

Universidade Federal de Alagoas
Instituto de Computação



Dissertação de Mestrado

**Modelo Computacional para Classificação de
Nódulos Pulmonares Utilizando Redes Neurais
Convolucionais**

Lucas Lins de Lima
l11@ic.ufal.com

Orientador:
Prof. Dr. Marcelo Costa Oliveira

Maceió, 2018

Lucas Lins de Lima

Modelo Computacional para Classificação de Nódulos Pulmonares Utilizando Redes Neurais Convolucionais

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Curso de Mestrado em Informática do Instituto de Computação da Universidade Federal de Alagoas.

Orientador:

Prof. Dr. Marcelo Costa Oliveira

Maceió, 2018

Catálogo na fonte
Universidade Federal de Alagoas
Biblioteca Central

Bibliotecário: Marcelino de Carvalho

L732m Lima, Lucas Lins de.

Modelo computacional para classificação de nódulos pulmonares utilizando redes neurais convolucionais / Lucas Lins de Lima. – 2019.

72 f. : il.

Orientador: Marcelo Costa Oliveira.

Dissertação (mestrado em Informática) - Universidade Federal de Alagoas. Instituto de Computação. Maceió, 2018.

Bibliografia: f. 67-72.

1. Neoplasias pulmonares. 2. Diagnóstico por computador. 3. Redes neurais (Computação). 4. Nódulos pulmonares. I. Título.

CDU: 004.032.26



Membros da Comissão Julgadora da Dissertação de Lucas Lins de Lima, intitulada: *"Modelo Computacional para Classificação de Nódulos Pulmonares Utilizando Redes Neurais Convolucionais"*, apresentada ao Programa de Pós-Graduação em Informática da Universidade Federal de Alagoas em 31 de janeiro de 2019, às 15h20, na sala de aula da Pós-graduação, do Instituto de Computação da UFAL.

COMISSÃO JULGADORA

Prof. Dr. Marcelo Costa Oliveira
UFAL – Instituto de Computação
Orientador

Prof. Dr. Thales Miranda Vieira
UFAL – Instituto de Computação
Examinador Interno

Prof. Dr. Paulo Mazzoncini de Azevedo Marques
USP – Universidade de São Paulo
Examinador Externo

Agradecimentos

À Deus por estar sempre comigo.

À minha família, em especial ao meu tio Quirino por ter sido um dos responsáveis a guiar-me na escolha deste curso, o qual escolhi certamente.

À minha namorada, minha princesa linda Francycelly Moura, pelo seu amor e carinho, paciência e apoio nos meus estudos.

Aos meus amigos ao longo do curso: Augusto Ícaro, Raniery, Ailton, que troco muitas idéias de estudo, Rodrigo Rozendo, Luciano Costa e Thiago Peixoto. E ao meu amigo ao longo da vida Helder Oliveira que me ajudou nas correções.

Aos meus orientadores ao longo do curso, Thales e Marcelo. E um obrigado especial ao meu orientador Marcelo pela oportunidade dada em seu laboratório, pela paciência nos ensinamentos e dedicação ao meu progresso nos estudos. Aos professores que aceitaram o convite de participarem da minha banca, Thales e Paulo. Ao Marcelo da secretaria do IC.

A todos os computadores e materiais que me serviram ao longo do curso.

A todos que me ajudaram direta ou indiretamente e torceram por mim.

Amém.

Resumo

O câncer é uma doença em que as células do organismo de repente começam a ter um crescimento desordenado e que com o tempo passam a invadir os tecidos e órgãos, espalhando-se (metástase) para outras regiões do organismo. Dentre os mais de 100 tipos de câncer, o câncer de pulmão já é o que mais mata em todo o mundo, onde em cada 5 mortes, 1 é causada por este câncer. Contudo, se o diagnóstico for realizado no início da doença as taxas de sobrevivência de 1 ano são de aproximadamente de 15-19%. A principal manifestação do câncer de pulmão se inicia através do nódulo pulmonar, que apresentam diâmetros maiores entre 3mm e 30mm. O diagnóstico do nódulo pulmonar é realizado principalmente através de imagens de Tomografia Computadorizada (TC), porém, realizar este diagnóstico ainda representa uma tarefa complexa e desafiadora para os especialistas, pois os nódulos podem estar localizados em estruturas complexas do pulmão e a quantidade de exames ou imagens que chegam para serem examinados é cada vez maior. Assim, é essencial a integração de uma ferramenta de Auxílio ao Diagnóstico por Computador (CADx) à interpretação de imagens médicas, que tem como objetivo agir como uma segunda opinião aos especialistas através de uma sugestão fornecida por um modelo computacional. Há vários trabalhos propostos na literatura para auxiliar melhor os especialistas no diagnóstico do nódulo pulmonar, seja utilizando atributos *radiomics* aliado a uma técnica de aprendizagem de máquina, seja utilizando técnicas de aprendizagem profunda, que têm chamado a atenção tanto na comunidade científica como também na indústria. Mais especificamente, Redes Neurais Convolucionais (RNCs) têm se tornado uma tendência para classificar imagens. Porém, ainda existe uma busca por um modelo computacional que melhore o desempenho de sistemas CADx. Nesse contexto, o objetivo deste trabalho foi pesquisar e desenvolver um modelo computacional para classificar nódulos pulmonares em benignos ou malignos utilizando uma RNC junto com uma técnica de otimização hiperparamétrica. O modelo conseguiu os seguintes resultados: sensibilidade de 95%, especificidade de 100%, acurácia de 85% e área sob a curva ROC (AUC) de 0,93, em um conjunto de imagens médicas composta de nódulos pulmonares sólidos de TC com diâmetro entre 3mm e 30mm. Os resultados mostraram a importância de se analisar a região do parênquima em uma proporção bem próxima dos nódulos para alcançar uma performance maior na classificação dos nódulos pulmonares, e também, que a análise do parênquima isolado leva a resultados melhores do que analisar apenas o nódulo ou o nódulo com o parênquima.

Palavras-chaves: Câncer de Pulmão; Auxílio ao Diagnóstico por Computador; Aprendizagem Profunda; Redes Neurais Convolucionais; Otimização Hiperparamétrica; Classificação.

Abstract

Cancer is a disease that occurs when the organism's cells suddenly begin to grow disorderly and over time invading tissues and organs, spreading (metastasizing) to other regions of the body. Among the more than 100 cancers, lung cancer is just the one that causes more deaths in the world, wherein every 5 deaths, 1 is caused by this cancer. However, if the diagnosis is made at the beginning of the disease, the 1-year survival rates are approximately of 15-19%. The main manifestation of lung cancer is the initiation of the pulmonary nodule, which has larger diameters between 3mm and 30mm. The diagnosis of the pulmonary nodule is performed mainly through Computed Tomography (CT) images, however, performing this diagnosis still represents a complex and challenging task for the specialists, since the nodules can be located in complex structures of the lung and the number of exams or images that come to be examined is increasing. Thus, it is essential to integrate a tool of Computer Diagnostic Assistance (CADx) in the interpretation of medical images, which aims to act as a second opinion to specialists through a suggestion provided by a computational model. There are a number of papers proposed in the literature to better assist specialists in pulmonary nodule diagnosis, whether using radiomics attributes associated with a machine learning technique or using deep learning techniques, which have called the attention both in the scientific community as well as in industry. More specifically, Convolutional Neural Networks (CNNs) has become a trend for images. However, there is still a search for a computational model that improves the performance of CADx systems. In this context, the objective of this work was to research and develop a computational model to classify lung nodules in benign or malignant using a CNN together with a hyperparametric optimization technique. The model obtained the following results: sensitivity of 95%, specificity of 100%, an accuracy of 85% and under the ROC curve (AUC) of 0.93 in a set of medical images composed of solid pulmonary nodules of CT with a diameter between 3mm and 30mm. The results showed the importance of analyzing the parenchyma region in a very close proportion of the nodules to achieve a higher performance in the classification of the pulmonary nodules, and also that the analysis of the isolated parenchyma leads to better results than analyzing only the nodule or the nodule with the parenchyma.

Keywords: Lung Cancer; Computer-Aided Diagnosis; Deep Learning; Convolutional Neural Networks; Hyperparameter Optimization; Classification.

Lista de Figuras

| | | |
|------|--|----|
| 1.1 | Exemplos de nódulos pulmonares benignos (a) e malignos (b) em 3D. Fonte: Imagem extraída de Silva et al. (SILVA; CARVALHO; GATTASS, 2005). | 2 |
| 1.2 | Exemplos de nódulos pulmonares (destacados em vermelho) com seus respectivos tamanhos e estrutura anatômica conectada a ele. Fonte: Imagem extraída de Mehdi et al. (ALILOU et al., 2014). | 3 |
| 2.1 | Espalhamento de células cancerígenas do pulmão. Fonte: Imagem extraída de (BBC, 2016). | 5 |
| 2.2 | Ilustração gráfica de um aparelho de TC convencional. O tubo de raios X gira em torno do paciente em forma de anel, emitindo um feixe colimado de raios X em forma de leque em direção aos detectores de raios X. Fonte: Imagem extraída de (BUSHBERG et al., 2011). | 6 |
| 2.3 | Demonstração de um espaço <i>voxel</i> . Fonte: Imagem extraída de (FALCÃO, 1993). | 6 |
| 2.4 | Exemplos de nódulos sólido (a), semi-sólido (b) e não sólido. Fonte: Imagem extraída de (MOSMANN et al., 2016). | 7 |
| 2.5 | Exemplos de lesões (apontadas pelas setas) de acordo com as três classificações no LIDC-IDRI (ARMATO et al., 2011). Fonte: Imagem extraída de (ARMATO et al., 2011). | 8 |
| 2.6 | Hierarquia de Inteligência Artificial, Aprendizagem de Máquina e Deep Learning. Fonte: elaborado pelo autor. | 10 |
| 2.7 | Desenho ilustrativo no neurônio humano. Os dendritos recebem impulsos elétricos de outros neurônios, que são repassados para o corpo celular, e nessa região as informações são combinadas e processadas, e dependendo da intensidade e frequência desses estímulos o Corpo celular transmite um novo impulso através de sinapses até os Dendritos de outros neurônios vizinhos pelo Axônio (FERNEDA, 2006). Fonte: Imagem extraída de (CADENE, 2016). | 10 |
| 2.8 | Figura ilustrativa do funcionamento de um neurônio artificial, onde em “Y” é o resultado calculado a partir dos nós X1, X2, . . . , Xn. Fonte: Imagem extraída de (CHARTRAND et al., 2017). | 11 |
| 2.9 | Exemplos de Arquiteturas de uma Rede Neural Artificial e de uma Aprendizagem Profunda. Fonte: elaborado pelo autor. | 12 |
| 2.10 | Ilustração de uma Rede Neural Convolucional. Fonte: Imagem extraída de (CHARTRAND et al., 2017). | 13 |
| 2.11 | Ilustração de um filtro de convolução aplicado em uma imagem com nódulo pulmonar. Fonte: Imagem extraída de (SILVA; PAIVA; SILVA, 2017). | 14 |

| | | |
|------|--|----|
| 2.12 | Exemplo onde a camada K está conectada as camadas anterior $K-1$ e posterior $K+1$, e cada neurônio ou mapa de características está conectado a três campos receptivos anteriores passando a saída a um campo receptivo posterior. Fonte: Imagem extraída de (SILVA; PAIVA; SILVA, 2017). | 14 |
| 2.13 | Subamostragem do tipo máximo. Fonte: elaborado pelo autor. | 15 |
| 2.14 | Subamostragem do tipo médio. Fonte: elaborado pelo autor. | 15 |
| 2.15 | Exemplo de curva ROC. Fonte: elaborado pelo autor. | 20 |
| 3.1 | Esquema geral aplicado neste trabalho. Fonte: elaborado pelo autor. | 26 |
| 3.2 | Processo de segmentação de um nódulo. Fonte: Imagem extraída de (FERREIRA-JUNIOR; OLIVEIRA; AZEVEDO-MARQUES, 2016) | 28 |
| 3.3 | Demonstração do processo de cálculo do diâmetro do nódulo. Fonte: Imagem extraída de (FILHO; MACHADO; OLIVEIRA, 2016). | 29 |
| 3.4 | Exemplos de nódulos benigno e maligno segmentados. Fonte: Imagem extraída de (ARMATO et al., 2011). | 30 |
| 3.5 | Ilustração da etapa de segmentação do parênquima proporcional (D) ao nódulo pulmonar. Fonte: Imagem extraída de (FILHO; MACHADO; OLIVEIRA, 2016) | 31 |
| 3.6 | Exemplificação dos três ROIs selecionados ($D/2$, $D/4$ e $D/8$) utilizando nódulo com parênquima (linha superior) e parênquima isolado (linha inferior). Fonte: elaborado pelo autor. | 32 |
| 4.1 | AUC obtidas na segmentação do parênquima na proporção $D/8$ por cortes (a) e por volume (b). Fonte: elaborado pelo autor. | 60 |

