54].

4.3.2 Unsupervised Learning

The unsupervised learning discover patterns and learn the structure of unlabeled data, that is, extract knowledge without the guidance of a known outcome [46, 47, 48]. The problems in unsupervised learning can be divided into:

- **Clustering:**is an exploratory data analysis technique that allow us to structuring information into groups,analyzing similarity of the features and establish relationships from the data.
- **Association:**pretend to discover rules that describe large amounts of data.

One example of unsupervised learning algorithm is the K-Means, which is described below.

4.3.2.1 K-means Algorithm

The K-means is one of the most popular clustering methods.

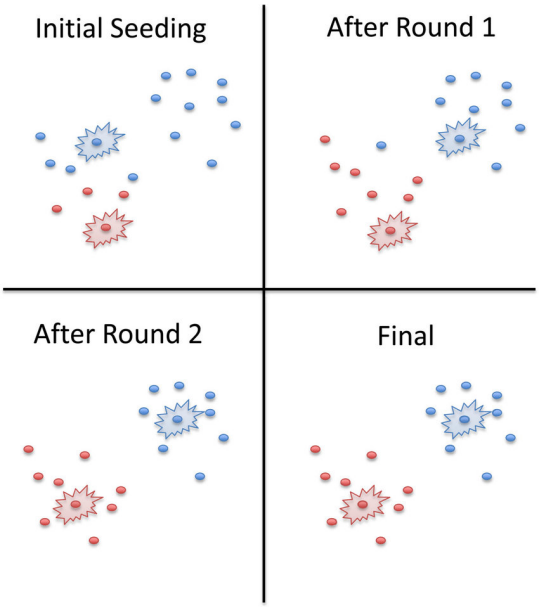

Figure 4.4: K-means example (K = 2). Clusters are distinguished by color, from [55].

For the initialization, the first task is to form the initial K clusters and divide the data points randomly into k subsets. Then the basic idea, is relocating each point to its new nearest center based on the Euclidean distance, update the clustering centers by calculating the mean and repeat the process until convergence criteria are satisfied [46].

4.3.3 Semi-supervised Learning

The semi-supervised learning is an approach that combines a small amount of labelled data and a large amount of unlabeled data. The algorithm would analyze the labelled data and try to label each of the unlabeled data points. This process can be repeated iteratively until the algorithm learns the structure of the data and make accurate predictions [47].

4.4 Deep Learning

In the past few years, the Deep Learning techniques have achieved very good results in challenges that conventional Machine Learning models not solve. The Deep Learning is a subfield of Machine Learning that belongs to the family of Artificial Neural Networks (ANN) algorithms. The term "deep" is representative of the numerous number of layers.
For many years, the design of pattern recognition or traditional ML models required careful engineering (meaningful feature extraction) and domain expertise to develop a system capable to transform the raw data into a representation that can be used to classify or detect patterns.
The Deep Learning has relieved some obstacles, eliminating the Feature extraction and Feature selection steps [56, 57]. The Feature Engineering step, now change from a human-side to a computer-side.
This is possible due to:

- availability of a huge amount of data;
- development of learning algorithms;
- advancements in central processing units (CPU) and graphics processing units (GPU), which allow us to run much larger models today.

The Deep Learning have a high capacity of learning hierarchical feature representations of data, that is, the concepts in the lower level layers of a neural network refers to basic representations and in a higher level forms abstract concepts.

4.4.1 Structure of an Artificial Neural Network

A Deep Learning model is composed by a set of layers (see Figure 4.5). The first layer is the input, which contains the variables that we are able to observe. Then, there are a series of hidden layers to extract features from the inputs.
Finally, the deep models are composed by the output layer that produce the results for a given input. Supervised learning is commonly used in Deep Learning. In this technique there are, generally, thousands of parameters and the main idea is to minimize the error between the outputs and the ground truth.
To update the weights, one of the most used method is *Stochastic Gradient Descent* (SGD), which consists in an iterative process that compute the outputs, errors, the average gradient and the update of the weights [46, 47]. The result of the weighted sum of the inputs from the previous layer is passed by a non-linear function or activation function, such as sigmoid function, rectified linear unit function (ReLU), hyperbolic tangent, and others. After the training phase, the performance of the model is measured in the test set (unseen data) to check the generalization ability of the system.

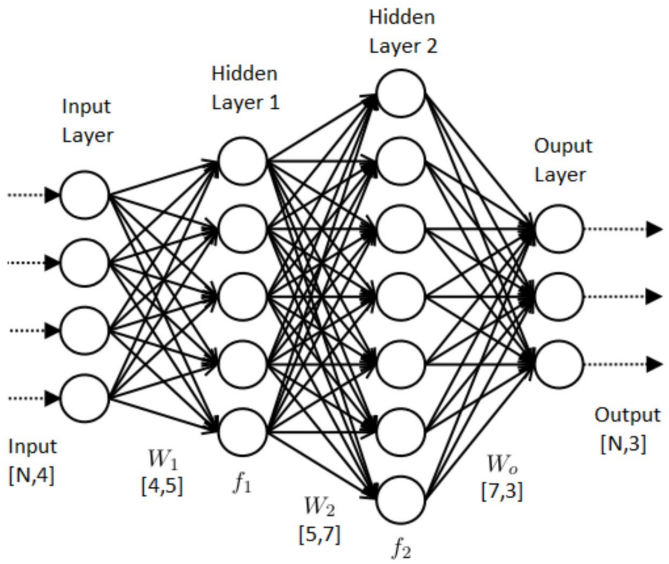

Figure 4.5: Artificial Neural Network architecture.In this case, the input layer is composed by 4 inputs that are multiplied by a weight matrix (W) and passed to the next layer. This happens in each layer of the neural network, and the weights are adjusted during the training phase. This representation contains two hidden layers (with 5 nodes in first layer and 7 nodes in second layer). Finally, the last layer, called output layer with 3 nodes. Image from [58].

4.4.2 Convolution Neural Networks

The Convolution Neural Networks (CNN) uses multiple building blocks (see Figure 4.6). The first layers are the convolution layers and max-pooling layers. The convolution layer is a fundamental layer in the CNN architecture, which performs the feature extraction. Generally consists of a combination of a convolution operation and an activation function. For convolution operations is used a kernel size, and it is applied across to the input. The product of each element between the kernel and the input is calculated at each position and then is summed to obtain the output value in the corresponding position forming the feature map [46, 59]. After a convolution layer it is used an activation function.

The pooling layer performs a downsampling, reducing the input size and retaining the most important information. The max-pooling is the most popular form and extract patches from the input features maps, calculate the maximum value in each patch and discards all the other values [61]. This layer introduces a translational invariance to small shifts and distortions.
The flatten layer or dense layer collapses the spatial dimensions of the multidimensional tensor of the previous layer into a one-dimensional tensor, which is given as input to the fully connected layer [62].
In the fully connected layer each neuron of the previous layer is connected to all neurons in the current layer. This layer combines the features in order to model the outputs [62].
Finally and according to the architecture presented in Figure 4.6 the softmax layer is often used in classification problems and turns a vector of numbers into a vector of probabilities[59]. Normally

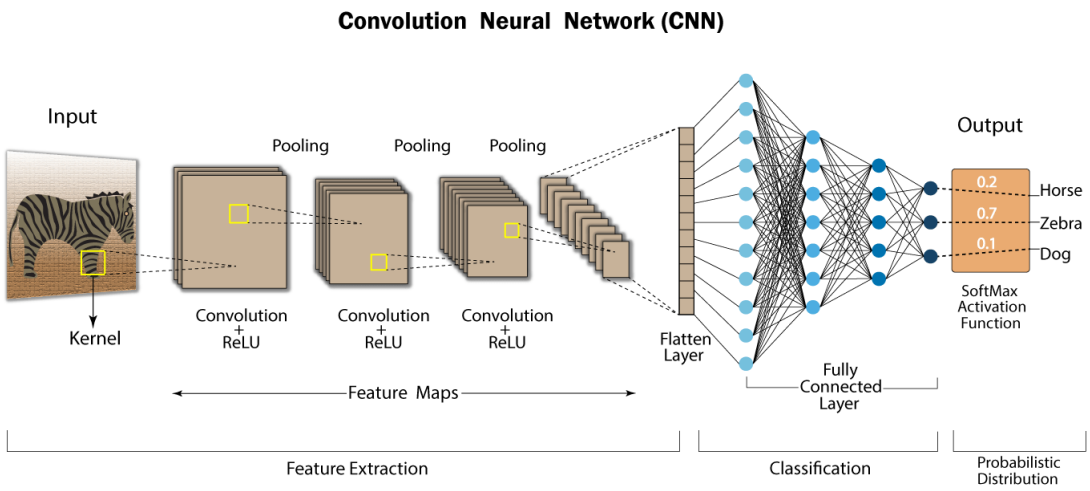

Figure 4.6: A typical CNN architecture, from [60].

in this layer is used a different activation function comparing to the previous layers, which needs to be selected according to the task.

4.4.2.1 Architectures

As the image progresses through the network it gets smaller and smaller but also more deeper and deeper, with more feature maps, depending on the filters used. Typical CNN architectures includes: AlexNet, VGG16, ResNet, GoogleNet and others.

AlexNet: The AlexNet was introduced in 2012 by Alex Krizhevesky and others [63], which won the most difficult ImageNet challenge for visual object recognition called ImageNet Large Scale Visual Recognition Challenge (ILSVRC). It is a deeper and wider CNN model, composed by eight layers, of which five are convolutional layers and the other three are fully connected layers. It uses ReLU activation functions for the hidden layers and softmax for the output layer.

VGG16: Proposed by Simonyan *et al.* [64], members of Visual Geometry Group (VGG) Lab of Oxford University in 2014. It was one of the best performing model in ILSVRC 2014. Creators of the model evaluated the networks and increased the depth using an architecture with small convolution filters (3x3). It is an object recognition and classification algorithm, composed by 16 layers that have weights, reason why this network is called VGG16.
Instead of having a large number of hyper-parameters, the team focus on having convolution layers with 3x3 filter size with stride 1, with same padding and max-pool layer of 2x2 filter with stride 2.

ResNet:The ResNet is an extremely deep network with 152 layers, which won the ILSVRC 2015 challenge. It was developed by Kaiming He *et al.* [65] in order to solve the problem of vanishing gradient. This architecture introduced a concept called residual blocks. In this network, it is used

a technique called skip connections, which connects activations of a layer to further layers by skipping some layers in between, which forms a residual blocks. In this dissertation, was used a ResNet-50, which used 50 residual networks.