Lista de Tabelas

| | | |
|------|---|----|
| 2.1 | Parâmetros ordinários e hiperparâmetros de alguns modelos de aprendizagem de máquina, como: Support Vector Machine (SVM), MLP e RNC. Fonte: (JUNIOR; LEMOS, 2018) | 16 |
| 2.2 | Ilustração de uma Matriz de Confusão. | 19 |
| 2.3 | Trabalhos relacionados e seus resultados: Sensibilidade (SEN), Especificidade (ESP), Acurácia (ACC) e AUC. Nos grupos G1, G2 e G3 foram destacados os trabalhos com os melhores resultados. | 25 |
| 3.1 | Número de nódulos sólidos com diâmetro entre 3-30mm presentes no BNP. . . | 29 |
| 3.2 | Número de cortes dos nódulos sólidos com diâmetro entre 3-30mm presentes no BNP. | 29 |
| 3.3 | Arquiteturas exploradas neste trabalho variando o números de camadas convolucionais, <i>max-pooling</i> e totalmente conectadas. | 33 |
| 3.4 | Valores explorados dos hiperparâmetros utilizados nas CNNs. | 34 |
| 3.5 | Abordagens avaliadas nas classificações dos nódulos por volume. Majoritária se refere a maioria, ou seja, se a maior quantidade de cortes foi classificada como benigna ou maligna ou se a quantidade para os dois lados foi igual. . . . | 35 |
| 4.1 | Valores dos hiperparâmetros escolhidos nas seis arquiteturas para a classificação dos nódulos pulmonares segmentados. | 37 |
| 4.2 | Valores dos hiperparâmetros utilizados nas seis arquiteturas para a classificação dos nódulos pulmonares com o parênquima na proporção $D/2$ | 37 |
| 4.3 | Valores dos hiperparâmetros escolhidos nas seis arquiteturas para a classificação dos nódulos pulmonares com o parênquima na proporção $D/4$ | 38 |
| 4.4 | Valores dos hiperparâmetros escolhidos nas seis arquiteturas para a classificação dos nódulos pulmonares com o parênquima na proporção $D/8$ | 39 |
| 4.5 | Valores dos hiperparâmetros escolhidos nas seis arquiteturas para a classificação do parênquima segmentado na proporção $D/2$ | 39 |
| 4.6 | Valores dos hiperparâmetros selecionados nas seis arquiteturas para a classificação do parênquima segmentado na proporção $D/4$ | 40 |
| 4.7 | Valores dos hiperparâmetros selecionados nas seis arquiteturas para a classificação do parênquima segmentado na proporção $D/8$ | 40 |
| 4.8 | Resultados obtidos analisando os nódulos segmentados pelos cortes de formas balanceadas e desbalanceadas. | 41 |
| 4.9 | Resultados obtidos analisando os nódulos e parênquima (caso $D/2$) de formas balanceadas e desbalanceadas. | 42 |
| 4.10 | Resultados obtidos analisando os nódulos e parênquima (caso $D/4$) por corte de formas balanceada e desbalanceada. | 43 |

| | | |
|------|--|----|
| 4.11 | Resultados obtidos analisando os nódulos e parênquima (caso $D/8$) por corte de formas balanceada e desbalanceada. | 44 |
| 4.12 | Resultados obtidos analisando o parênquima segmentado (caso $D/2$) por corte de formas balanceada e desbalanceada. | 46 |
| 4.13 | Resultados obtidos analisando o parênquima segmentado (caso $D/4$) por corte de formas balanceada e desbalanceada. | 47 |
| 4.14 | Resultados obtidos analisando o parênquima segmentado (caso $D/8$) por corte de formas balanceada e desbalanceada. | 48 |
| 4.15 | Melhores resultados na avaliação por corte. | 49 |
| 4.16 | Resultados obtidos analisando os nódulos segmentados por volume nas abordagens 1, 2, 3, 4 e 5. | 50 |
| 4.17 | Resultados obtidos analisando os nódulos com o parênquima na proporção $D/2$ por volume nas abordagens 1, 2, 3, 4 e 5. | 52 |
| 4.18 | Resultados obtidos analisando os nódulos com o parênquima na proporção $D/4$ por volume nas abordagens 1, 2, 3, 4 e 5. | 53 |
| 4.19 | Resultados obtidos analisando os nódulos com o parênquima na proporção $D/8$ por volume nas abordagens 1, 2, 3, 4 e 5. | 55 |
| 4.20 | Resultados obtidos analisando o parênquima segmentado por volume na proporção $D/2$ nas abordagens 1, 2, 3, 4 e 5. | 56 |
| 4.21 | Resultados obtidos analisando o parênquima segmentado por volume na proporção $D/4$ nas abordagens 1, 2, 3, 4 e 5. | 58 |
| 4.22 | Resultados obtidos analisando o parênquima segmentado por volume na proporção $D/8$ nas abordagens 1, 2, 3, 4 e 5. | 59 |
| 4.23 | Melhores resultados na avaliação por corte. | 60 |
| 4.24 | Comparação com outros trabalhos publicados com foco em classificar nódulos pulmonares por meio de imagens de TC. | 62 |

Lista de Abreviaturas

| | |
|-----------|--|
| AM | Aprendizagem de Máquina |
| TC | Tomografia Computadorizada |
| CADx | do inglês <i>Computer-Aided Diagnosis</i> |
| CNN | do inglês <i>Convolutional Neural Networks</i> |
| COPD-Gene | do inglês <i>Chronic Obstructive Pulmonary Disease Genetic Epidemiology</i> |
| DL | do inglês <i>Deep Learning</i> |
| DNN | do inglês <i>Deep Neural Network</i> |
| KNN | do inglês <i>K-nearest neighbor</i> |
| LOOCV | do inglês <i>leave-one-out cross-validation</i> |
| LIDC | do inglês <i>Lung Image Database Consortium</i> |
| LIDC-IDRI | do inglês <i>Lung Image Database Consortium and Image Database Resource Initiative</i> |

| | |
|------|---|
| ML | do inglês <i>Machine Learning</i> |
| MLP | do inglês <i>Multilayer Perceptron</i> |
| MP | do inglês <i>Max-pooling</i> |
| NLST | do inglês <i>National Lung Cancer Screening Trial</i> |
| PSO | do inglês <i>Particle Swarm Optimzation</i> |
| RBF | do inglês <i>Radial Basis Function Neural Network</i> |
| RF | do inglês <i>Random Forest</i> |
| ROI | do inglês <i>Region of Interest</i> |
| RNA | Redes Neurais Artificiais |
| RNC | Redes Neurais Convolucionais |
| SAE | do inglês <i>Stacked AutoEncoder</i> |
| SMBO | do inglês <i>Sequential Model-Based Optimization</i> |
| SVM | do inglês <i>Support Vector Machine</i> |
| TC | Tomografia Computadorizada |
| TPE | do inglês <i>Tree-of-Parzen-Estimators</i> |

Sumário

| | | |
|----------|---|-----------|
| 1 | Introdução | 1 |
| 1.1 | Objetivo | 4 |
| 1.2 | Estrutura do Trabalho | 4 |
| 2 | Fundamentação Teórica | 5 |
| 2.1 | Nódulos Pulmonares em Imagens de Tomografia Computadorizada | 5 |
| 2.2 | Base LIDC-IDRI | 7 |
| 2.3 | Diagnóstico Auxiliado por Computador | 8 |
| 2.4 | Aprendizagem Profunda | 9 |
| 2.4.1 | Redes Neurais Convolucionais | 13 |
| 2.5 | Otimização Hiperparamétrica | 16 |
| 2.5.1 | Algoritmo Tree-of-Parzen-Estimators | 17 |
| 2.6 | Avaliação de desempenho | 18 |
| 2.7 | Trabalhos Relacionados | 21 |
| 3 | Materiais e Métodos | 26 |
| 3.1 | Banco de Nódulos Pulmonares | 27 |
| 3.2 | Seleção dos Nódulos | 28 |
| 3.3 | Segmentação do Parênquima | 30 |
| 3.4 | Avaliação do Parênquima | 31 |
| 3.5 | Classificação | 32 |
| 4 | Resultados e Discussão | 36 |
| 4.1 | Resultados por Corte | 41 |
| 4.2 | Resultados por Volume | 49 |
| 5 | Conclusão | 64 |
| 5.1 | Trabalhos Futuros | 65 |
| 5.2 | Contribuições Científicas | 65 |
| | Referências | 71 |

1 Introdução

Segunda o INCA (INCA, 2018), o câncer é o nome dado a um conjunto de mais de 100 doenças que têm em comum o crescimento desordenado (maligno) de células que invadem os tecidos e órgãos, podendo espalhar-se (metástase) para outras regiões do corpo. Dividindo-se rapidamente, estas células tendem a ser muito agressivas e incontroláveis, determinando a formação de tumores (acúmulos de células cancerosas) ou neoplasias malignas. Por outro lado, um tumor benigno significa simplesmente uma massa localizada de células que se multiplicam vagarosamente e se assemelham ao seu tecido original, raramente constituindo um risco de vida.

O câncer de pulmão possui a maior taxa de mortalidade dentre os tipos de câncer em todo mundo, sendo responsável por 1 em cada 5 mortes. Somente para o ano de 2012 foram estimados 1.8 milhões de casos de câncer de pulmão, destes 1.6 milhões foram a óbito. (TAMMELI; LAM, 2014). O INCA estima que 31.270 novos casos acontecerão no Brasil em 2018, sendo 18.740 nos homens e 12.530 nas mulheres (INCA, 2018).

A principal causa deste tipo de câncer é o tabagismo, seguido de outros fatores como poluição atmosférica, exposição a substâncias químicas e influência genética (INCA, 2018). Cerca de metade dos pacientes com câncer de pulmão morre dentro de um ano, devido ao fato que na maioria das vezes a doença é diagnosticada em estágios avançados, comprometendo outras regiões do organismo (ZAPPA; MOUSA, 2016). As taxas de sobrevivência do paciente diagnosticado com câncer pulmonar em estágio avançado (doença já com metástases no estágio IV) de 1 ano é de aproximadamente de 15-19%, no entanto, as taxas de sobrevivência de 1 ano podem ser aumentadas para faixa de 81-85% quando esta doença é diagnosticada em estágio inicial (estágio I) (KNIGHT et al., 2017). Logo, é necessário realizar o diagnóstico de nódulos pulmonares potencialmente cancerígenos o quanto antes pelos especialistas para aumentar as chances de sobrevivência dos pacientes, e assim também evitar desgaste físico e psicológico do paciente diante do protocolo de rastreamento do câncer de pulmão, onde o paciente poderá levar até meses para medir mudanças de tamanho, forma ou textura do nódulo, e também ficará se expondo a radiação a cada exame de Tomografia Computadorizada (TC).

A região do parênquima em torno do nódulo pode abrigar doenças microscópicas em tumores pulmonares primários que podem indicar aumento de risco de metástases a distância (DOU. et al., 2018). Estudos patológicos têm demonstrado que o tumor pulmonar pode se espalhar através do sangue e da vasculatura linfática, bem como através de espaços aéreos no parênquima pulmonar (KADOTA; AL, 2015; SHIMADA; AL, 2010). Inclusive, trabalhos recentes

((FILHO; MACHADO; OLIVEIRA, 2016; DILGER et al., 2015; DOU. et al., 2018)) demonstraram que tecidos parenquimatosos extratumorais do pulmão em torno de tumores primários podem estar infiltrados pelo câncer por metástases, apresentando ser uma importante região para ajudar no diagnóstico de pacientes com risco de ter metástases decorrente do câncer pulmonar.

A principal manifestação do câncer de pulmão, em fase inicial, é através do nódulo pulmonar, que pode ser definido como uma opacidade focal arredondada e irregular, parcialmente delimitada, cujo diâmetro varia entre 3mm e 30mm (WORMANNS; HAMER, 2015). Lesões com diâmetros maiores que 30 mm são classificadas como massas e geralmente tem natureza malignas (FUJIMOTO; WISTUBA, 2014). O diagnóstico do nódulo pulmonar é realizado através de imagens de Tomografia Computadorizada (TC), cuja técnica de diagnóstico por imagem tem sido a principal ferramenta visual utilizada pelos radiologistas para detectar nódulos pulmonares (DICIOTTI; AL., 2010).

As imagens de TC têm sido amplamente aceitas para detecção de tumores pulmonares, pois fornecem imagens 3D em alta resolução com alto contraste refletindo as diferenças de intensidade, textura e forma dos tumores, contribuindo para os radiologistas analisarem com maior precisão alguns detalhes acerca dos nódulos identificados, tais como: tamanho (altura e volume), calcificação benigna (presente ou ausente) e forma (arredondada ou espiculada) (Figura 1.1) (HUA et al., 2015a; WANG et al., 2016; HENSCHKE et al., 1999). E assim, quando um nódulo for detectado, os especialistas podem ficar acompanhando os pacientes através de imagens de TC ou realizar biópsia para determinar o estado de malignidade de um nódulo pulmonar suspeito de ser maligno (REEVES; XIE; JIRAPATNAKUL, 2015).