GoogleNet: The GoogleNet was developed by Szegedy *et al.* [66] from Google Research and won the ILSVRC 2014 challenge, by building a much deeper network than previous CNNs. This architecture is composed by smaller networks called inception modules. These modules use 1x1 convolutions to perform dimensionality reduction. The second set of layers use different kernel sizes (3x3 and 5x5) to capture patterns in different scales. This architecture use nine inception modules with three layers each.

4.4.3 Overfitting and Techniques to Avoid It

The overfitting refers to a situation where model learns the training data too well, which have negatively impact in the performance of the model on new data. This is one of the main challenges in Machine Learning (ML) and Deep Learning (DL) because an overfit model is not generalizable.
In the Deep Learning is important to evaluate the loss and the accuracy on training and validation set to recognize the overfitting. There are several techniques to minimize the overfitting [47, 62]. One of the best solution is obtaining more training data to generalizes better. However, this is difficult because can be very expensive. Other solutions includes the regularization with dropout, data augmentation and Batch normalization and also reducing the model complexity.
The dropout refers to a technique that randomly omits units during the training phase according to a dropout rate. The Batch normalization was first introduced by Sergey Ioffe *et al.*[67], in the 2015 paper, and is an additional layer which normalizes the output of the previous layer, that is, normalizes the activations of a given input before passing to the next layer in the network. Data augmentation refers to the process of modifying the training data by applying random transformations to them, for example, rotations, translations, flipping, resizing operations and others. An example of a set of augmentations is shown in Figure 4.7. Because of that modifications the model is constantly seeing different examples in training iterations.

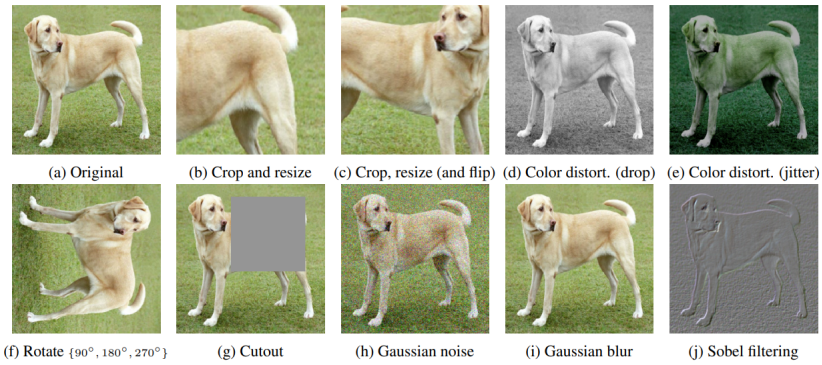

Figure 4.7: Data augmentation technique. This technique can be used to artificially expand the size of a training dataset.Image from [68].

4.5 Evaluation Metrics

This section describes the several evaluation metrics used to evaluate the quality of the obtained models. The algorithm performance is generically evaluated by comparing a score, indicating how a sample belongs to a certain class with a reference standard (ground-truth). From the confusion matrix it is possible to extract four outcomes, respectively:

- **True Positive (TP):** when the predicted value is positive and the actual value is positive.
- **True Negative (TN):** when the predicted value is negative and the actual value is negative.
- **False Positive (FP):** is considered if the predicted value is positive and the actual value is negative.
- **False Negative (FN):** is considered if the predicted value is negative and the actual value is positive.

With these results it is possible to perform several classification metrics, presented below.

- **Accuracy:** is represented as a ratio between the number of all correct predictions and the total number of examples. With imbalanced data should be used the balanced accuracy.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{4.3}$$

- **Recall or True Positive Rate (TPR):** ratio between the number of true positives and the number of positive samples.This performance metric express how well the model deals with the positives. Recall it is also called as sensitivity.

$$Recall = \frac{TP}{TP+FN} \tag{4.4}$$

- **Specificity or True Negative Rate (TNR):** ratio between the correct negative predictions over all negative samples. The perception behind specificity is how well the model deals with the negatives.

$$Specificity = \frac{TN}{TN+FP} \tag{4.5}$$

- **Precision:** refers to the ratio between the number of positive predictions that were correct (TP) and the total amount of positive predictions.

$$Precision = \frac{TP}{TP+FP} \tag{4.6}$$

- **F1-score:** is the harmonic mean of Precision and Recall.

$$F1Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{4.7}$$

- **Receiver Operating Characteristic Curve (ROC):** is a graphical plot that illustrates the performance of the model at all classification thresholds and is defined by two parameters, respectively, the Recall (TPR) and Specificity (FPR).
- **Area Under the Curve (AUC):** is the measurement of the area underneath the entire ROC curve. This performance metric allows us to compare directly the performance of several classifiers [69], which the model with higher AUC is a better model.

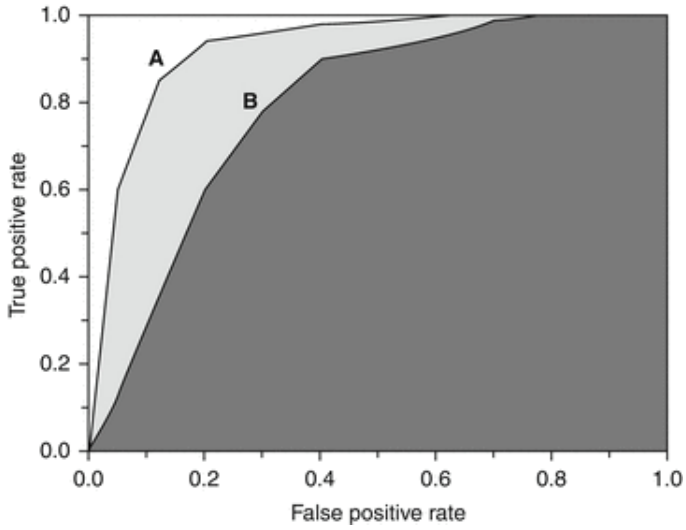

Figure 4.8: Area under the ROC Curve. The figure shows the ROC curves for two classifiers (classifier A and classifier B). Note that the classifier A has a larger AUC value than classifier B.Image from [70].

Chapter 5

Towards an Automatic Pancreatic Ductal Adenocarcinoma Detection System in Computed Tomography

This chapter is dedicated to the proposed deep learning-based pipeline with the objective to detect the Pancreatic Ductal Adenocarcinoma in abdominal contrast-enhanced CT (CE-CT). First, is presented the method overview, respectively, the region of interest detection and Pancreatic Ductal Adenocarcinoma classification. In each of the steps is described the model CNN architecture which approach is based on, a detailed description of the training and the parameters and the analysis of the results.

5.1 Method Overview

The proposed method to detect the Pancreatic Ductal Adenocarcinoma in abdominal CT images is detailed in this chapter. Our approach can be divided in:

1. ROI detection, to analyze further the regions of diagnostic interest for cancer detection.

2. Pancreatic ductal adenocarcinoma classification by distinguish the pancreatic tissue as healthy or unhealthy.

Our method is developed and evaluated using the NIH - Pancreas CT [37] and TCIA CPTAC-PDA [38] datasets, both available publicly. Please refer to section 3.3 for a comprehensive description of the datasets.
The TCIA CPTAC-PDA contains medical images from many imaging modalities, so it was filtered according to the inclusion criteria, presented below, resulting in 59 patients.

1. Dataset has minimal noise, contextually appropriate and/or the exam of the patient corresponds to an abdominal CT.

2. Dataset of abdominal CT scans, in different phases of contrast including the portal-venous phase.

3. Complete inclusion of the pancreas in abdominal CT scans.

The figure 5.1 shows a schematic representation of the process of the filtration of this dataset.

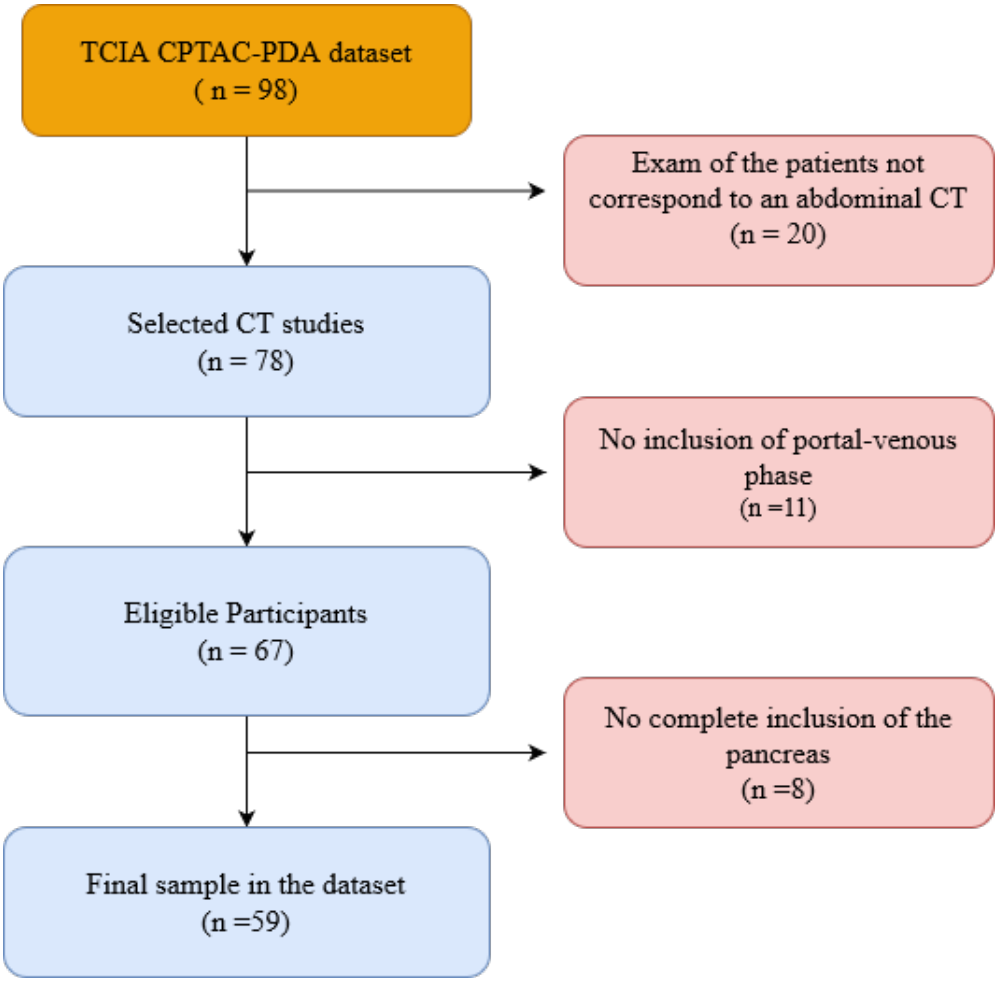

Figure 5.1: Study selection flowchart of the TCIA CPTAC-PDA dataset. From 98 patients in the initial database, the final number included in the study is 59 patients.