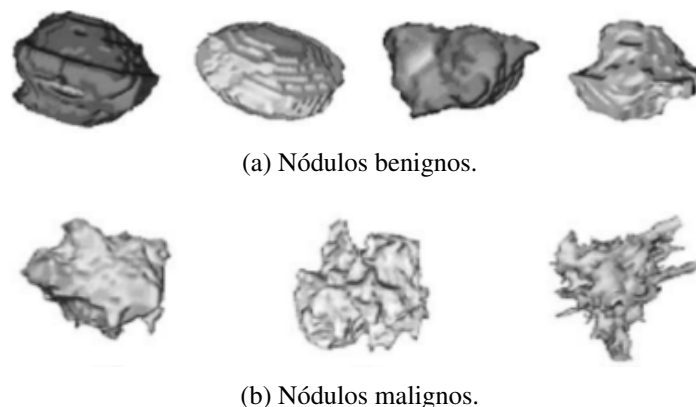
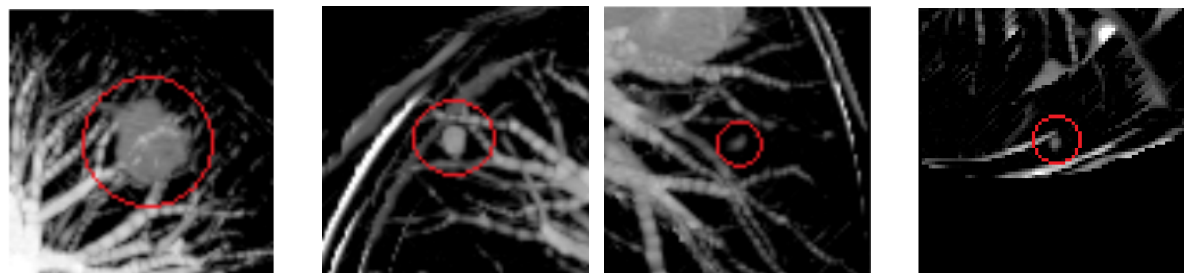


Figura 1.1: Exemplos de nódulos pulmonares benignos (a) e malignos (b) em 3D. Fonte: Imagem extraída de Silva et al. (SILVA; CARVALHO; GATTASS, 2005).

Contudo, diagnosticar o câncer de pulmão pode ser uma tarefa complexa e desafiadora para os radiologistas, pois, os nódulos podem ser pequenos, podem estar localizados em estruturas anatômicas complexas do pulmão e ainda apresentar contraste similar aos vasos e a pleura (Figura 1.2). Além disso, os radiologistas enfrentam outras dificuldades como: condições psicológicas como fadiga e pressa, condições externas como ruído nas imagens e luminosidade ruim

do ambiente de trabalho, experiência ou níveis de treinamentos diferentes e nem sempre suficientes e aumento contínuo do volume de imagens, tornando a tarefa do diagnóstico do câncer um processo realmente cansativo, e sujeito a falhas de diagnóstico, onde interpretações errôneas podem levar a cirurgias desnecessárias e erros de diagnóstico (CHUQUICUSMA et al., 2017; FERREIRA-JUNIOR; OLIVEIRA; AZEVEDO-MARQUES, 2016).



(a) 23.1mm conectado a pleura. (b) 8.3mm conectado a pleura. (c) 6.2mm isolado. (d) 5.7mm isolado.

Figura 1.2: Exemplos de nódulos pulmonares (destacados em vermelho) com seus respectivos tamanhos e estrutura anatômica conectada a ele. Fonte: Imagem extraída de Mehdi et al. (ALILOU et al., 2014).

Logo, é adequado e desejável a integração de uma ferramenta de auxílio computadorizado ao processo de diagnóstico e interpretação de imagens médicas (do inglês *Computer-Aided Diagnosis - CADx*). Sistemas CADx tem como um dos objetivos auxiliar os especialistas a melhorarem a acurácia do diagnóstico, agindo como uma segunda opinião ao especialista por intermédio de sugestão fornecida por um computador sem realizar o diagnóstico final (FERREIRA-JUNIOR; OLIVEIRA; AZEVEDO-MARQUES, 2018; CHUQUICUSMA et al., 2017).

Até recentemente, a eficácia de sistemas CADx dependia do tradicional e meticuloso *pipeline*: pré-processamento de imagens, segmentação dos nódulos pulmonares, extração de atributos, seleção dos atributos mais relevantes, e em seguida, a aplicação de uma técnica de Aprendizagem de Máquina (AM) (do inglês *Machine Learning - ML*). Além disso, muitas vezes o especialista precisava ser consultado no processo de elaboração dos atributos, tais como: contraste da lesão, padrões de espiculação e a nitidez da borda (KOOI; AL, 2017).

Nos últimos anos, as técnicas de Aprendizagem Profunda (AP, do inglês *Deep Learning - DL*) tem atraído a atenção da comunidade científica para resolver muitos problemas em diferentes áreas, como: processamento de linguagem natural, reconhecimento de fala, visão computacional, e mais especificamente, classificação e recuperação de imagens (YU et al., 2013; GUO et al., 2016). AP tem a vantagem de extrair atributos, selecionar as características mais relevantes e classificar o dado de entrada em uma única execução, sem precisar executar cada um desses passos como nas tradicionais técnicas de AM, e várias variantes metodológicas de DL têm sido aplicadas em subáreas de informática aplicada à saúde (RAVÌ et al., 2017). Logo, AP se torna uma técnica de AM atraente com a capacidade de aumentar o desempenho com-

putacional, e ainda encurtar o tedioso *pipeline* dos tradicionais sistemas CADx. Redes Neurais Convolucionais (RNC, do inglês *Convolutional Neural Network* - CNN) é um modelo arquitetural de AP e classificam imagens a partir de um modelo arquitetural eficiente com base nos seguintes componentes: convoluções, função de ativação e subamostragem (CHARTRAND et al., 2017).

Trabalhos como de Dilger et al. (DILGER et al., 2015), de Feliz et al. (FILHO; MACHADO; OLIVEIRA, 2016) e de Dou (DOU. et al., 2018) mostraram um ganho de performance na classificação dos tumores pulmonares ao levar em consideração a região do parênquima nas suas respectivas análises junto dos nódulos. Porém, o número de trabalhos que tem investigado o potencial da região do parênquima no processo de classificação de nódulos pulmonares ainda é limitado, necessitando a literatura de um aprofundamento no tema. Além disso, pelo melhor conhecimento desse autor, técnicas de DL ainda não foram aplicadas na região do parênquima para a classificação de nódulos pulmonares em benigno ou maligno.

1.1 Objetivo

O objetivo deste trabalho foi desenvolver um modelo computacional para classificar nódulos pulmonares sólidos em benignos ou malignos utilizando uma RNC junto com uma técnica de otimização hiperparamétrica.

1.2 Estrutura do Trabalho

A estrutura da dissertação está da seguinte forma:

- **Capítulo 2 - Fundamentação Teórica:** Este capítulo apresenta os principais conceitos utilizados neste trabalho como a base utilizada, bem como as técnicas utilizadas na etapa de classificação, e apresenta os trabalhos relacionados na literatura que trabalharam com a classificação de nódulos pulmonares;
- **Capítulo 3 - Materiais e Métodos:** Este capítulo apresenta como foi desenvolvido o modelo proposto para a classificação dos nódulos pulmonares;
- **Capítulo 4 - Resultados e Discussão:** Este capítulo apresenta os resultados obtidos e a discussão em torno deles e perante aos resultados obtidos pelos trabalhos relacionados;
- **Capítulo 5 - Conclusão:** Este capítulo finaliza este trabalho apresentando as conclusões e os planejamentos futuros.

2 Fundamentação Teórica

2.1 Nódulos Pulmonares em Imagens de Tomografia Computadorizada

Em imagens de Tomografia Computadorizada (TC), um nódulo pulmonar solitário (NPS) pode ser definido como uma opacidade arredondada, ou parcialmente delimitada, com diâmetro menor que 3 cm, e geralmente com densidades de partes moles ou cálcica (SILVA et al., 2010). A partir de um NPS os especialistas investigam a possibilidade deste nódulo vir a ser um tumor maligno, onde este tem a capacidade de invadir tecidos próximos, de penetrar na corrente sanguínea ou vasos linfáticos, formar tumores secundários (metástases). Logo, precisa ser contido o quanto antes para evitar que o paciente venha a óbito (Figura 2.1) (ALBERTS et al., 2010).

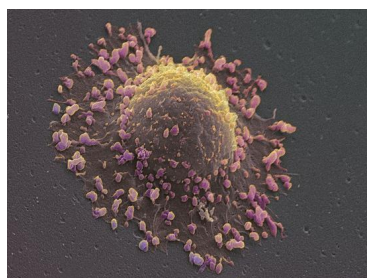


Figura 2.1: Espalhamento de células cancerígenas do pulmão. Fonte: Imagem extraída de (BBC, 2016).

A partir de imagens de TC é possível identificar nódulos entre 1mm a 2mm e visualizar os seus detalhes. Logo, o exame de TC de rotina tem se tornado o ponto inicial para a detecção de um câncer pulmonar (ERASMUS et al., 2010; DICIOTTI; AL., 2010). A TC tem como objetivo obter uma representação em 3D da estrutura interna de um objeto, como um órgão do organismo, a partir de emissões de raios X de muitas diferentes direções (Figura 2.2). As imagens são geradas em tons de cinza uniformemente espaçados de partes do corpo ou de órgãos selecionados, sendo posteriormente “empilhadas” para produzir um volume 3D de uma seção do corpo. Dependendo do *scanner* utilizado, é possível utilizar tubos de raios X mais econômicos, reduzindo assim os custos e potencialmente a dosagem (GONZALEZ; WOODS, 2010).

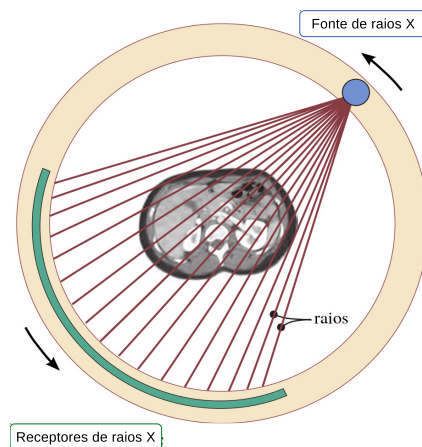


Figura 2.2: Ilustração gráfica de um aparelho de TC convencional. O tubo de raios X gira em torno do paciente em forma de anel, emitindo um feixe colimado de raios X em forma de leque em direção aos detectores de raios X. Fonte: Imagem extraída de (BUSHBERG et al., 2011).

O empilhamento das imagens de TC, mantendo o espaço original entre elas, pode ser idealizado de forma que cada *pixel* tenha a capacidade de representar o volume de um *voxel*. O *voxel* tem como característica principal possuir o valor constante em torno de uma malha de pontos central (OLIVEIRA, 2002). Um conjunto formado por *voxels* forma uma representação digital de região cuboide em estudo no paciente denominando-se “espaço *voxel*”, onde cada um desses *voxels* tem normalmente um número inteiro associado representado pelo tom em nível de cinza do *pixel* na imagem correspondente (Figura 2.3).

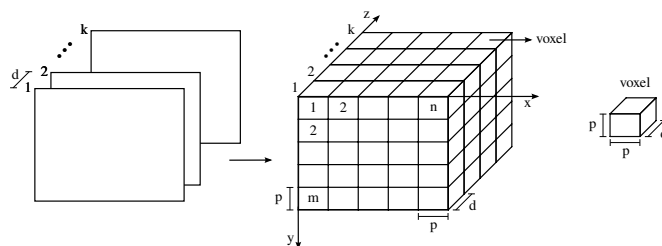


Figura 2.3: Demonstração de um espaço *voxel*. Fonte: Imagem extraída de (FALCÃO, 1993).

Em imagens de TC os nódulos pulmonares podem ser visualizados em diferentes tonalidades de intensidades, classificando-os em três categorias: sólidos, semi-sólidos ou não sólidos, também chamados de vidro-fosco (do inglês *ground glass*). Nódulos sólidos são densos e possuem uma alta intensidade, fazendo-os serem mais fáceis de serem identificados por possuírem uma aparência sólida (Figura 2.4a). Nódulos semi-sólidos constituem de parte sólida mais internamente e parte nebulosa em seu contorno, sendo assim menos densos e com intensidade média (Figura 2.4b). Já os nódulos não sólidos possuem uma aparência nebulosa, sem partes

sólidas, que mesmo possuindo uma intensidade muito menor ainda são visíveis em imagens de TC por ainda possuírem uma densidade maior do que o restante do tecido do parênquima pulmonar (Figura 2.4c).

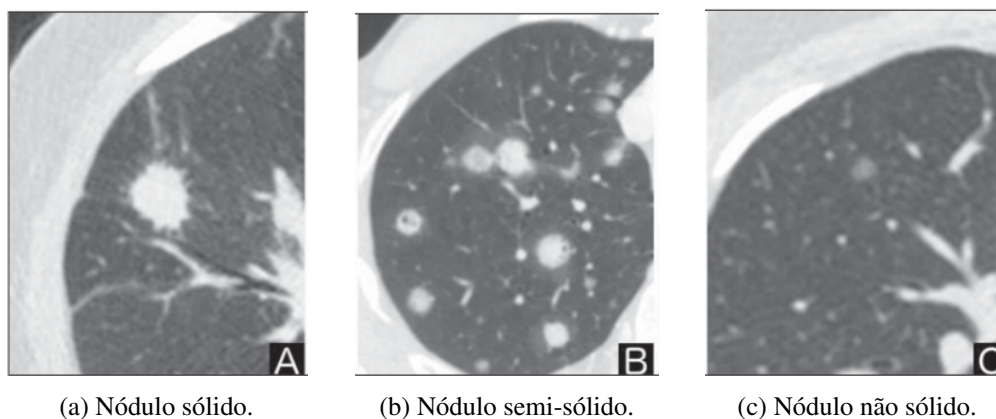


Figura 2.4: Exemplos de nódulos sólido (a), semi-sólido (b) e não sólido. Fonte: Imagem extraída de (MOSMANN et al., 2016).