Both the datasets are mixed randomly. Each of the steps, will be detailed in the remaining of this chapter.

5.2 First Stage: Region of Interest Detection

As stated in Chapter 2, in the medical imaging the pancreas vary in shape and size between the patients and have poorly-defined margins and low contrast comparing with adjacent anatomical structures. Medical images are larger and complex than non-medical images used for image processing. Filtering out the no relevant information is an important step to analyzing the important

regions of the images, and it is beneficial for further processing and analysis and reducing the time spent in processing the images.

Previously, in the section 4.4.2 was described some CNN architectures. For this step, was used a pre-trained Faster R-CNN ResNet-50 FPN model. The ROI delimitation task would be composed by two processes: the data preprocessing and the ROI pancreas detection.

5.2.1 Data Preprocessing

The data preprocessing includes the annotation procedure, the split of the dataset and the normalization of the data.

First, was performed the annotation procedure of the regions of the interest.The annotations were performed in portal-venous phase CT scans for better pancreas visibility. To ensure that the bounding box cover the entire pancreas, the delimitation of the ROI was conducted in three anatomical planes (axial plane, coronal plane and sagittal plane). For this task was used a graphical user interface based on Streamlit [1] to manually draw the bounding boxes that compose the reference standard (ground truth). It is important to note that the pancreas is a small organ in the abdomen, so the number of images containing the organ only represents about 30% of the total of the images in each abdomen CT. The annotations were performed by a radiographer, respectively, the author of this dissertation, who have six years of experience in radiology from University Hospital Center of São João. All bounding boxes are manually drawn, which requiring highly labor-intensive work.

The dataset is randomly split into 70% for training, 15% for validation and 15% to compose the test set. The composition of each set are presented in Tables 5.1 and 5.2. The split was made by exam and not by slices, so that there are no slices of the same exam in the several sets. This ensures that the validation and testing set are not performed on images similar to those used in the training set.

Table 5.1: Number of the patients after the split of the dataset.

	Healthy patients	Unhealthy patients	Total patients
Training set	57	41	98
Validation set	12	9	21
Testing set	13	9	22

Table 5.2: Number of the images after the split of the mixed dataset.

	Healthy images	Unhealthy images	Total images
Training set	13561	10602	24163
Validation set	2644	2245	4889
Testing set	3123	3176	6299

[1] More information at: https://streamlit.io/

The CT images were processed from DICOM format to a 3D volume for each exam. Each exam was corrected the orientation of the exam, and it is applied also an abdomen window of [40,400] HU, to enhance image details (see figure 5.2). Further, is applied a normalization, which scales values of the HU units in [0,1] range. To use the Faster R-CNN ResNet-50 FPN it is required that the input have three channels, that is, the input images are converted from grayscale to the RGB format. Then, is applied the recommended normalization for that model, Normalize (mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.225]).

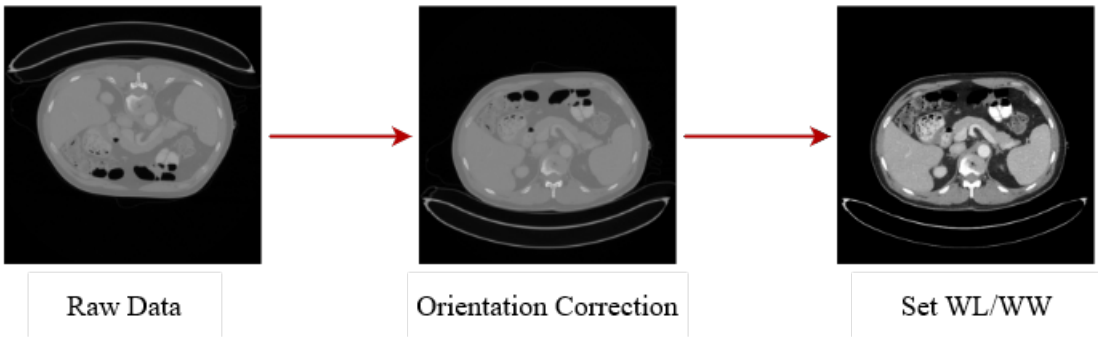

Figure 5.2: The process flow of the raw data. Raw volume data were applied an orientation correction and then was applied an abdomen window. Window width (WW) and Window level (WL).

5.2.2 ROI Pancreas Detection

A variety of deep learning models have been proposed for object detection. These models can be classified in two main categories:

- one-stage approaches, including YOLO and SSD (Single Shot MultiBox Detector) to simultaneously detect the location and classify the target object.

- two-stage detector, for example Faster R-CNN and Feature Pyramid Network (FPN). That models first find the region proposal, classify and regress the location to place an anchoring box to frame the target. The Faster R-CNN detects small objects because a variety of sizes of anchors are used in a single grid.

A perception system usually performs multiple tasks such as object detection, semantic segmentation among others. However, having separate models for each task involves a very high computational cost and memory, which in most of the cases is impractical. A solution is the multi-task learning, in which multiple learning tasks are solved at the same time. This result in an improved learning efficiency and prediction accuracy for the task specific models. A top performing framework for object detection and classification and instance segmentation is the Faster R-CNN with Feature Pyramid Network [71] to learn multi-scale representations.

5.2.2.1 Proposed model

For the purpose, was used a Faster R-CNN based on the strategy of A. Nugaliyadde *et al.*[72] with ResNet and FPN as backbone, which is described below. The Faster R-CNN model consists of two parts: a Region Proposal Network (RPN) and a ROI pooling.These two parts shared a hidden layer, which is a deep convolutional neural network. The proposed approach, use a shared feature representation based on ResNet-50 and FPN. The figure 5.3 show the architecture of the Faster R-CNN ResNet-50 FPN [73] used in this task.

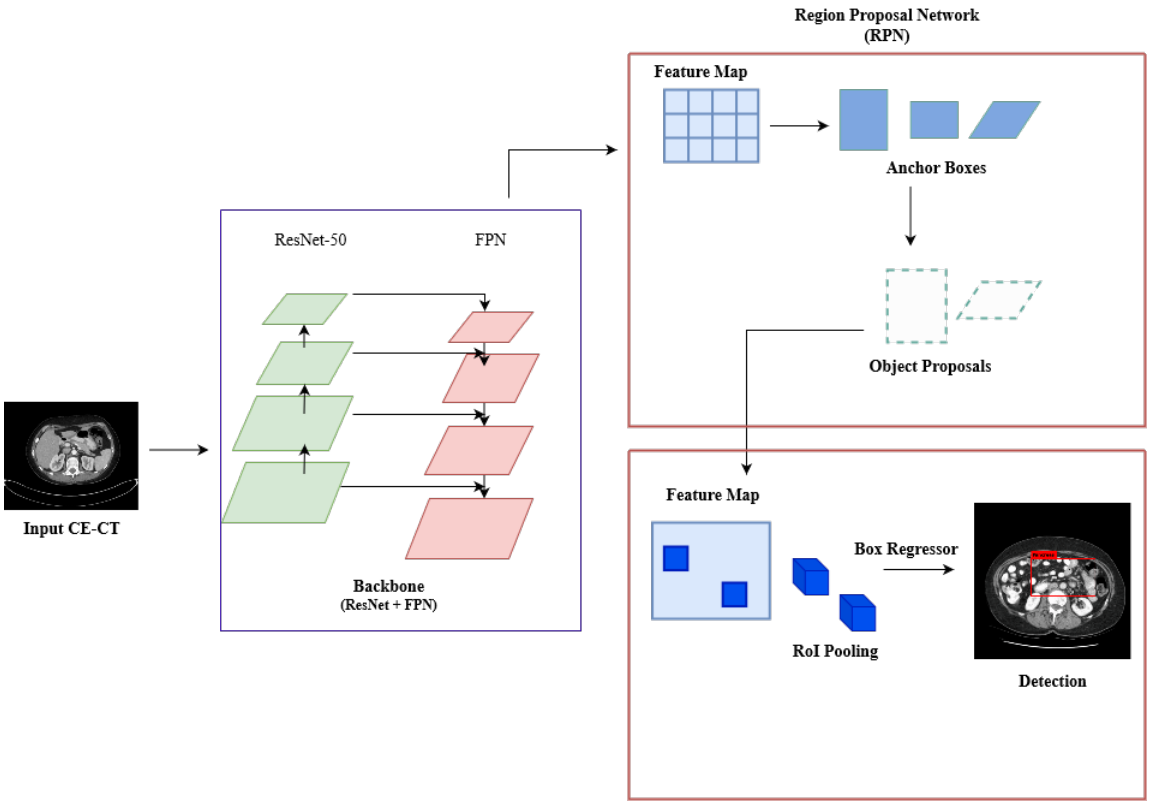

Figure 5.3: The architecture of Faster R-CNN containing a ResNet-50 and FPN as backbone.

The RPN input is an image and the output is a set of rectangular region proposals that represent an area that contains an object. After inputting the image, the last layer's feature map is obtained using the deep convolutional network and then a sliding window sweeps over the entire feature map. Each point on the feature map represents an anchor. In the sliding window there are k reference frames. The reference frames are transformed into actual region proposals depending on the parameters output using the sliding window. Then, the region proposals are extracted using a ROI pooling to give regions of interest of the same size and posteriorly imported into a final layer for classification and detection.

At the last layer of ResNet-50 is built the 4-level Feature Pyramid in a top-down manner by upsampling feature maps from higher pyramid levels and merging them via lateral connections with corresponding feature representations from the ResNet. In the Feature Pyramid Network, low-level features have less semantic information, but the location information is accurate. In

the high-level features semantic information is abundant, but the information on locations can be eliminated, so some algorithms use multi-scale features for prediction. Each level of the pyramid consist of more complex richer features.