2.2 Base LIDC-IDRI

A base de dados de TC pulmonares *Lung Image Database Consortium e Image Database Resource Initiative* (LIDC-IDRI) é uma base pública disponível na internet como resultado da junção entre o *Lung Image Database Consortium* (LIDC) e o *Image Database Resource Initiative* (IDRI) (ARMATO et al., 2011).

O LIDC-IDRI dispõe de 1.018 exames e 244.527 imagens de TC torácicas (no padrão DICOM) com lesões identificadas e classificadas por quatro radiologistas. O projeto LIDC-IDRI existe desde 2001 como o objetivo de desenvolver a maior base pública de TC torácicas para disponibilizar um recurso de pesquisa acessível de maneira aberta pela internet para o desenvolvimento, treinamento e avaliação de sistemas computacionais de diagnóstico e interpretação de imagens médicas (do inglês *Computed-Aided Diagnosis - CADx*) em nódulos pulmonares. As imagens estão em níveis de cinza, com resolução espacial de 512x512 elementos de imagens (pixel) e profundidade de 16 bits. Os exames foram gerados por diferentes instituições, logo, tomógrafos e protocolos distintos foram utilizados, gerando assim uma complexidade de semelhança entre as imagens (ARMATO et al., 2011; FERREIRA-JUNIOR; OLIVEIRA; AZEVEDO-MARQUES, 2016; SILVA; PAIVA; SILVA, 2016)

As lesões do LIDC-IDRI foram identificadas e classificadas de forma separada por quatro radiologistas em três categorias.

- Nódulos ≥ 3 mm de diâmetro (3.4b);
- Nódulos < 3 mm de diâmetro (3.2b);

- Não-nódulos ≥ 3 mm de diâmetro (3.2c).

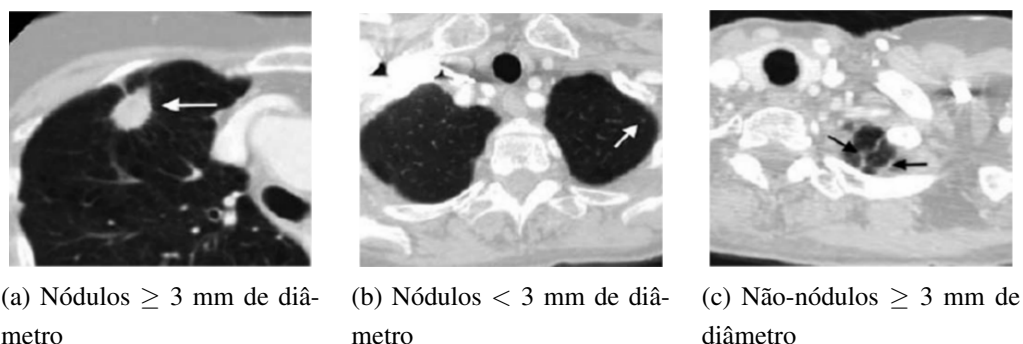


Figura 2.5: Exemplos de lesões (apontadas pelas setas) de acordo com as três classificações no LIDC-IDRI (ARMATO et al., 2011). Fonte: Imagem extraída de (ARMATO et al., 2011).

Os radiologistas após classificarem os nódulos, classificaram-os segundo nove características subjetivas: calcificação, estrutura interna, lobulação, malignância, margem, esfericidade, espiculação, sutileza e textura. No final, 7.371 lesões foram marcadas como nódulos por pelo menos um radiologista, e destas lesões, 2.669 foram classificadas por pelo menos um radiologista como nódulos, enquanto 928 lesões foram classificadas pelos quatro radiologistas como nódulo com diâmetro ≥ 3 mm, e os contornos dos nódulos demarcados por posições cartesianas 2D (x, y) em cada um dos seus cortes (ARMATO et al., 2011).

2.3 Diagnóstico Auxiliado por Computador

Sistemas computacionais de diagnóstico e interpretação de imagens médicas (do inglês *Computer-Aided Diagnosis - CADx*) tem como objetivo auxiliar os radiologistas na interpretação de imagens médicas, contribuindo para aumentar a acurácia e diminuir o tempo necessário de diagnóstico, visto que, o volume de imagens gerado a cada ano aumenta exponencialmente. Ou seja, esses sistemas agem como uma “segunda opinião” aos especialistas através de uma sugestão fornecida por um computador sem dar o diagnóstico final do exame (FERREIRA-JUNIOR; OLIVEIRA; AZEVEDO-MARQUES, 2018; CHUQUICUSMA et al., 2017; GILLIES; KINAHAN; HRICAK, 2016).

Para o âmbito do câncer pulmonar, sistemas CADx podem auxiliar na detecção e classificação de nódulos em benignos ou malignos, ou definir o estágio do câncer, fornecendo ao especialista um diagnóstico preliminar que possa auxiliá-lo na sua decisão final. Vale destacar que, sistemas CADx devem ser entendidos como sistemas que auxiliam, portanto, não podem ter seu desempenho comparado a de um ser humano, mas devem promover uma melhora das interpretações médicas com a capacidade de processamento de informações do computador (AZEVEDO-MARQUES, 2001; DOI, 2007).

Sistemas CADx auxiliam os especialistas a partir de informações extraídas das imagens, e assim usam estas informações em análises quantitativas realizadas por um computador. Estas informações são extraídas por meio de técnicas de processamento de imagens ou modelos matemáticos, que assim geram os atributos/características de forma matemática, e o desempenho de sistemas CADx para classificar nódulos pulmonares depende diretamente da escolha dos atributos que serão extraídos das imagens, e da escolha de uma técnica que vai pegar esses atributos para realizar a classificação (FILHO; MACHADO; OLIVEIRA, 2016). Esta técnica que fará a classificação dos atributos será uma técnica de Aprendizagem de Máquina (AM), onde AM pode ser definido como uma subárea da Inteligência Artificial (IA), cujo objetivo é o desenvolvimento de técnicas computacionais a partir de um processo de aprendizado para realizar alguma tarefa de forma automática (BISHOP, 2007; MITCHELL, 1997). Dentre as técnicas de IA, há a Aprendizagem Profunda (AP).

2.4 Aprendizagem Profunda

Entre as décadas de 1960 e 1970 surgiram os primeiros sistemas precursores da técnica de Aprendizagem Profunda (AP) (do inglês *Deep Learning* - DL), porém a expressão “Deep Learning” só foi criada em 2006 (SCHMIDHUBER, 2015). A popularidade da AP nos últimos anos tem crescido devido a três motivos: poder de processamento paralelo nas CPUs e GPUS, custo dos equipamentos e avanços nos algoritmos de AP (SCHMIDHUBER, 2015) (DENG, 2014).

AP é um subcampo de Aprendizagem de Máquina (AM), que por sua vez é um subcampo de Inteligência Artificial (IA) (Figura 2.6), que abrange métodos computacionais compostos de várias camadas de processamento, em que cada camada fica responsável por uma tarefa, cujo objetivo é aprender representações de dados com altos níveis de abstração. Esse aprendizado é executado a partir de dados submetidos a rede, que servirão como dados para adquirir treinamento, para assim poder conseguir executar tarefas como identificar e classificar um novo dado desconhecido a partir da aprendizagem adquirida (GUO et al., 2016). O conceito “Deep Learning” tem esse nome porque é uma técnica que busca “aprender” padrões, características dos dados, daí o nome “Learning”, através de múltiplas camadas que buscam extrair informações cada vez mais abstratas dos dados, e dessa extração de informações cada vez mais abstrata e “profunda” vem o nome “Deep”, formando assim a nomenclatura do conceito de Deep Learning.

DL oferece métodos que podem ser empregados de forma supervisionada e não supervisionada ou até mesmo as duas formas juntas (SCHMIDHUBER, 2015; ERHAN et al., 2010). A forma não supervisionada oferece a vantagem de se poder trabalhar com bases de dados não rotulados, que é uma dificuldade geralmente encontrada nas bases encontradas, principalmente no campo da medicina, mas, de todo jeito, em algum momento alguém (ou algo) precisa verificar (rotular) a clusterização para ver o que foi feito! (SHIN et al., 2016).

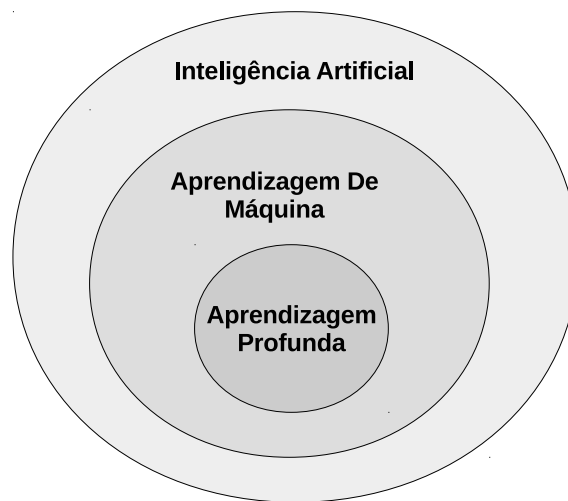


Figura 2.6: Hierarquia de Inteligência Artificial, Aprendizagem de Máquina e Deep Learning. Fonte: elaborado pelo autor.

AP é um tipo de Rede Neural Artificial (RNA) devido ao fato de a AP executar seus processamentos através de uma RNA, que são compostos por nós que buscam se comportar de forma semelhante a um neurônio humano (Figura 2.7). Assim, como em uma RNA, podemos dizer que uma AP pode ser descrita como um grafo direcionado composto por nós e arestas, onde esses nós correspondem aos neurônios e as arestas que ligam os nós correspondem às sinapses, que é por onde passam os estímulos nervosos dos neurônios.

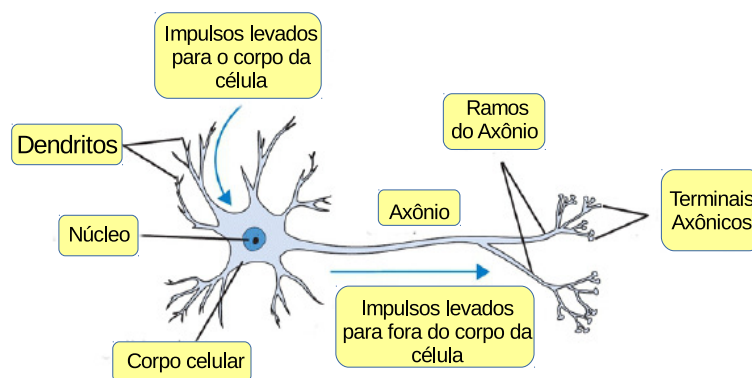


Figura 2.7: Desenho ilustrativo no neurônio humano. Os dendritos recebem impulsos elétricos de outros neurônios, que são repassados para o corpo celular, e nessa região as informações são combinadas e processadas, e dependendo da intensidade e frequência desses estímulos o Corpo celular transmite um novo impulso através de sinapses até os Dendritos de outros neurônios vizinhos pelo Axônio (FERNEDA, 2006). Fonte: Imagem extraída de (CADENE, 2016).

Nas RNAs, os nós (equivalentes aos neurônios humanos) recebem como entrada o resultado calculado por uma função “Z” (Figura 2.8). Nesta função Z é calculado um somatório, onde cada nó conectado ao nó “Z” tem um valor multiplicado por um peso, pesos estes que podem

sofrer ajustes com o objetivo de fazer a rede adquirir um maior aprendizado com relação a base de treinamento, e este somatório corresponde a quantidade de nós conectados ao nó de saída, redes sem realimentação (*feedforward*), que receberá como entrada esse somatório como mestrado na equação 2.1. O resultado do somatório é somado a um viés (“*b*” da equação 2.1) e depois aplicado a uma função de ativação (Figura 2.8) para ser computado um único valor, e dependendo do resultado este nó na rede pode ou não ser “disparado”. Se for “disparado” propaga o resultado para o nó seguinte, caso contrário não propaga o resultado, e há várias funções de ativação para calcular “*Z*” com o objetivo de normalizar o resultado, funções estas geralmente são não-lineares e deriváveis (LITJENS et al., 2017; SONODA; MURATA, 2015).

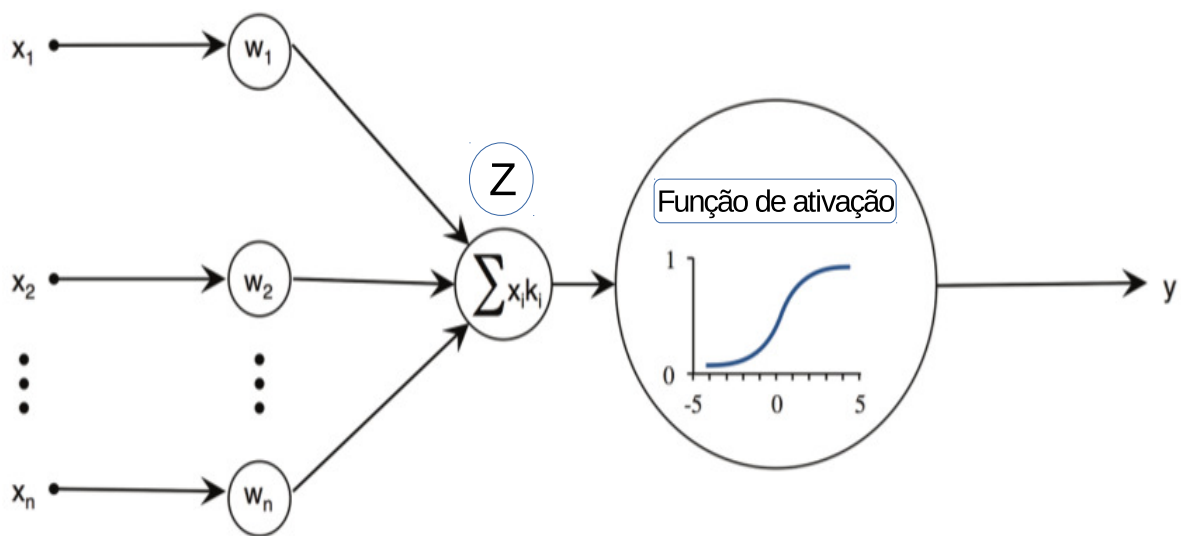


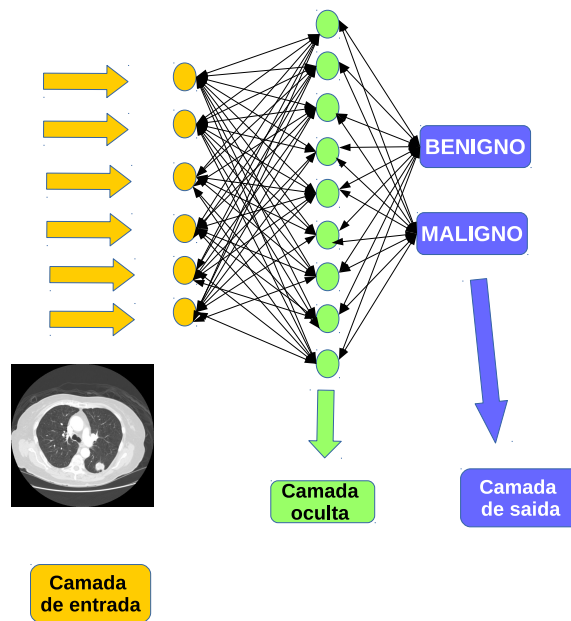
Figura 2.8: Figura ilustrativa do funcionamento de um neurônio artificial, onde em “*Y*” é o resultado calculado a partir dos nós X_1, X_2, \dots, X_n . Fonte: Imagem extraída de (CHARTRAND et al., 2017).

$$Z = \sum_{i=1}^n x_i w_i + b \quad (2.1)$$

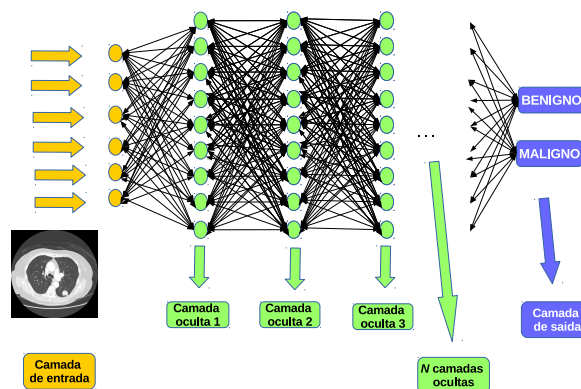
, onde X é vetor correspondente aos valores dos nós ($X_1, X_2, X_3, \dots, X_n$) e W corresponde aos pesos ($W_1, W_2, W_3, \dots, W_n$) aplicados a cada nó.

O que distingue uma RNA de uma arquitetura de AP é o número de camadas ou a profundidade da rede, distinção essa também conhecida na literatura através do termo *Credit Assignment Path* (SCHMIDHUBER, 2015). Em uma rede neural artificial costumamos encontrar apenas uma camada *hidden* ou “oculta”, a camada compreendida entre as camadas de entrada e de saída, e podemos chamar esse tipo de rede de rede “*shallow*” ou “raza” por conter apenas uma camada oculta com um número variado de nós (Figura 2.9a). Já na AP podemos encontrar n camadas ocultas, como podemos ver na arquitetura da figura 2.9b, onde temos 3 camadas ocultas, e esta arquitetura é um exemplo para ilustrar a classificação de um nódulo pulmonar entre

benigno ou maligno. Neste exemplo ou em qualquer outro que encontramos de AP aplicada a imagens, a camada oculta um é responsável por detectar a presença ou ausência de arestas, a segunda camada oculta tem a característica de detectar formas, padrões através da identificação de arranjos particulares de arestas, independentemente de pequenas variações nas posições das bordas, e a última camada oculta, a terceira, pode montar formas ou padrões em combinações maiores correspondendo a partes de objetos conhecidos (LECUNN; BENGIO; HINTON, 2015).



(a) Arquitetura de uma Rede Neural Artificial.



(b) Arquitetura de uma Aprendizagem Profunda.

Figura 2.9: Exemplos de Arquiteturas de uma Rede Neural Artificial e de uma Aprendizagem Profunda. Fonte: elaborado pelo autor.

2.4.1 Redes Neurais Convolucionais

Modelos seguindo a arquitetura do tipo RNC tem tido um impacto maior na área de informática médica, sendo considerados a abordagem mais efetiva atualmente para classificar imagens (RAVI et al., 2017).

RNCs são uma arquitetura específica entre as várias existentes que seguem o conceito de AP, com o foco específico em classificação de imagens. As RNCs combinam três ideias arquitetônicas para garantir algum grau de invariância de deslocamento, escala e distorção, que são: campos receptivos locais, compartilhamento de pesos e subamostragem espacial ou temporal (LECUN et al., 1998). As RNCs são modelos baseados na inspiração biológica, ou seja, possuem a capacidade de aprender características invariantes a rotação, translação e transformações afins de forma hierárquica, e foram propostas pela primeira vez por Fukushima em 1980 (FUKUSHIMA, 1980; LECUN; KAVUKCUOGLU; FARABET, 2010). Para uma rede de AP ser considerada do tipo RNC basta conter pelo menos uma camada de convolução entre as suas camadas, onde normalmente uma RNC é composta por 3 tipos de camadas principais: convolutiva, subamostragem e totalmente conectada, onde as camadas de convolução e subamostragem são responsáveis por extrair características das imagens e a camada totalmente conectada é responsável pelo processo de classificação dos padrões da imagem de entrada (Figura 2.10).

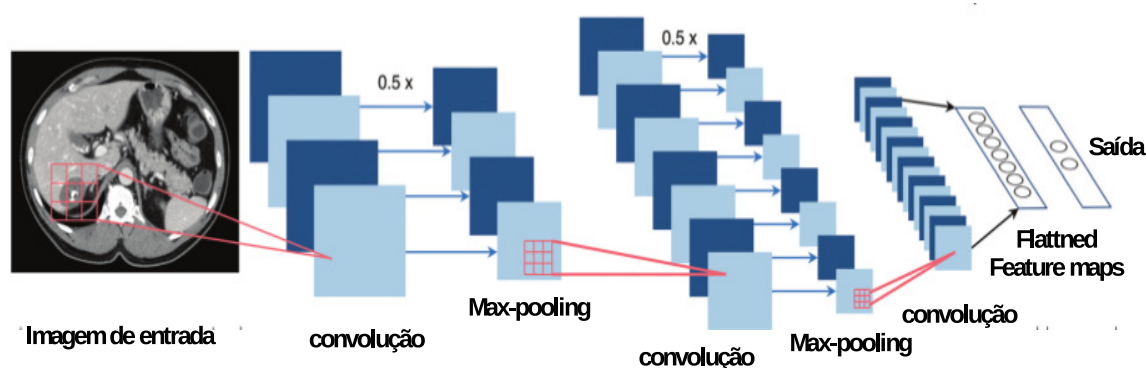


Figura 2.10: Ilustração de uma Rede Neural Convolutional. Fonte: Imagem extraída de (CHARTRAND et al., 2017).

A camada de convolução tem o papel de aplicar operações de convolução, sendo composta por filtros que são aplicados na imagem de entrada para extrair características para o treinamento da rede (Figura 2.11). Os filtros são ajustados durante a fase de treinamento para que se encontre as características mais relevantes para a construção da rede e o resultado obtido da aplicação de cada um desses filtros é denominado de mapa de características (do inglês *feature map*). Com isso, é garantido que uma determinada característica da imagem seja detectada pelo mapa de características independentemente de sua posição na imagem. Assim, a camada de convolução recebe como entrada um conjunto de características e gera a aplicação dos filtros de convolução onde cada um dos filtros tem o objetivo de destacar uma característica diferente,

como na primeira camada, por exemplo, em que as linhas e gradientes em diferentes orientações são destacadas (LECUN; KAVUKCUOGLU; FARABET, 2010).

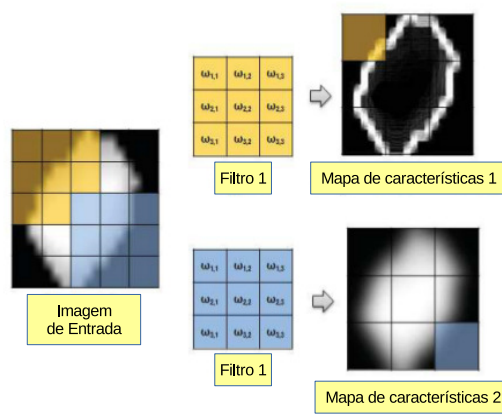


Figura 2.11: Ilustração de um filtro de convolução aplicado em uma imagem com nódulo pulmonar. Fonte: Imagem extraída de (SILVA; PAIVA; SILVA, 2017).

A entrada de uma RNC é uma imagem em 2D ou um volume de imagens 2D que irão representar um objeto 3D, assim, é destinado aos filtros uma “pequena área” 2D menor que a imagem original, geralmente uma janela 3x3 ou 5x5; e cada neurônio da camada n é conectado apenas aos neurônios desta “pequena área” da camada anterior, como podemos ver na figura 2.12.

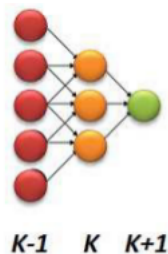


Figura 2.12: Exemplo onde a camada K está conectada as camadas anterior $K-1$ e posterior $K+1$, e cada neurônio ou mapa de características está conectado a três campos receptivos anteriores passando a saída a um campo receptivo posterior. Fonte: Imagem extraída de (SILVA; PAIVA; SILVA, 2017).

A equação de um *feature map* pode ser definido como a seguinte equação 2.2 (LECUN; KAVUKCUOGLU; FARABET, 2010):

$$(M_p) = (W_p * I) + b_p \quad (2.2)$$

onde, p corresponde ao índice do neurônio, M_p é o mapa de características de saída p , W_p é o filtro aplicado, I é o mapa de características de entrada, b_p é o viés e “*” significa o processo de convolução espacial na imagem 2D.

As camadas de subamostragem quando inseridas em uma RNC são adicionadas logo após uma camada de convolução e tem como objetivo fornecer invariância às imagens de entrada levemente diferentes e a reduzir progressivamente a resolução espacial ou dimensão dos mapas de características assegurando as *features* mais importantes, reduzindo assim a quantidade de parâmetros, o tempo de treinamento na rede, e controlando o *overfitting* (quando um modelo aprende muito bem a classificar os dados do treinamento e se mostra ineficaz nos dados de testes). Este tipo de camada permite a aprendizagem ideal de *features* invariantes à deslocamento e distorções em cada *feature map* melhorando a capacidade de generalização do conhecimento da rede (LECUN et al., 1998; CRUZ-ROA et al., 2014; PEIXOTO; CÁMARA-CHÁVEZ; MENOTTI, 2015). Dentre os tipos de subamostragens, os mais comuns são de máximo (do inglês *max-pooling*) e médio (do inglês *mean-pooling*).

O tipo de subamostragem máximo ou mais conhecido como *max-pooling* (MP) calcula o filtro de maior valor em uma janela 2D de determinado tamanho, geralmente 2 x 2, e assim gera a nova saída (Figura 2.13).

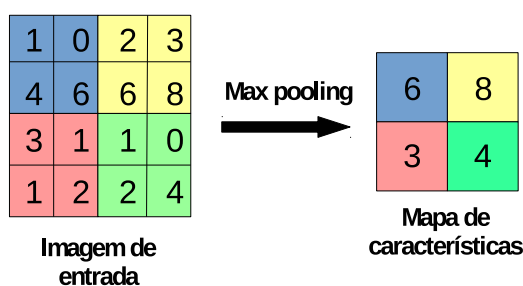


Figura 2.13: Subamostragem do tipo máximo. Fonte: elaborado pelo autor.

Já o tipo de subamostragem médio ou *mean pooling* calcula a média das intensidades em uma janela 2D de tamanho fixo, normalmente 2x2, e gera o novo conjunto de *features maps* (Figura 2.14).