5.2.2.2 Training Details

The project was developed on Python 3.7.13 and implemented in Pytorch version 1.12.1 [74]. The model was trained using Google Colab, which provides access to powerful GPUs and requires no configuration. The ResNet-FPN based joint architecture, was trained using the Stochastic Gradient Descent (SGD) optimizer. The chosen learning rate for the Faster R-CNN based architecture applied to the dataset was 0.005. A learning rate scheduler[2] was also employed, which consisting in a reduction of learning rate by a factor of 0.1 every three epochs. The number of epochs was chosen in an empirical way, that is, according to how many epochs were needed to reach a minimum in the validation loss, which was considered as reached after four epochs of no improvements. In the dataset was used a batch size of 5, due to the hardware characteristics. The batch size defines the number of samples that will be propagated through the network. During training, the Faster R-CNN ResNet-50 FPN expects as input the ground truth bounding boxes and the respective labels (class label for each ground truth bounding box). The weights used are the result of training performed for the dataset Common Objects in Context (COCO) 2017 [75]. The model and the weights used belong to the TorchVision library[3].
The model weights with the lower loss value on the validation set was chosen during training phase and saved. The Faster R-CNN ResNet-50 loss function is composed by four losses: box loss, objectness loss, box regression loss and classification loss. The first two losses belong to the Region Proposal Network and the last two losses belong to the ROI pooling layer. The box loss and the box regression loss represents how well the algorithm can locate the center of an object and how well the predicted bounding box covers an object. It used as metric the Smooth L1 loss.
The objectness loss is essentially a measure of the probability that an object exists in a proposed region of interest. If the objectivity is high, this means that the image window is likely to contain an object. This part use the Binary Cross Entropy as metric.
The classification loss represents how well the algorithm can predict the correct class of a given object, which use the Cross Entropy as metric.
The Smooth L1[4] also known as Huber loss, it is given by:

$$\ell(x,y) = L = \{l_1,...,l_N\}^T \tag{5.1}$$

[2]More information at:https://pytorch.org/docs/stable/generated/torch.optim.lr_scheduler.StepLR.html
[3]More information at:https://pytorch.org/vision/stable/index.html
[4]More information at:https://pytorch.org/docs/stable/generated/torch.nn.SmoothL1Loss.html

$$l_n = \begin{cases} 0.5(x_n - y_n)^2/beta, & if|x_n - y_n| \leq beta \\ |x_n - y_n| - 0.5 \times beta, & otherwise \end{cases} \tag{5.2}$$

where N is the size of the batch. This loss creates a criterion that uses a squared term if the absolute error falls below 1 and an absolute term otherwise. It is less sensitive to outliers than the mean square error loss and in some cases prevents exploding gradients.
The Cross Entropy Loss[5] is a criterion that measures the cross entropy between the predicted and the actual value and is computed as:

$$\ell(x,y) = L = \{l_1, \ldots, l_N\}^\top, \quad l_n = -w_{y_n} \log \frac{\exp(x_{n,y_n})}{\sum_{c=1}^{C} \exp(x_{n,c})} \cdot 1\{y_n \neq ignore_index\} \tag{5.3}$$

where x is the input, y is the target, w is the weight, C is the number of classes, and N spans the mini-batch dimension.
The Binary Cross Entropy [6] creates a criterion that measures the binary cross entropy between the target and the input probabilities and was defined as follows:

$$\ell(x,y) = L = \{l_1, \ldots, l_N\}^\top, \quad l_n = -w_n [y_n \cdot \log x_n + (1 - y_n) \cdot \log(1 - x_n)] \tag{5.4}$$

where N is the batch size, x is the input, y is the target, w is the weight.
The loss functions are shown in figures 5.4 and 5.5. In the figure 5.6 is presented a plot of the loss of the model, which is the result of the sum of the mentioned losses.

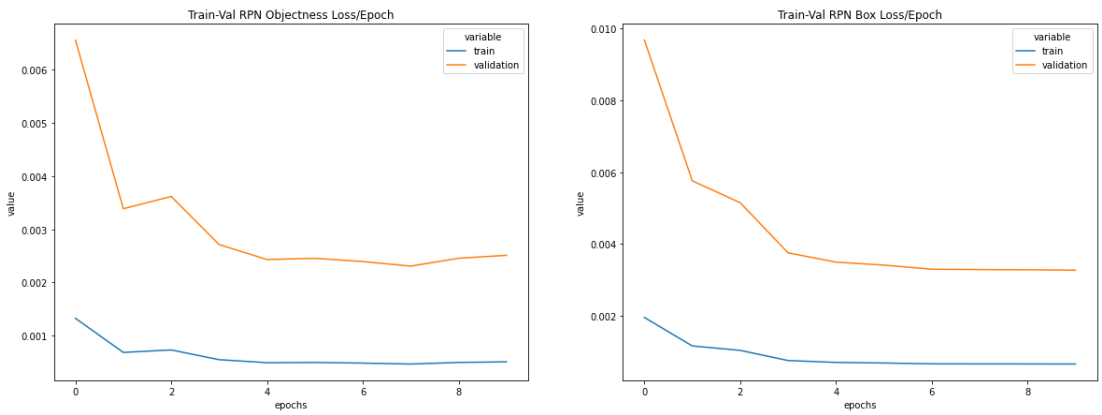

Figure 5.4: Model losses for Region Proposal Network layer of the Faster R-CNN over the training epochs for the training and validation set. The **left** image shows the plot of objecteness loss, which used a binary cross entropy loss. The **right** image shows the plot of box loss, which use the Smooth L1 loss.

[5]More information at:https://pytorch.org/docs/stable/generated/torch.nn.CrossEntropyLoss.html
[6]More information at:https://pytorch.org/docs/stable/generated/torch.nn.BCELoss.html

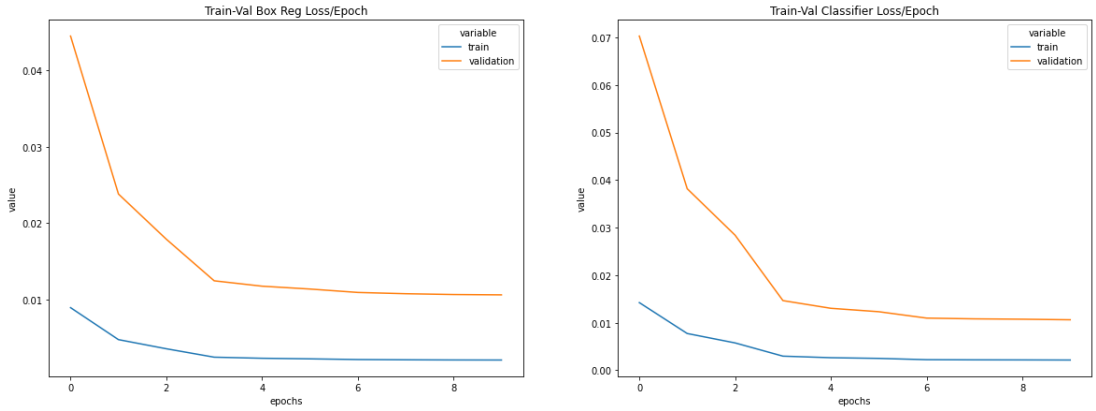

Figure 5.5: Model losses for ROI pooling layer of the Faster R-CNN over the training epochs for the training and validation set. The **left** image shows the plot of box regression loss, which used the Smooth L1 loss. The **right** image shows the plot of classifier loss, which use the cross entropy loss.

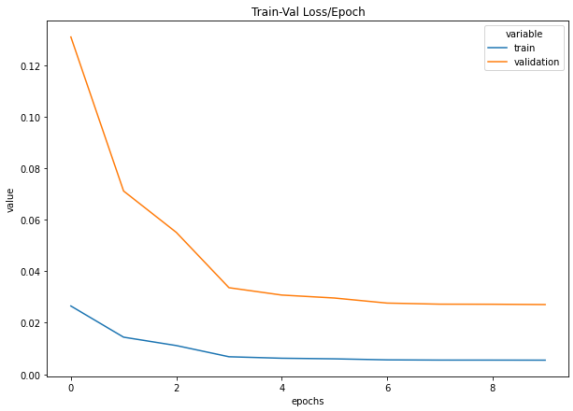

Figure 5.6: Total loss plot of the Faster R-CNN over the training epochs for the training and validation set. The total loss is the result of the sum of the previous loss functions.

It was then used to perform inference and evaluation on the test set, which model returns the predicted boxes, the predicted labels and scores of each detection.

5.2.2.3 Results

Regarding the performance metric for this task it was used some similarity metrics, respectively the Jaccard Index (JI), Sorensen-Dice Index (DSC) and Overlap coefficient (OV).
The Jaccard index (IoU) is similar to the Sorensen-Dice coefficient and is defined as:

$$JI = \frac{|V_{gt} \cap V_{pred}|}{|V_{gt}| + |V_{pred}| - |V_{gt} \cap V_{pred}|} \tag{5.5}$$

The Sorensen-Dice coefficient measures the volumetric overlap ratio between the ground truth and network predictions. It is defined as follows:

$$DSC = \frac{2|V_{gt} \cap V_{pred}|}{|V_{gt}| + |V_{pred}|} \tag{5.6}$$

The Overlap coefficient is defined as the size of the union of set A and set B over the size of the smaller set between A and B.

$$OV = \frac{2|V_{gt} \cap V_{pred}|}{min(|V_{gt}|, |V_{pred}|)} \tag{5.7}$$

where V_{gt} and V_{pred} represents the bounding boxes sets of manual annotations and network predictions, respectively. All the similarity measurements ranges from 0 to 1. The closer the measurement is to 1 the better the prediction effect. The results on the test set with the above metrics are shown in Table 5.3.

Table 5.3: Average results on the test set for Faster R-CNN ResNet-50 FPN model.

Method	Jaccard Index	DSC Coefficient	Overlap Coefficient
Faster R-CNN Resnet-50 FPN	0.807	0.920	0.929

The sensitivity, specificity, accuracy and precision of the model were analyzed, and the values are shown in Table 5.4.

Table 5.4: Detection performance of our model.

Method	Sensitivity	Specificity	Accuracy	Precision
Our model	**0.890**	**0.941**	**0.920**	**0.900**

After analysis of the statistical results, the resulted images are available for comparison. Figures 5.7 and 5.8 shows the visualized results in the experiment based on the Jaccard index or Intersection-over-Union (IoU) index. The results were obtained using only the bounding boxes with a confidence interval $\geq 90\%$.