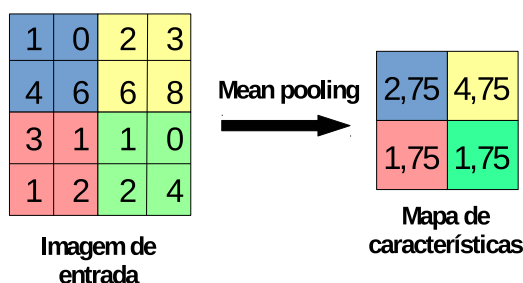


Figura 2.14: Subamostragem do tipo médio. Fonte: elaborado pelo autor.

Após a extração dos atributos da imagem de entrada usando as camadas de convolução e subamostragem, os atributos são entregues a camada totalmente conectada, que dependendo da rede pode conter mais de uma. O objetivo desta camada é classificar a imagem de entrada, e se assemelha a uma rede neural simples como a *Multilayer Perceptron* (MLP).

2.5 Otimização Hiperparamétrica

As técnicas de Aprendizagem de Máquina (AM), possuem um certo número de hiperparâmetros que se relacionam com o objetivo de obter um melhor desempenho de predição em etapas como: pré-processamento, regularização, treinamento e estrutura de um modelo. A busca por valores de hiperparâmetros ótimos possibilita alcançar um melhor desempenho e também a evitar problemas como sobre-ajuste (do inglês *overfitting*).

Otimização de hiperparâmetros (OH) é crucial para o ganho de performance de qualquer técnica de aprendizagem de máquina, podendo levar experimentos com resultados inicialmente ruins a satisfatórios, com isso, OH tem se tornado fundamental no “mundo” da aprendizagem de máquina, principalmente em técnicas de AP, onde a quantidade de hiperparâmetros é maior e dependendo da quantidade de camadas esse número tende a crescer ainda mais (Tabela 2.1). OH pode ser representada pela seguinte equação 2.3:

| Modelos | Parâmetros Ordinários | Hiperparâmetros |
|------------|-----------------------|---|
| SVM | vetores de suporte | tipo de <i>kernel</i> constante regularização constante de tolerância |
| MLP | pesos e <i>bias</i> | número de camadas e neurônios função de ativação |
| RNC | pesos e <i>bias</i> | número de camadas e seu arranjo filtros (quantidade, dimensão e deslocamento) tipo de subamostragem e dimensão da janela funções de ativação otimizador Batch Size Número de épocas |

Tabela 2.1: Parâmetros ordinários e hiperparâmetros de alguns modelos de aprendizagem de máquina, como: Support Vector Machine (SVM), MLP e RNC. Fonte: (JUNIOR; LEMOS, 2018)

$$v^* = \arg \min_{x \in T} f(x) \quad (2.3)$$

onde, $f(x)$ representa uma “pontuação objetiva” ou “função objetivo” a se minimizar (como a taxa de erro) avaliada no conjunto de validação, v^* representa o conjunto de hiperparâmetros que buscará o menor valor da “função objetivo”, e x pode ser qualquer valor no domínio de T. A “pontuação objetiva” é exatamente uma ou mais de uma métrica que desejamos minimizar (como a taxa de erro) ou a maximizar (como a acurácia). Com base nesta equação, é desejado encontrar o melhor algoritmo que consiga encontrar os melhores valores ou valores ótimos dos hiperparâmetros de uma técnica de AM em um conjunto de validação com base na “pontuação

objetiva”.

Há várias técnicas de otimização de hiperparâmetros propostas na literatura, como por exemplo: algoritmo genético, busca em grade (do inglês *grid search*), busca aleatória/randômica (do inglês *random*), e o Tree-of-Parzen-Estimators (TPE) (SILVA; PAIVA; SILVA, 2017; BERGSTRÄ; YAMINS; COX, 2013b; PINAYA et al., 2016). A busca em grade avalia todas as combinações possíveis dentro de um espaço de valores de parâmetros ou hiperparâmetros. Já a busca randômica é uma técnica em que combinações aleatórias são utilizadas para encontrar a melhor solução para o modelo de classificação proposto (BERGSTRÄ; BENGIO, 2012). Ou seja, as técnicas de busca em grade ou randômica não aplicam nenhum algoritmo “inteligente” na busca dos melhores valores de hiperparâmetros de uma técnica de ML.

Modelos de Otimização Sequencial Baseada em Modelo (Sequential Model-based Optimization - SMBO) são uma formalização de otimização Bayesiana, onde “sequencial” refere-se à execução de testes um após o outro, onde resultados anteriores são acompanhados em um histórico de pares (*score* e hiperparâmetros) para encontrar os melhores valores para os hiperparâmetros de uma técnica de AM, e esta característica não está presente nas técnicas de busca em grade e randômica. Existem várias variantes dos métodos SMBO, entre eles o Gaussian Processes, Random Forest Regressions e o TPE. Modelos SMBO se diferem pela construção do modelo de probabilidade (Eq. 2.4), que é a representação de probabilidade da função objetivo construída usando avaliações anteriores.

$$p(\textit{score} \mid \textit{hyperparameters}) \quad (2.4)$$

onde, *score* representa a “função objetivo” e *hyperparameters* os hiperparâmetros do modelo de AM em questão. Na literatura também é chamado de “substituto” para a “função objetivo” e é também representada pela equação 2.5 (DEWANCKER; MCCOURT; CLARK, 2015).

$$p(y \mid x) \quad (2.5)$$

2.5.1 Algoritmo Tree-of-Parzen-Estimators

A técnica TPE é um modelo baseado em uma Otimização Sequencial Baseada em Modelo (Sequential Model-based Optimization - SMBO) e foi proposto por Bergstra et al. (BERGSTRÄ et al., 2011). A TPE constrói um modelo de probabilidade aplicando a regra de Bayes e atualizando um modelo de probabilidade.

Assim como o modelos SMBO, o algoritmo TPE constrói o modelo de probabilidade a partir dos resultados passados e decide o próximo conjunto de hiperparâmetros para avaliar a “função objetivo” maximizando a melhoria esperada (BERGSTRÄ et al., 2011). Ou seja, com base

na minimização ou maximização da função objetivo, o TPE vai otimizar os hiperparâmetros a partir de um certo número de iterações definido. Em vez de usar a representação $p(y|x)$, o TPE segue o modelo representado na equação 2.6:

$$p(y|x) = \frac{p(x|y) * p(y)}{p(x)} \quad (2.6)$$

onde, $p(x|y)$ é a probabilidade de hiperparâmetros dado a pontuação (*score*) na “função objetivo”, e em resumo é expressado por (Eq. 2.7):

$$p(x|y) = \begin{cases} l(x) & \text{se } y < y^* \\ g(x) & \text{se } y \geq y^* \end{cases} \quad (2.7)$$

onde $y < y^*$ significa um valor menor do que a função objetivo, $y \geq y^*$ significa um valor igual ou acima da função objetivo. Com isso, podemos fazer duas distribuições diferentes dos hiperparâmetros: o valor da função objetivo pode ser menor do que um *threshold*, $l(x)$, ou maior igual do que um *threshold*, $g(x)$. E a partir destas equações (2.6 e 2.7) a técnica TPE busca os melhores valores dos hiperparâmetros selecionados em uma AM.

Logo, o TPE avalia com a função objetivo e registra o resultado em um par (*score* e hiperparâmetros), formando assim um histórico de resultados. Em seguida, o algoritmo constrói $l(x)$ e $g(x)$ usando o histórico para chegar a um modelo de probabilidade da função objetivo que melhora a cada iteração.

A técnica TPE junto com a técnica aleatória estão disponíveis no framework Hyperas, e oferecem a disponibilidade da exploração de qualquer hiperparâmetro em uma AP ou em uma RNC (MAXUMPERIA, 2017). Ainda, Bergstra et al. (BERGSTRA; YAMINS; COX, 2013a) demonstrou em seus experimentos que a técnica TPE atingiu taxas de erros menores que a técnica randômica, mostrando ser mais interessante do que a técnica randômica.

2.6 Avaliação de desempenho

A avaliação de um modelo computacional aplicado para uma classificação binária na área da saúde, é realizado para determinar quão bom é este modelo em discriminar a presença ou ausência, no caso desta dissertação, de um câncer pulmonar.

Em técnicas de aprendizagem de máquina, um classificador prediz amostras do conjunto de teste como verdadeira ou falsa. Nesta predição, quatro classificações são possíveis a partir da matriz de confusão (Tabela 2.2):

- Verdadeiro Positivo (VP): amostra corretamente classificada como maligna;

| | | Valor Predito | |
|------------------|-------------|---------------|-------------|
| | | Benigno (-) | Maligno (+) |
| Valor Verdadeiro | Benigno (-) | VN | FP |
| | Maligno (+) | FN | VP |

Tabela 2.2: Ilustração de uma Matriz de Confusão.

- Falso Positivo (FP): amostra era benigna e foi classificada como maligna;
- Verdadeiro Negativo (VN): amostra foi corretamente classificada como benigna;
- Falso Negativo (FN): amostra era maligna e foi classificada como benigna .

A partir da matriz de confusão algumas métricas podem ser extraídas/calculadas, como: sensibilidade (Equação 2.8), especificidade (Equação 2.9), acurácia (Equação 2.10), F-Measure (Equação 2.11), e taxa de erro (Equação 2.12):

$$Sensibilidade = \frac{TP}{TP + FN} \quad (2.8)$$

$$Especificidade = \frac{TN}{TN + FP} \quad (2.9)$$

$$Acurácia = \frac{TP + TN}{TP + FN + FP + FN} \quad (2.10)$$

$$F - Measure = \frac{2 * precisão * recall}{preciso + recall} \quad (2.11)$$

$$Erro = 1 - Acurácia \quad (2.12)$$

, onde TP é a taxa de verdadeiro positivo, FN é a taxa de falso negativo, TN é a taxa de verdadeiro negativo e FP é a taxa de falso positivo, e as equação de Precisão e Recall são:

$$Precisão = \frac{TP}{TP + FP} \quad (2.13)$$

$$Recall = \frac{TP}{TP + FN} \quad (2.14)$$

A métrica sensibilidade avalia a proporção de verdadeiros positivos sobre o total de nódulos malignos existentes; a especificidade avalia a proporção de verdadeiros negativos sobre o total de nódulos benignos existentes; a acurácia significa a proporção de nódulos classificados corretamente sobre o total de nódulos existentes; *f-measure* considera tanto a precisão quanto a revocação no seu cálculo e pode ser definida como a média harmônica ponderada da precisão e

da revocação na base de teste; e a taxa de erro representa a proporção de nódulos classificados de forma errada.

Uma maneira de avaliar um classificador que possuem classes binárias é através do uso das curvas ROC (do inglês *Receiving Operating Characteristics*), que mostram como o número de amostras positivas corretamente classificadas, Taxa de Verdadeiros Positivos (TVP) (do inglês *True Positive Rate*) (Equação 2.15) variam de acordo com o número de amostras negativas incorretamente classificadas como positivas, Taxa de Falsos Positivos (TFP) (do inglês *False Positive Rate*) (Equação 2.16) (LANTZ, 2015).

$$TVP = \frac{TP}{TP + FN} \quad (2.15)$$

$$TFP = \frac{FP}{FP + TN} \quad (2.16)$$

As curvas ROC podem ser visualizadas através de um gráfico bidimensional através dos eixos x (TVP ou 1 - Especificidade) e y (TFP ou Sensibilidade) (Figura 2.15). Onde, a área sob a curva (do inglês *area under the curve* (AUC)) produz valores entre 0 e 1, servindo para comparar o desempenho entre os algoritmos de classificação, e quanto mais próximo de 1 for o valor da AUC melhor será o desempenho alcançado pelo classificador (FACELI et al., 2011). Os pontos que compõem as curvas ROC indicam as taxas positivas verdadeiras em diferentes limiares de falsos positivos.

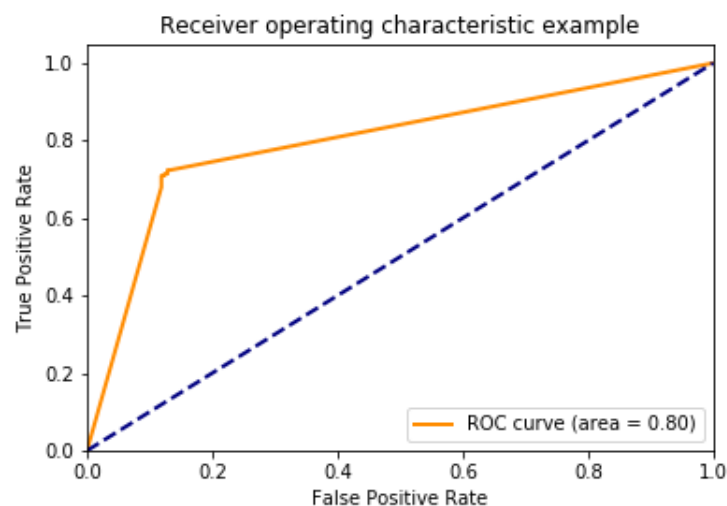


Figura 2.15: Exemplo de curva ROC. Fonte: elaborado pelo autor.

Na figura 2.15 pode ser observado algumas características do espaço ROC. Pegando a curva ROC de cor laranja, quanto mais próximo for do ponto (1,0) pior será o classificador, e quanto mais próximo do for do ponto (0,1) mais perfeito será o classificador (LANTZ, 2015).

Dentre as estratégias de amostragem existentes, há a forma *hold-out*, em que os dados de uma base são separados nos conjuntos de treinamento, validação e teste, e a proporção da sepa-

ração dos dados nestes conjuntos é definido pelo autor, não um padrão definido na literatura, e entre as opções há por exemplo: 70% (treinamento), 10% (validação) e 10% (teste). Então, o classificador adquire o seu aprendizado a partir da base de treinamento, valida os resultados na base de validação e busca aprimoramentos no aprendizado para enfim avaliar os resultados na base de teste.

2.7 Trabalhos Relacionados

Neste capítulo serão apresentados da literatura que focaram na classificação de nódulos pulmonares em benignos ou malignos e que também abordaram o parênquima pulmonar através de imagens de TC. A maioria dos trabalhos relacionados teve como foco o uso de técnicas de ML na classificação de nódulos pulmonares, porém, os últimos trabalhos publicados revelam uma certa tendência para o uso de AP. Primeiro serão apresentados os trabalhos que utilizaram técnicas tradicionais de AM (grupo 1), em seguida, serão apresentados os trabalhos que utilizaram técnicas de AP (grupo 2), e por último serão apresentados os trabalhos que abordaram o parênquima (grupo 3) (Tabela 2.3):

- Grupo 1:

Obayya et al. (OBAYYA; GHANDOUR, 2015) utilizou 100 nódulos (50 benignos e 50 malignos) do LIDC (do inglês *Lung Image Database Consortium*) com diâmetro não especificado. Obayya et al. propôs dois conjuntos de atributos extraídos: atributos derivados de análise de textura baseados em matriz de co-ocorrência de escala de cinza e atributos derivados da transformada de *Curvelet*. Na classificação foi utilizado a rede neural RBF (do inglês *Radial Basis Function Neural Network*) de forma *hold-out*, onde os conjuntos de treinamento e de teste foram compostos por 50% cada dos 100 nódulos utilizados. Os resultados alcançados foram uma taxa de acurácia de 96%, sensibilidade de 96% e especificidade de 96% utilizando os atributos da transformada de *Curvelet*.

Silva et al. (SILVA; AL, 2016) utilizaram 1.405 nódulos (1.011 benignos e 394 malignos) do LIDC com diâmetro maior que 3mm. Os autores propuseram uma metodologia utilizando os índices taxonômicos e árvores filantrópicas como descritores de textura, onde ao todo 54 atributos foram extraídos, e o algoritmo de máquina de vetores de suporte foi usado no processo de classificação. Os nódulos foram divididos em duas regiões: partes mais internas e mais externas dos nódulos. O melhor resultado alcançado foi uma taxa de acurácia de 88.44%, especificidade de 90.06%, sensibilidade de 84.22% e uma área sob a curva ROC de 0.871 utilizando índices taxonômicos.

Wang et al (WANG et al., 2016) utilizaram a base LIDC-IDRI (do inglês *Lung Image Database Consortium and Image Database Resource Initiative*) composta de 593 nódulos (271 benignos e 322 malignos), sem tamanho especificado. Os autores propuseram uma

metodologia utilizando uma análise de *radiomics* composta de 150 atributos de textura, intensidade e forma. Em seguida, foi utilizada a técnica *Random Forest* para selecionar os atributos mais relevantes. Na classificação foi utilizado o algoritmo genético para a otimização dos parâmetros de uma SVM, de forma *hold-out*, onde no conjunto de treinamento foi utilizando 200 nódulos benignos e 200 nódulos malignos, e no conjunto de teste foi utilizando 71 nódulos benignos e 122 nódulos malignos. Os resultados obtidos foram uma taxa de acurácia de 76.10%, taxa de sensibilidade de 74.60% e taxa de especificidade de 78.90%.

Dhara et al. (DHARA et al., 2016) utilizaram a base LIDC-IDRI composta de 891 nódulos (279 benignos, 349 indeterminados e 263 malignos), sem tamanho especificado. Foi extraído um conjunto de 49 atributos de atributos de forma 2D e 3D, atributos de borda 3D e atributos de textura 2D e 3D. Foi analisado três combinações diferentes: na primeira nódulos benignos e malignos, na segunda nódulos indeterminados como benignos e na terceira nódulos indeterminados como malignos. Utilizando o classificador SVM, com uma validação cruzada *5-fold*, os melhores resultados alcançados foram na primeira configuração com área sob a curva ROC de 0.950, sensibilidade de 89.73% e especificidade de 86.36%.

Ferreira Junior et al. (FERREIRA-JUNIOR; OLIVEIRA; AZEVEDO-MARQUES, 2018) utilizou a base LIDC-IDRI composta de 1.171 nódulos (745 benignos e 426 malignos) com diâmetro entre 3-30mm, e foram separados para os experimentos de forma balanceada (300 benignos e 300 malignos) e de forma desbalanceada. Ferreira Junior extraiu 36 atributos de textura 3D (AT3D) e 12 atributos de nitidez de borda 3D (ANB3D) para gerar um vetor de 48 atributos para cada nódulo. Em seguida, utilizou três técnicas de seleção de atributos para gerar mais três vetores de atributos: análise de significância estatística, método de filtragem baseado em correlação e *wrapper*, gerando assim quatro vetores de atributos para nódulo. As classificações foram efetuadas usando validação cruzada *10-fold* e os melhores resultados foram AUC de 0.858, sensibilidade de 70.2%, especificidade de 85.6% e acurácia de 80.0% usando RF na base desbalanceada formada por todos os 48 atributos.

Reeves et al. (REEVES; XIE; JIRAPATNAKUL, 2015) utilizou os bancos Lung Cancer Action Program (ELCAP) e National Lung Cancer Screening Trial (NLST) compostos juntos de 736 nódulos desbalanceados (324 benignos e 412 malignos) com diâmetro entre 3-29mm. Um total de 46 atributos divididos em quatro grupos foram extraídos: morfológico, densidade, curvatura de superfície e gradiente de borda. O melhor resultado obtido foi uma AUC de 0.772 (desvio padrão de 0.031) utilizando SVM Radial em uma validação cruzada *5-fold*. Reeves também avaliou os nódulos com diâmetro de 5-14mm de forma balanceada, e o melhor resultado foi uma AUC de 0.708 (desvio padrão de 0.062) em uma validação cruzada *5-fold*.

- Grupo 2:

Hua et al (HUA et al., 2015b) utilizou a base do LIDC-IDRI composta de 2.545 cortes com diâmetro maior que 3mm, onde todos os cortes foram redimensionados para uma resolução de 32x32. Segundo os autores este foi o primeiro trabalho a explorar técnicas de deep learning na aplicação de classificação de tumor, onde analisaram a comparação entre as redes Deep Belief Network (DBN) e CNN, e os descritores: 1º: fractais junto com o classificador K-nearest neighbor (KNN) e 2º: geométricos (*scale-invariant feature transform* (SIFT) mais o *local binary pattern* (LBP)) com o classificador Máquina de Vetores de Suporte (do inglês *support vector machine* - (SVM)). O melhor resultado obtido foi utilizando redes DBN com sensibilidade de 73.4% e especificidade de 82.2% por meio de *leave-one-out cross-validation* (LOOCV) na fase de treinamento.

Shen et al. (SHEN et al., 2017) utilizou o banco de dados LIDC-IDRI composto de 1.375 nódulos, 880 com suspeitas de malignância baixa (SMB) e 495 com suspeitas de malignância alta (SMA), com diâmetro entre 3-30mm. Shen et al. propôs uma técnica chamada de Multi-crop Convolutional Neural Network (MC-CNN), onde 9 diferentes configurações foram investigadas nas arquiteturas variando o número de kernels convolucionais de cada camada convolucional e a posição de uma camada *pooling* multi-crop para a classificação dos nódulos pulmonares. Os desempenhos foram avaliados através de uma validação cruzada *5-fold*, onde em cada rodada da validação cruzada havia originalmente 825 nódulos (528 SMB e 297 SMA) no conjunto de treinamento e 275 nódulos (176 SMB e 99 SMA) tanto no conjunto de validação como no conjunto de teste. Os resultados alcançados foram sensibilidade de 77%, especificidade de 93%, uma acurácia de 87.14% e AUC de 0.93.

Silva et al. (SILVA; PAIVA; SILVA, 2017) utilizou a base LIDC-IDRI composta de 3.243 nódulos (1.830 benignos e 1.413 malignos) com diâmetros entre 3-30mm. Primeiro, Silva et al. aplicou o algoritmo Otsu baseado no *particle swarm optimization* (PSO) para separar os nódulos em duas sub-regiões (partes mais interna e mais externa), formando assim três bases: nódulos, sub-região 1 e sub-região 2. Depois, Silva aplicou o algoritmo genético em uma CNN para encontrar as melhores quantidades de filtros nas camadas convolucionais e de número de unidades na camada totalmente conectada nas três bases simultaneamente. Na avaliação de forma *hold-out* foi separado 3.043 nódulos (1.730 benignos e 1.313 malignos) gerando 20.288 cortes (7.960 benignos e 12.328 malignos) para o treinamento e 200 nódulos (100 benignos e 100 malignos) gerando 1.343 cortes (350 benignos e 993 malignos) para o teste, o conjunto de validação teve 10% dos cortes do conjunto de treinamento. Os resultados alcançados foram uma sensibilidade de 94.66%, especificidade de 95.14%, acurácia de 94.78% e AUC de 0.949.

Song et al. (SONG et al., 2017) utilizou a base LIDC-IDRI composta de 9.106 nódulos com diâmetro entre 3-30mm. Song comparou três tipos de AP: Convolutional Neural

Network (CNN), Deep Neural Network (DNN) e Stacked AutoEncoder (SAE). Nos experimentos, 4.581 nódulos foram separados para o treinamento e 10% da base de treino foi separado para o caso de teste, em torno de 448 nódulos. Dentre os três tipos de AP avaliados, a CNN obteve os melhores resultados, com acurácia de 84.15%, sensibilidade de 83.96% e especificidade de 84.32%.

- Grupo 3:

Dilger et al. (DILGER et al., 2015) utilizou os banco de dados National Lung Screening Trial (NLST) e Chronic Obstructive Pulmonary Disease Genetic Epidemiology (COPD-Gene) compostos de 50 nódulos (28 benignos e 22 malignos) com diâmetro entre 4-30mm. Dilger et al. selecionou 47 atributos de um total de 358 a partir de quatro grupos (intensidade, forma, tamanho e textura) tanto do nódulo como também da região do parênquima pulmonar. Os melhores resultados alcançados por Dilger et al. foram alcançados na inclusão dos atributos de parênquima junto aos atributos dos nódulos, com sensibilidade de 90.91% (desvio padrão de 0.00%), especificidade de 92.86% (desvio padrão de 0.00%), acurácia de 92.0% (desvio padrão 0.00%) e AUC de 0,938 (desvio padrão de 0,009) utilizando uma Rede Neural Artificial (RNA) através de uma *leave-one-case-out cross-validation* (LOOCV), mostrando a importância do uso do parênquima.

Felix et al. (FILHO; MACHADO; OLIVEIRA, 2016) utilizou o banco de dados LIDC-IDRI composto de 214 nódulos balanceados (107 benignos e 107 malignos) com diâmetro entre 5-10mm. Os seguintes atributos foram extraídos do nódulo: 14 Atributos de Intensidade 3D (AI), 9 Atributos de Forma 3D (AF), 36 AT e 12 ANB. Do parênquima foram extraídos os seguintes atributos: 14 AI, 9 AF e 12 ANB. Em seguida foi feita a seleção dos atributos mais relevantes utilizando o Algoritmo Genético Evolutivo (AGE). O melhor resultado obtido foi uma AUC de 0.875 (desvio padrão de 0.048) com uma validação cruzada *10-fold* utilizando MLP e atributos integrados do nódulo e do parênquima.

| Grupo | Trabalho | SEN (%) | ESP (%) | ACC (%) | AUC |
|--------------|--------------------------------------|----------------|----------------|----------------|--------------|
| G1 | Obayya et al. (2015) | 96 | 96 | 96 | - |
| G1 | Silva et al. (2016) | 84.22 | 90.06 | 88.44 | 0.871 |
| G1 | Wang et al. (2016) | 74.60 | 78.90 | 76.10 | - |
| G1 | Dhara et al. (2016) | 89.73 | 86.36 | - | 0.950 |
| G1 | Ferreira Junior et al. (2018) | 70.2 | 85.6 | 80.0 | 0.858 |
| G1 | Reeves et al. (2015) | - | - | - | 0.708 |
| G2 | Hua et al. (2015b) | 73.4 | 82.2 | - | - |
| G2 | Shen et al. (2017) | 77 | 93 | 87.14 | 0.93 |
| G2 | Silva et al. (2017) | 94.66 | 95.14 | 94.78 | 0.949 |
| G2 | Song et al. (2017) | 83.96 | 84.32 | 84.15 | - |
| G3 | Dilger et al. (2015) | 90.91 | 92.86 | 92.0 | 0.938 |
| G3 | Felix et al. (2016) | - | - | - | 0.875 |

Tabela 2.3: Trabalhos relacionados e seus resultados: Sensibilidade (SEN), Especificidade (ESP), Acurácia (ACC) e AUC. Nos grupos G1, G2 e G3 foram destacados os trabalhos com os melhores resultados.