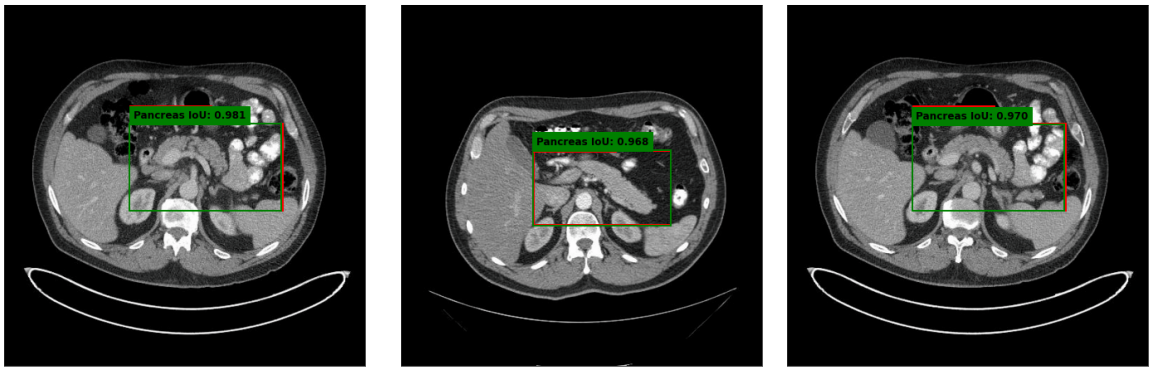

Figure 5.7: The visualized results in the experiment, which corresponds to the correct results, that is, with the highest IoU. The green bounding boxes corresponds to the predictions and the red bounding boxes corresponds to the ground-truth.

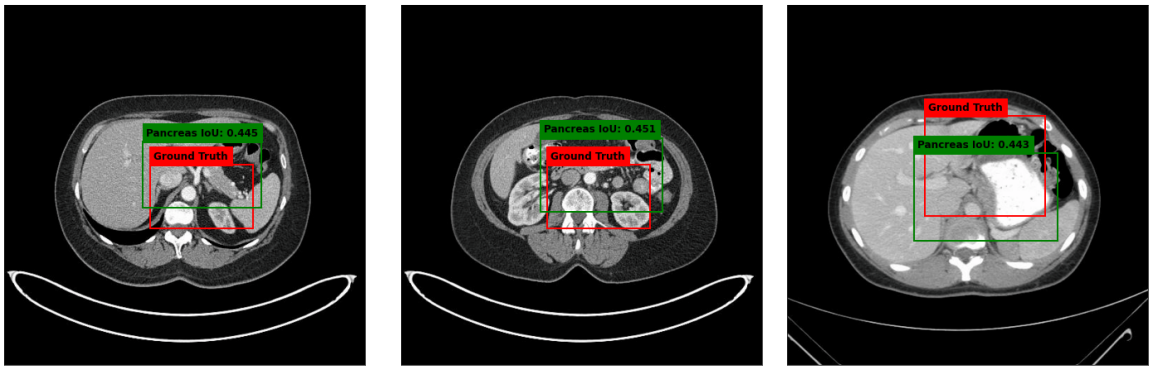

Figure 5.8: The visualized results in the experiment, which corresponds to the incorrect results, that is, with the lowest IoU (IoU < 0.5).The green bounding boxes corresponds to the predictions and the red bounding boxes corresponds to the ground-truth.

Analyzing and observing the detection and classification results of the AI system more straightforwardly we classified all the labeled regions of the test set into TP, FP, TN and FN outcomes, which is detailed in Figure 5.9 in the confusion matrix. The classification confusion matrix shows a negative predictive value (TN/ TN+ FN) of 0,94.

Figure 5.9: The confusion matrix of the system.

5.2.2.4 Discussion

Our datasets are manually revised. In the TCIA CPTAC-PDA dataset were removed some patients, as documented in the figure 5.1. In this task was used a model based in the transfer learning technique [76, 77], which is an approach that applies knowledge learned from a different domain, for example the ImageNet, to build a pre-trained model and then apply the trained model to target images, for example in abdominal CT images. In medical imaging the target domain is often small, so it is impractical to properly fine-tune the deeper layer. It was used a backbone (ResNet-50 + FPN) as feature extractor to capture high-level and informative details from images. These features, can be further used to train a classifier for the target domain, allowing the building of a high performance transfer learning model using a small dataset. The proposed architecture reached an accuracy of 92%. The figures 5.5 and 5.6 shows the results obtained in this experiment and by visual inspection and quantitatively by Jaccard index, can conclude that the model fit the manual annotation very well. However, it is important the identification of bad results, that is, results with the lowest Jaccard index, to understand why that occurred and debug the model. Taking this into account, it is possible to explain the bad results of the Jaccard index: the ground truth was generated as one bounding box containing the pancreas for each CT volume, which, possibly decreases the performance of the Jaccard index. A solution can be performing the bounding boxes adjusted to visible area of the pancreas in each slice.

5.3 Second Stage: Pancreatic Ductal Adenocarcinoma Classification

5.3.1 Introduction

This stage intend to study a method of identification of the Pancreatic Ductal Adenocarcinoma, which is the major type of the pancreatic cancer accounting for about 85% of the cases [31] and can be detected it by checking the abdominal CT scans. The goal of this part of the work is discriminating abnormal pancreas from normal ones. This is a classification task, and it is used the pre-trained AlexNet model with the weights used in ImageNet-1K dataset, from the TorchVision library.

5.3.2 Data Preprocessing

The split process was made the same as in the previous section. Firstly, the dataset was randomly split into 70% for training, 15% for validation and 15% to compose the test set (see Table 5.5). The Table 5.6 represents the number of images of each dataset, after the split of the CT volumes into images. Since the deep neural networks are easily affected by the large and irrelevant background information due to the small size of the pancreas in the whole-slice of a CT scan, the data was cropped by the pancreas bounding boxes, using the ROI of ground truth. This process is illustrated in the figure 5.10.

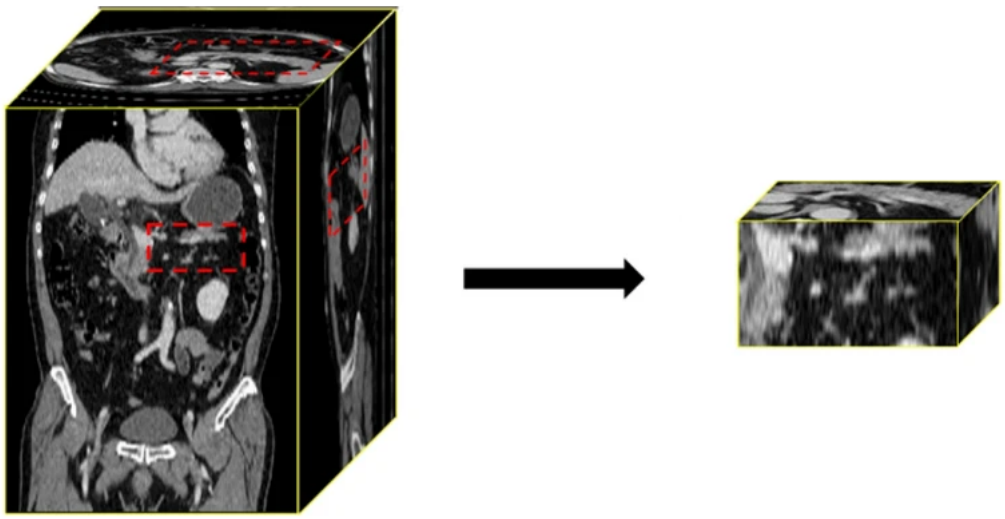

Figure 5.10: The ground truth were cropped for training data generation.

The data preprocessing process was the same as in the previous section. It is recommended that the images have the same size for which the model was trained, so the input have a size

of (224,448,224), respectively, the depth, width and height. The CT volume was divided into slices, creating 448 images with size of (224,224). Similar to the model used in the first stage, this model only accepts images with three channels (RGB format), as input.

Table 5.5: Number of the patients after the split of the dataset.

	Healthy patients	Unhealthy patients	Total patients
Training set	57	41	98
Validation set	12	9	21
Testing set	13	9	22

	Healthy images	Unhealthy images	Total images
Training set	25536	18368	43904
Validation set	5376	4032	9408
Testing set	5824	4032	9856

dataset.

5.3.3 Pancreatic Ductal Adenocarcinoma Classification

A CT-scanned images cropped with the region of interest ($\mathbf{X}$) is a L x W x H matrix, where L are the length, W are the width, H are the height of a parallelepiped, respectively. Our goal is to design a discriminative function $p(\mathbf{X}) \in \{0,1\}$, with 1 indicating that this a person suffers of PDAC and 0 otherwise.

5.3.3.1 Network Arquiteture

The architecture of the AlexNet [17] is described in Chapter 4 and represented in figure 5.11. In brief, the most important feature is not the number of layers but the below parameters:

- ReLU non-linearity: ReLU means rectified linear unit and is considered to be the most applied activation function in deep learning. It can be understood, easily as $f(x) = max(0,x)$, which implies that if there is any negative value it will be made 0 but for any positive value the function f(x) will return the positive value.
- Multiple GPUs: The AlexNet allow the usage of multiple GPUs at the same time, which reduce the training time but also help in training larger models.

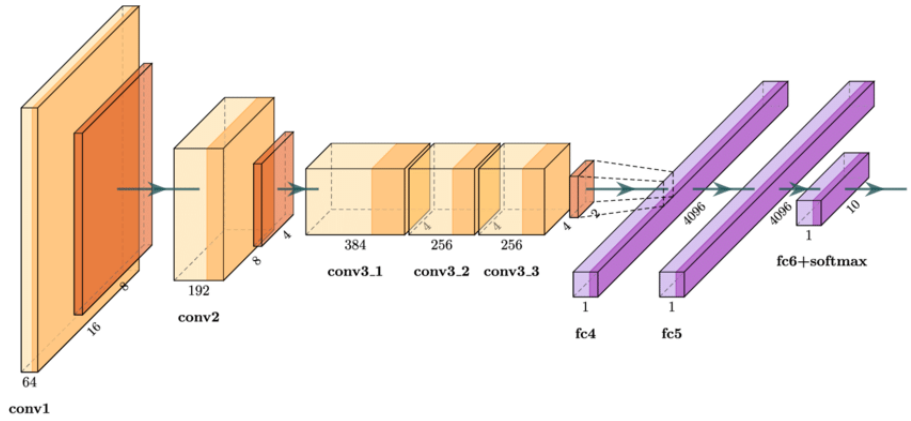

Figure 5.11: The AlexNet architecture.Image from [1].

5.3.3.2 Details of Learning

In the pre-trained AlexNet model from the ImageNet Challenge [78] was used the Stochastic Gradient Descent (SGD) as optimizer. The starting learning rate for the AlexNet model was 0.0001, as the default value. A learning rate scheduler was employed that adjust the learning rate between epochs as the training progresses. In this case, consists in a reduction of learning rate by a factor of 0.1 every seven iterations. The model was trained during 40 epochs and the best epoch is chosen when is reached the minimum validation loss. It was selected the epoch number 8 with a validation loss of 0.127. Regarding the evolution of the classifier's accuracies and loss dataset, the classifier stabilizes faster, as shown in figure 5.12.

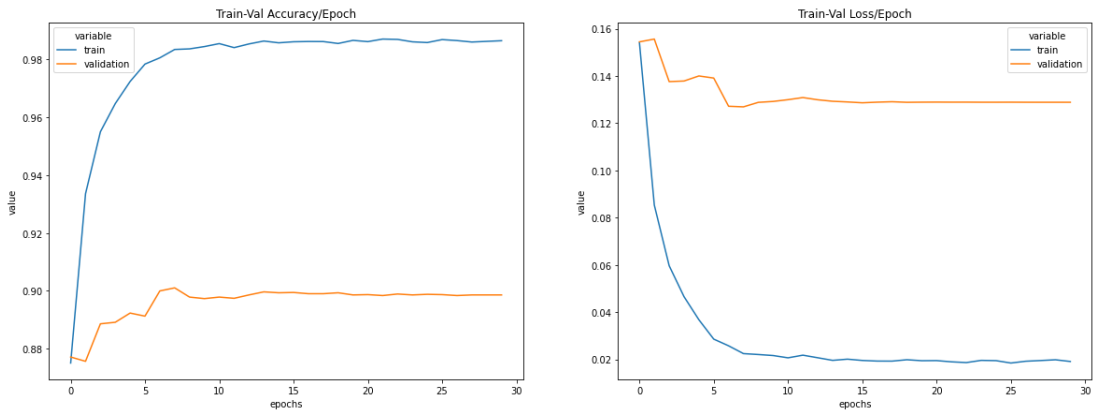

Figure 5.12: Model loss and accuracy plot for training and validation set.

For this approach a different strategy for the loss function was designed. It was applied the Multi-class margin loss[7], which is represented below:

$$loss(x,y) = \frac{\sum_i^C \max(0, margin - x[y] + x[i])^p}{N} \tag{5.8}$$

[7]More information at: https://pytorch.org/docs/stable/generated/torch.nn.MultiMarginLoss.html

where x is the input, y is the output, C is the number of classes and N is the batch size.
In this model was used a batch size of 16 and since it is an unbalanced dataset was applied the Weighted Random Sampler[8] , which apply the oversampling technique. This technique is used to adjust the class distribution of a dataset, by randomly increasing the number of the minority observations until is reached a balanced dataset. Then, is performed the evaluation on the test set which returns a probability for every possible class.

5.3.3.3 Results

Regarding the training phase for the best epoch, saved in validation set, were evaluated the performance of the model based on the metrics described in section 4.5 and presented in Table 5.7. The PDAC classification model was tested on 9856 CT images, which corresponds to a total of 22 patients.

Table 5.9: Performance metrics of the AlexNet model.

System	Sensitivity	Specificity	Accuracy	Precision
Han Ma *et al.* [34]	0.923	0.979	0.952	-
Kao-Lang Liu *et al.* [35]	0.790	0.976	0.832	-
Our method	**0.900**	**0.924**	**0.920**	**0.890**

The overall sensitivity of the proposed model is 90%, showing promising results. The figure 5.13 present the confusion matrix and in the figure 5.14 is represented the ROC curve for the model.

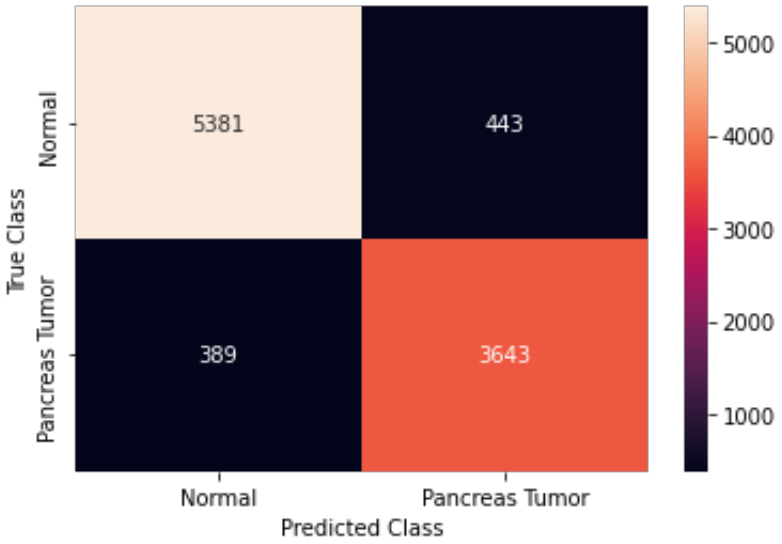

Figure 5.13: Confusion matrix of the system.

[8]More information at: https://pytorch.org/docs/stable/data.html

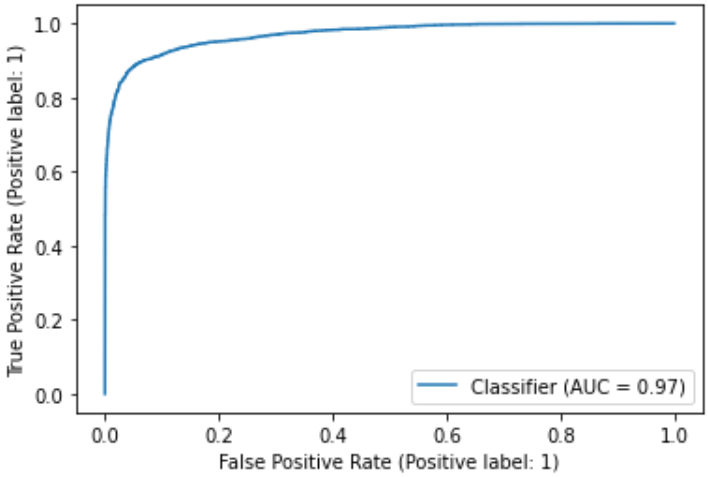

Figure 5.14: Receiver operating characteristic for the automatic system.

The figure 5.15 to figure 5.18 shows examples of correct and incorrect results for both classes, predicted by our method.

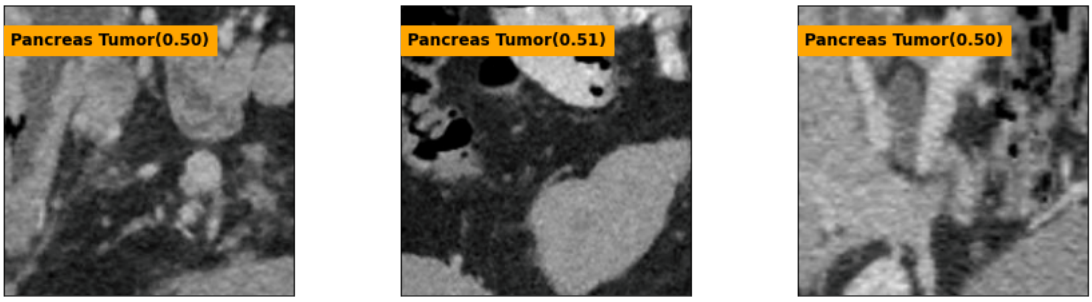

Figure 5.15: The visualized results in the experiment, which corresponds to the incorrect predictions belonging to unhealthy class. The value between parenthesis corresponds to the Jaccard index.

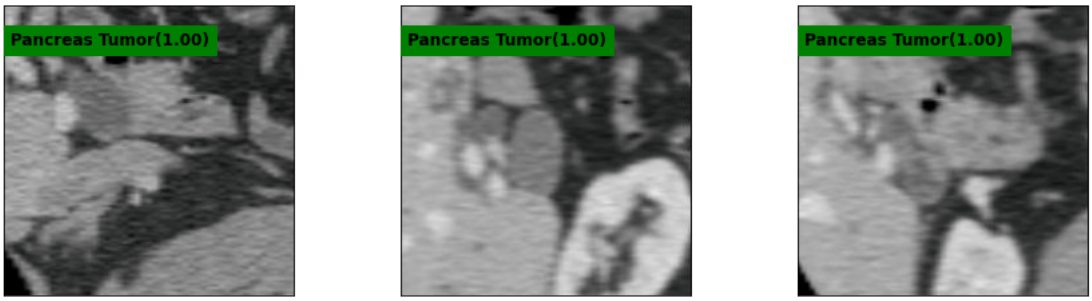

Figure 5.16: The visualized results in the experiment, which corresponds to the correct results belonging to unhealthy class.The value between parenthesis corresponds to the Jaccard index.

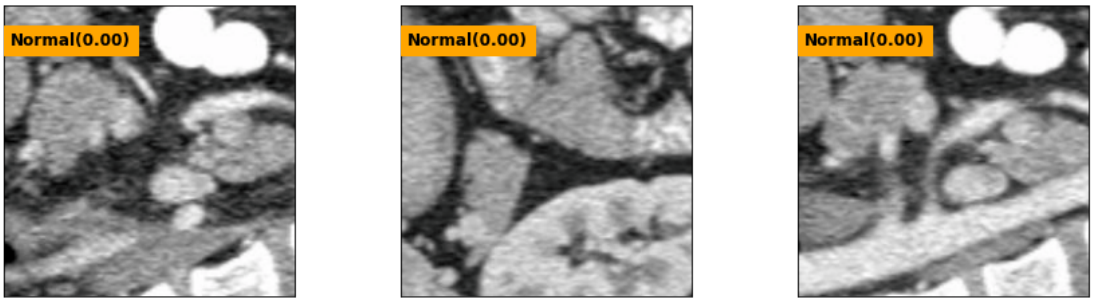

Figure 5.17: The visualized results in the experiment, which corresponds to the incorrect predictions belonging to normal class.The value between parenthesis corresponds to the Jaccard index.

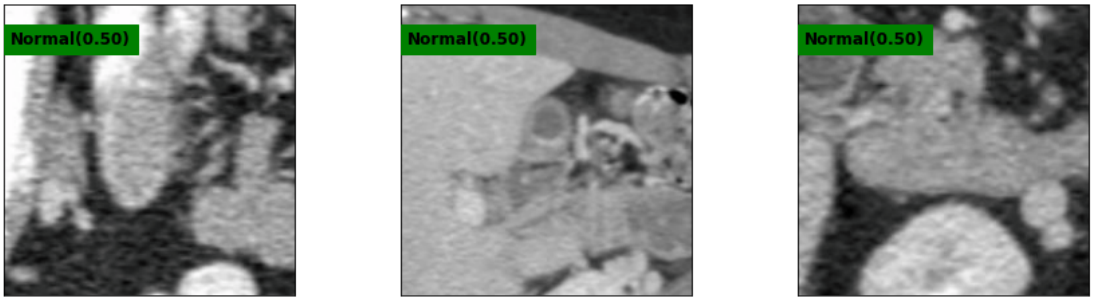

Figure 5.18: The visualized results in the experiment, which corresponds to the correct results belonging to normal class.The value between parenthesis corresponds to the Jaccard index.

Analyzing the results there is scope for improvements if a larger dataset is used and if the dataset has more diverse cases of the pancreatic cancer. However, considering the dataset size and the lack of the diversity cases the module learned well.

5.3.3.4 Discussion

In this step was used CT images in portal-venous phase, since is the phase that discriminate better the pancreas. The bounding box allow cropping the majority of the volumetric space from any input CT scan. However, for the future work it is important to analyze the tridimensional information of the pancreas or make a classification in each slice of the CT, in case of bidimensional model. The PDAC classifier is based on pre-trained AlexNet model, applying the transfer learning technique. Our model was compared with two studies, but not surpass the study of Han Ma *et al.*[34]. Previous studies showed that CT has a sensitivity of 76% to 92% and an accuracy of 85% to 95% for diagnosis of pancreatic cancer according to the ability of the doctors. Our results indicate that our model have the same detection performance [79, 80].

5.4 Conclusion Remarks

To conclude this chapter, the pancreatic cancer is a very fatal disease, and it is difficult to detect. Several studies are going on for early detection of this disease. The above algorithm is a stepping

stone to predict the same, and it can be better if the data is available in quantity and consistently. Both proposed models were functional, and they were not able to improve the state-of-the-art results, for example the ones published by Han Ma *et al.* [34]. However, this work demonstrated the effectiveness of using our proposed network for transfer learning. Model need an even higher variety of examples to properly learn. Also, it is not still clear if that approach will lead to the success, considering these tasks. For all of these reasons, it is still needed further research .

Chapter 6

Conclusions and Future Work

6.1 Overview

Data is the core ingredient in every Deep Learning pipeline. Incredibly difficult to come by labelled datasets and expensive to collect and annotate one. The outstanding performance level of deep learning-based models has actually led to the development of such systems. The second opinion computed-aided systems are useful to increase the overall cancer detection performance of the physicians. In this dissertation, the focus was designing data-efficient and knowledge-aware deep learning techniques towards medical machine intelligence. Consider that, there are several challenges in the field of medical imaging analysis, for example, the unsatisfactory performance in small targets, insufficient training data, high annotation cost, lack of specific domain knowledge and among others. When clinical experts interpret the medical images, they often rely on prior knowledge about anatomy. On the other hand, convolutional methods are limited in incorporating such domain specific knowledge. Taking into account, the main goal of the current work was investigate and present a two-step approaches for the PDAC detection in abdominal CT images. The two-step approach was achieved by using object detection networks for the pancreas localization and PDAC classification. This approach proved to be promising, because the performance metrics are similar to obtained in the state-of-the-art.

6.2 Future Work

The Deep learning has led remarkable achievements in the field of medical image analysis but the gap between research settings and real world clinical settings remains huge. In the future, we aim to investigate the following research directions:

1. Apply deep neural networks aware of knowledge priors.

2. Learning from heterogeneous data, because different equipments, image acquisition protocols and different patients populations contributes to the fact of, generally, the databases are disjoint. For maximizing the data utilization and make learned representations more robust

and universal it is important to adopt the use of dataset fusion techniques, that leverage multiple datasets to train a universal model (for example: multi-task learning and self-supervised learning).

3. Apply multi-phase learning algorithm to detect PDAC, because the dual-phase imaging is clinically used for better diagnosis of the pancreatic disease.

4. Individual annotations of the pancreas, the tumor and adjacent anatomical structures, such as the major arteries and veins.

References

[1] Nicola Strisciuglio, Manuel Lopez Antequera, and Nicolai Petkov. Enhanced robustness of convolutional networks with a push–pull inhibition layer. In *Neural Computing and Applications*, pages 1–15. Springer, 2020.

[2] J. Jameson, Dennis Kasper, Dan Longo, Anthony Fauci, Stephen Hauser, and Joseh Loscalzo. *Harrisson ´s: Principles of Internal Medicine - Volume I*. McGraw-Hill Education, Twentieth edition, 2018.

[3] Shallu Midha, Saurabh Chawla, and Pramod Kumar Garg. Modifiable and non-modifiable risk factors for pancreatic cancer: A review. *Cancer Letters*, pages 269–277, October 2016.

[4] Huiyun Zhu, Tuo Li, Yiqi Du, and Min Li. Pancreatic cancer: challenges and opportunities. *BMC Medicine*, page 214, November 2018.

[5] Georgios Kaissis and Rickmer Braren. Pancreatic cancer detection and characterization—state of the art cross-sectional imaging and imaging data analysis. *Translational Gastroenterology and Hepatology*, May 2019.

[6] Linda C. Chu, Michael G. Goggins, and Elliot K Fishman. Diagnosis and detection of pancreatic cancer. *The Cancer Journal*, November/December 2017.

[7] Ana Caldeira. Cancro do pâncreas: a crueza dos números e como inverter a tendência. *Revista Saúde e Bem-Estar*, pages 16–18, January 2020.

[8] Lulu Zhang, Santosh Sanagapalli, and Alina Stoita. Challenges in diagnosis of pancreatic cancer. *World Journal of Gastroenterology*, pages 2047–2060, May 2018.

[9] Fernando Araújo. Despacho n.º 8254/2017, September 2017. Available in: https://dre.pt/dre/detalhe/despacho/8254-2017-108189401.

[10] Richard L. Drake, A. Wayne Vogl, and Adam W. M. Mitchell. *Gray ´s Anatomy for Students*. Elsevier Inc., Fourth edition, 2020.

[11] Javier Casillas, Joe U. Levi, Alexander O. Quiroz, Roberto Ruiz-Cordero, Monica T. Garcia-Buitrago, and Danny Sleeman. *Multidisciplinary Teaching Atlas of the Pancreas:Radiological, Surgical, and Pathological Correlations*. Springer, 2016.

[12] Hans G. Beger, Andrew L. Warshaw, Ralph H. Hruban, Markus W. Büchler, Markus M. Lerch, John P. Neoptolemos, Tooru Shimosegawa, and David C. Whitcomb. *The Pancreas:An Integrated Textbook of Basic Science, Medicine, and Surgery*. Willey Blackwell, Tird edition, 2018.

[13] Morton A. Meyers, Chusilp Charnsangavej, and Michael Oliphant. *Meyers' Dynamic Radiology of the Abdomen:Normal and Pathologic Anatomy*. Springer, Sixth edition, 2005.

[14] John P. Neoptolemos, Raul Urrutia, James L. Abbruzzese, and Markus W. Büchler. *The Pancreatic Cancer*. Springer, Second edition, 2010.

[15] American Cancer Society. Cancer facts and figures: 2022. *Atlanta: American Cancer Society*, 2022.

[16] International Agency for Research on Cancer. Portugal: Globocan 2018. *The Global Cancer Observatory*, May 2018.

[17] R. Lakshmana Kumar, R. Indrakumari, B. Balamurugan, and Achyut Shankar. *Exploratory Data Analytics for Healthcare*. CRC Press, First edition, 2022.

[18] Eric Peter Tamm, Priya Ranjit Bhosale, Raghu Vikram, Leonardo Pimentel de Almeida Marcal, and Aparna Balachandran. Imaging of pancreatic ductal adenocarcinoma:state of the art. *World Journal of Radiology*, pages 98–105, March 2013.

[19] Jerrold T. Bushberg, J. Anthony Seibert, Edwin M. Leidholdt, and John M. Boone. *The Essential Physics of Medical Imaging*. Lippincott Williams Wilkins, Tird edition, 2012.

[20] National Comprehensive Cancer Network (NCCN). NCCN guidelines for patients - pancreatic cancer, February 2021. Available in: https://www.nccn.org/patientresources/patient-resources/guidelines-for-patients/guidelines-for-patients-details?patientGuidelineId=33.

[21] Naveen Anand and Peter Darwin. *Imaging Diagnostics in Pancreatic Cancer:A Clinical Guide*. Springer, 2021.

[22] Vikas Chaudhary and Shahina Bano. Imaging of the pancreas: Recent advances. *Indian Journal of Endocrinology and Metabolism*, pages 25–32, July 2011.

[23] Ahmed M. Zaky, Christopher L. Wolfgang, Matthew J. Weiss, Ammar A. Javed, Elliot K. Fishman, and Atif Zaheer. Tumor-vessel relationships in pancreatic ductal adenocarcinoma at multidetector ct: Different classifcation systems and their influence on treatment planning. *Radiographics*, pages 93–112, january-February 2017.

[24] Omnia Elsayed, Khaled Mahar, Mohamed Kholief, and Hatem A. Khaterr. Automatic detection of the pulmonary nodules from ct images. *SAI Intelligent Systems Conference*, November 2015.

[25] Soon Ho Yoon, Jeong Min Lee, Jae Yoon Cho, Kyung Bun Lee, Ji Eun Kim, Seung Kyoung Moon, Soo Jin Kim, Jee Hyun Baek, Seung Ho Kim, Se Hyung Kim, Jae Young Lee, Joon Koo Han, and Byung Ihn Choi. Small (20 mm) pancreatic adenocarcinomas: Analysis of enhancement patterns and secondary signs with multiphasic multidetector ct. *Radiology - RSNA Journal*, pages 442–452, May 2011.

[26] Alexander Weston. How to segment a pancreas ct, June 2021. Available in: https://towardsdatascience.com/how-to-segment-ct-pancreas-3a390acb3c70.

[27] American Joint Committee on Cancer. *AJCC Cancer Staging Manual*. American College of Surgeons, Eighth edition, 2018.

[28] Wazir Muhammad, Gregory R. Hart, Bradley Nartowt, James J. Farrell, Kimberly Johung, Ying Liang, and Jun Deng. Pancreatic cancer prediction through an artificial neural network. *Frontiers in Artificial Intelligence*, May 2019.

[29] Holger R. Roth, Le Lu, Nathan Lay, Adam P. Harrison, Amal Farag, Andrew Sohn, and Ronald M. Summers. Spatial aggregation of holistically-nested convolutional neural networks for automated pancreas localization and segmentation. *IEEE Trans. on Medical Imaging*, pages 1145–1159, January 2017.

[30] Mizuho Nishio, Shunjiro Noguchi, and Koji Fujimoto. Automatic pancreas segmentation using coarse-scaled 2d model of deep learning: Usefulness of data augmentation and deep u-net. *Applied Sciences*, May 2020.

[31] Fengze Liu, Lingxi Xie, Yingda Xia, Elliot Fishman, and Alan Yuille. Joint shape representation and classification for detecting pdac. In Heung-Il Suk, Mingxia Liu, Pingkun Yan, and Chunfeng Lian, editors, *Machine Learning in Medical Imaging*, pages 212–220. Springer, 2019.

[32] Po-Ting Chen, Dawei Chang, Huihsuan Yen, Kao-Lang Liu, Su-Yun Huang, Holger Roth, Ming-Shiang Wu, Wei-Chih Liao, and Weichung Wang. Radiomic features at ct can distinguish pancreatic cancer from noncancerous pancreas. *Radiology: Imaging Cancer*, July 2021.

[33] Zhengdong Zhang, Shuai Li, Ziyang Wang, and Yun Lu. A novel and efficient tumor detection framework for pancreatic cancer via ct images. *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 1160–1164, February 2020.

[34] Han Ma, Zhong-Xin Liu, Jing-Jing Zhang, Feng-Tian Wu, Cheng-Fu Xu, Zhe Shen, Chao-Hui Yu, and You-Ming Li. Construction of a convolutional neural network classifier developed by computed tomography images for pancreatic cancer diagnosis. *World Journal of Gastroenterology*, page 5156–5168, September 2020.

[35] Kao-Lang Liu, Tinghui Wu, Po-Ting Chen, Yuhsiang M.Tsai, Holger Roth, Ming-Shiang Wu, Wei-Chih Liao, and Weichung Wang. Deep learning to distinguish pancreatic cancer tissue from non-cancerous pancreatic tissue: a retrospective study with cross-racial external validation. *Lancet Digital Health*, pages 303–313, June 2020.

[36] Garima Suman, Anurima Patra, Panagiotis Korfiatis, Shounak Majumder, Suresh T Chari, Mark J Truty, Joel G Fletcher, and Ajit H Goenka. Quality gaps in public pancreas imaging datasets: Implications challenges for ai applications. *Pancreatology*, pages 1001–1008, August 2021.

[37] Evrim B. Turkbey Le Lu Jiamin Liu Holger R. Roth, Amal Farag and Ronald M. Summers. Data from pancreas-ct. *The Cancer Imaging Archive*, 2016. Available in: https://doi.org/10.7937/K9/TCIA.2016.tNB1kqBU.

[38] National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC). Radiology data from the clinical proteomic tumor analysis consortium pancreatic ductal adenocarcinoma [cptac-pda] collection [data set]. *The Cancer Imaging Archive*, 2018. Available in: https://doi.org/10.7937/k9/tcia.2018.sc20fo18.

[39] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. Pearson Education, Fourth edition, 2018.

[40] Sameera V. Mohd Sagheer and Sudhish N. George. A review on medical image denoising algorithms. *Biomedical Signal Processing and Control*, August 2020.

[41] Yansun Xu, John B. Weaver, Dennis M. Healy, and Jian Lu. Wavelet transform domain filters: A spatially selective noise filtration technique. *IEEE TRANSACTIONS ON IMAGE PROCESSING*, pages 747–758, NOVEMBER 1994.

[42] Parthima Guruprasad, Kushal Mahalingpur, and Manjesh.T.N. Overview of different thresholding methods in image processing. *3rd National Conference on ETACC*, June 2020.

[43] Richard Szeliski. *Computer Vision: Algorithms and Applications*. Springer London Ltd, 2011.

[44] Isaac Bankman. *Handbook of Medical Image Processing and Analysis*. Elsevier Inc., Second edition, 2009.

[45] Guorong Wu, Dinggang Shen, and Mert R. Sabuncu. *Machine Learning and Medical Imaging*. Elsevier Inc., 2016.

[46] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016.

[47] Jason Brownlee. *Master Machine Learning Algorithms:Discover How They Work and Implement Them From the Scratch*. Jason Brownlee, First edition, 2016.

[48] Sebastian Raschka and Vahid Mirjalili. *Python Machine Learning*. Packt Publishing, Tird edition, 2019.

[49] Carl Kingsford and Steven L Salzberg. What are decision trees? *Nature Biotechnology*, pages 1011–1013, September 2008.

[50] Leo Breiman. Random forests. *Machine Learning*, pages 5–32, October 2001.

[51] Prajeet Singh Kalchuri. Decision tree random forest explained with project, April 2021. Available in: https://medium.com/geekculture/decision-tree-random-forest-explained-with-project-c2c8ef1597ee.

[52] Thomas Hofmann, Bernhard Scholkopf, and Alexander J. Smola. Kernel methods in machine learning. *The Annals of Statistics*, pages 1171–1220, July 2008.

[53] Marc G. Genton. Classes of kernels for machine learning: A statistics perspective. *Journal of Machine Learning Research*, pages 299–312, December 2001.

[54] Dusty Phillip, Fabrizio Romano, Phuong Vo.T.H, Martin Czygan, Robert Layton, and Sebastian Raschka. *Python: Real-World Data Science*. Packt Publishing, 2016.

[55] Justin T. Page, Joshua A Udall, and Zach Liechty. Bambam: Genome sequence analysis tools for biologists. November 2014.

[56] Dinggang Shen, Guorong Wu, and Heung-Il Suk. Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, pages 221–248, June 2017.

[57] Bradley J. Erickson, Panagiotis Korfiatis, Zeynettin Akkus, and Timothy L. Kline. Machine learning for medical imaging. *RadioGraphics*, pages 505–515, February 2017.

[58] Pankajray. Artificial neural network (ann), June 2020. Available in: https://medium.com/ai-knowledge/artificial-neural-network-ann-ed6fa5b9c1b0.

[59] Rikiya Yamashita, Mizuho Nishio, Richard Kinh Gian Do, and Kaori Togashi. Convolutional neural networks: an overviewand application in radiology. *nsights into Imaging*, pages 611—-629, June 2018.

[60] Chirag Goyal. 20 questions to test your skills on cnn (convolutional neural networks), May 2021. Available in: https://www.analyticsvidhya.com/blog/2021/05/20-questions-to-test-your-skills-on-cnn-convolutional-neural-networks/.

[61] Pardeep Kumar and Amit Kumar Singh. *Machine Learning for Intelligent Multimedia Analytics:Techniques and Applications*. Springer Nature Singapore, 2021.

[62] Adrian Rosebrock. *Deep Learning for Computer Vision with Python*. PyImageSearch, First edition, 2017.

[63] Md Zahangir Alom, Tarek M. Taha, Chris Yakopcic, Stefan Westberg, Paheding Sidike, Mst Shamima Nasrin, Brian C Van Essen, Abdul A S. Awwal, and Vijayan K. Asari. The history began from alexnet: A comprehensive survey on deep learning approaches. *Deep Learning for Big Data analytics on High Performance Computing System*, March 2018.

[64] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition, September 2014. Available in: https://arxiv.org/abs/1409.1556.

[65] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016.

[66] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, and Vincent Vanhoucke1 Andrew Rabinovich4 Dragomir Anguelov1, Dumitru Erhan1. Going deeper with convolutions. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

[67] Sergey Ioffe and Christian Szegedy. Batch normalization: accelerating deep network training by reducing internal covariate shift. *ICML'15: Proceedings of the 32nd International Conference on International Conference on Machine Learning*, pages 448–456, July 2015.

[68] Thijs Kooi. Photometric data augmentation in projection radiography, April 2021. Available in: https://medium.com/lunit/photometric-data-augmentation-in-projection-radiography-bed3ae9f55c3.

[69] Andrew P.Bradley. The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern Recognition*, pages 1145–1159, July 1997.

[70] Francisco Melo. Area under the roc curve. In Werner Dubitzky, Olaf Wolkenhauerand Kwang-Hyun Cho, and Hiroki Yokota, editors, *Encyclopedia of Systems Biology*. Springer, 2013.

[71] Tsung-Yi Lin, Piotr Dollar, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In *Computer Vision and Pattern Recognition Conference*, pages 2117–2125, 2017.

[72] A Nugaliyadde, Kok Wai Wong, Jeremy Parry, Ferdous Sohel, Hamid Laga, Upeka V. Somaratne, Chris Yeomans, , and Orchid Foster. Rcnn for region of interest detection in whole slide images. In *Communications in Computer and Information Science*, pages 2–10. Springer, 2020.

[73] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in Neural Information Processing Systems (NIPS 2015)*. Springer, 2016.

[74] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in pytorch. In *Conference on Neural Information Processing Systems (NIPS 2017)*, 2017.

[75] COCO-common objects in context. Available in: https://cocodataset.org/home.

[76] Yucheng Zhang, Edrise M. Lobo-Mueller, Paul Karanicolas, Steven Gallinger, Masoom A. Haider, and Farzad Khalvati. Prognostic value of transfer learning based features in resectable pancreatic ductal adenocarcinoma. *Frontiers in Artificial Intelligence*, pages 1–9, October 2020.

[77] Garry Choy, Omid Khalilzadeh, Mark Michalski, Synho Do, Anthony E. Samir, Oleg S. Pianykh, J. Raymond Geis, Pari V. Pandharipande, James A. Brink, and Keith J. Dreyer. Current applications and future impact of machine learning in radiology. *Radiology - RSNA Journal*, page 318–328, June 2018.

[78] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems 25 (NIPS 2012)*, page 84–90. Association for Computing Machinery, 2017.

[79] Michele Scialpi, Alfonso Reginelli, Alfredo D'Andrea, Sabrina Gravante, Giuseppe Falcone, Paolo Baccari, Lucia Manganaro, Barbara Palumbo, and Salvatore Cappabianca. Pancreatic tumors imaging: An update. *International Journal of Surgery*, page 142–155, January 2016.

[80] Rupert W Prokesch, Lawrence C Chow, Christopher F Beaulieu, Roland Bammer, and R Brooke Jeffrey Jr. Isoattenuating pancreatic adenocarcinoma at multi-detector row ct: secondary signs. *Radiology*, page 764–768, September 2002.