3 Materiais e Métodos

A Figura 3.1 apresenta uma visão geral da metodologia proposta. Neste trabalho foi utilizado um banco de dados de nódulos pulmonares desenvolvido em nosso laboratório de pesquisa por Ferreira Junior et al. (FERREIRA-JUNIOR; OLIVEIRA; AZEVEDO-MARQUES, 2016) composto por imagens de TC do pulmão do LIDC-IDRI (Seção 3.1). A base de imagens foi pré-processada em três etapas: seleção dos nódulos (Seção 3.2), segmentação do parênquima (Seção 3.3), e avaliação do parênquima (Seção 3.4). A seção de classificação é composta por quatro etapas: 100 iterações/repetições de (A) treinamento utilizando uma CNN, (B) avaliação na base de validação e (C) ajuste dos hiperparâmetros da CNN, para enfim ser feito a avaliação na base de teste para geração dos resultados (Seção 3.5).

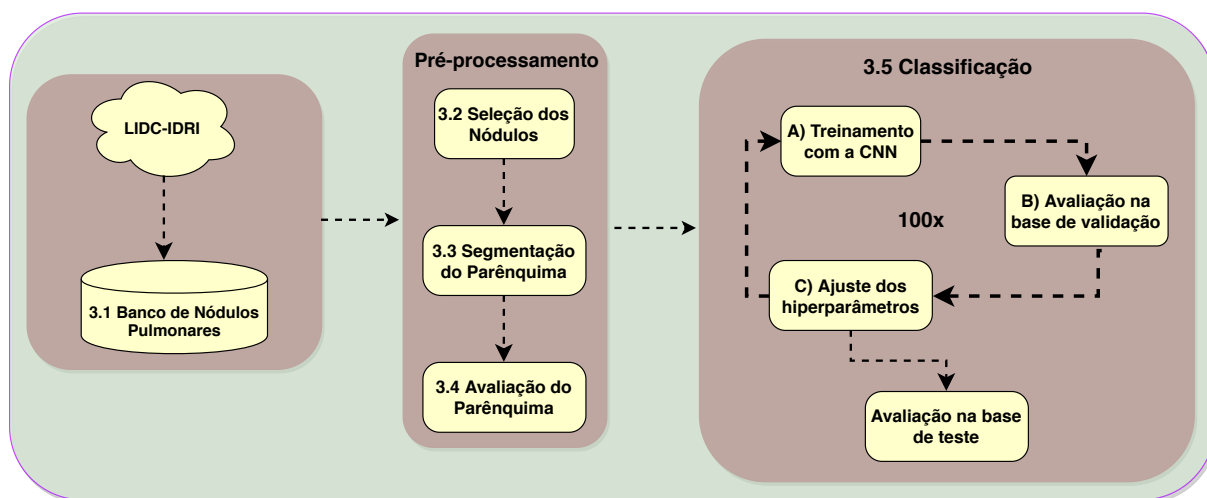


Figura 3.1: Esquema geral aplicado neste trabalho. Fonte: elaborado pelo autor.

Todos os experimentos deste trabalho foram realizados em uma Unidade de Processamento Gráfica (GPU) Nvidia Titan X com 12 GB de RAM, 3.584 núcleos, velocidade de 1,5 GHz e arquitetura Pascal. Os frameworks utilizados foram: Hyperas (MAXUMPERIA, 2017) (versão 0.4) com o Hyperopt (BERGSTRA; YAMINS; COX, 2013b) (version 0.1.1) como backend e Keras (CHOLLET et al., 2015) (versão 2.0.9) com o Tensorflow (ABADI et al., 2015) (versão 1.3.0) como *backend*.

3.1 Banco de Nódulos Pulmonares

Neste trabalho foi utilizado o repositório público LIDC-IDRI (ARMATO et al., 2011) de imagens de TC de câncer pulmonar, onde as lesões foram anotadas, identificadas e classificadas por quatro radiologistas. Esta base é composta atualmente por 1.018 exames, 1.308 estudos e 244.529 imagens.

Os radiologistas identificaram os nódulos de acordo com algumas características subjetivas e uma delas é a probabilidade de malignância em uma escala de 1 a 5, seguindo as seguintes condições:

- **Malignância 1:** probabilidade alta de ser benigno;
- **Malignância 2:** probabilidade moderada de ser benigno;
- **Malignância 3:** probabilidade intermediária;
- **Malignância 4:** probabilidade moderada de ser maligno;
- **Malignância 5:** probabilidade alta de ser maligno.

O LIDC-IDRI é um banco de dados onde todas as lesões foram manualmente segmentadas utilizando as marcações dos radiologistas. Contudo, o LIDC-IDRI é uma coleção não organizada em esquema de banco de dados, assim, não existe correlação entre imagens, anotações e classificação dos nódulos pelos radiologistas nos exames. Além disso, o LIDC-IDRI não contém informações a respeito do tamanho do nódulo. Logo, foi optado em usar o Banco de Nódulos Pulmonares (BNP) desenvolvido por Ferreira Junior et al. (FERREIRA-JUNIOR; OLIVEIRA; AZEVEDO-MARQUES, 2016). Tal banco usa uma abordagem NoSQL orientado a documentos (STRAUCH, 2011) com o MongoDB (TIWARI, 2011), onde todas as imagens e os dados dos exames estão armazenadas de forma organizada e de fácil acesso. Como o LIDC-IDRI tem anotações de quatro radiologistas, Ferreira Junior escolheu as anotações de apenas um radiologista para evitar redundâncias, sendo aquele que identificou o maior número de lesões em cada exame.

O BNP consta atualmente com 752 exames e 1.944 nódulos segmentados corte a corte manualmente no formato PNG, onde as imagens originais de TC do tórax estão armazenadas no padrão DICOM (Figura 3.2) (FERREIRA-JUNIOR; OLIVEIRA, 2015). Para aproximar as imagens ao janelamento da escala de cinza do pulmão, foram utilizados *level* de -600 e nível de janela de 1600. Neste trabalho os nódulos com probabilidade de malignância 3 foram desconsiderados por possuírem probabilidade de malignância indefinida, deixando-nos com 1.171 nódulos a disposição. Nódulos com probabilidades de malignância 1 e 2 foram considerados benignos, e nódulos com probabilidades de malignância 4 e 5 foram considerados malignos.

Os motivos pela escolha do LIDC-IDRI foram:

1. O fato da base ser pública;
2. A grande quantidade de exames já com as anotações e com os nódulos demarcados pelos radiologistas;
3. A extensa quantidade de trabalhos publicados na literatura utilizando o LIDC-IDRI para detecção, segmentação e diagnóstico de nódulos pulmonares;
4. Possibilidade de comparar os nossos resultados com outros trabalhos, visto que, o LIDC-IDIR é a base mais utilizada pela literatura;
5. O fácil acesso às imagens e informações dos exames pelo BNP.

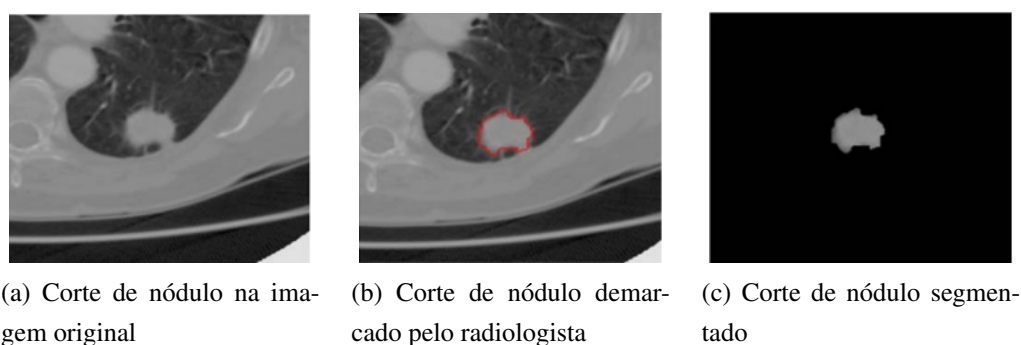


Figura 3.2: Processo de segmentação de um nódulo. Fonte: Imagem extraída de (FERREIRA-JUNIOR; OLIVEIRA; AZEVEDO-MARQUES, 2016)

3.2 Seleção dos Nódulos

O cálculo dos tamanhos dos nódulos foi realizado e armazenado no nosso BNP por Lima Filho (FILHO; MACHADO; OLIVEIRA, 2016). Onde, o cálculo de um nódulo foi extraído através de uma medição 2D, em que foram calculadas as distâncias euclidianas entre as coordenadas dos pontos mínimo e máximo no plano axial ao longo dos eixos x e y de todos os cortes de cada nódulo (Figura 3.3), e a distância de maior valor foi armazenada no nosso BNP como o tamanho do respectivo nódulo (BARTHOLMAI et al., 2015).

Um nódulo pulmonar é definido como uma opacidade focal cujo diâmetro varia de 3mm a 30mm (WORMANNS; HAMER, 2015). Logo, este foi o limiar definido, de 3mm a 30mm, um limiar inclusive adotado na maioria dos trabalhos na literatura que utilizam como base o LIDC-IDRI (HUA et al., 2015a; SILVA; PAIVA; SILVA, 2016). Para este trabalho foram considerados apenas os nódulos sólidos, assim ficamos com um total de 1.006 nódulos (Tabela 3.1).

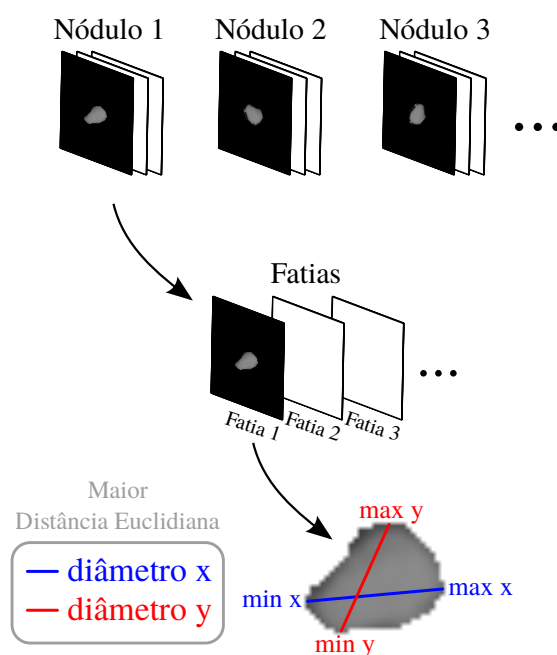


Figura 3.3: Demonstração do processo de cálculo do diâmetro do nódulo. Fonte: Imagem extraída de (FILHO; MACHADO; OLIVEIRA, 2016).

| | Benigno | | Maligno | | |
|------------------------------|---------|-----|---------|-----|-------|
| Probabilidade de Malignância | 1 | 2 | 4 | 5 | Total |
| Número de Nódulos | 304 | 394 | 171 | 137 | 1.006 |
| Soma | 698 | | 308 | | |

Tabela 3.1: Número de nódulos sólidos com diâmetro entre 3-30mm presentes no BNP.

O número de nódulos benignos e malignos presentes no BNP são diferentes e desbalanceados. Logo, como o BNP apresenta um número ainda baixo de nódulos para uma CNN, apenas 1.006 (698 benignos e 308 malignos), foram utilizados todos os cortes desses nódulos afim de aumentar o número de amostras. Partindo de um total de 5.020 cortes, foram separados 2.510 cortes benignos e malignos com diâmetros entre 3-30mm para garantir o balanceamento dos dados em todos os conjuntos (treinamento, validação e teste) nos experimentos e evitar viés nas classificações (Tabela 3.2). Por fim, foi avaliado neste trabalho os nódulos segmentados com a resolução de 64x64 (Figura 3.4).

| | Benigno | | Maligno | | |
|------------------------------|---------|-------|---------|-------|-------|
| Probabilidade de Malignância | 1 | 2 | 4 | 5 | Total |
| Número de Cortes | 1.115 | 1.395 | 1.169 | 1.341 | 5.020 |
| Soma | 2.510 | | 2.510 | | |

Tabela 3.2: Número de cortes dos nódulos sólidos com diâmetro entre 3-30mm presentes no BNP.

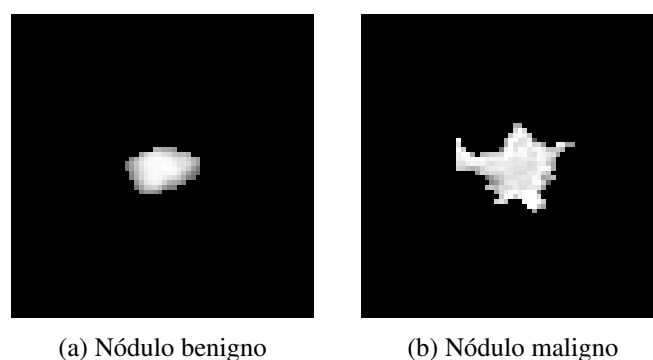


Figura 3.4: Exemplos de nódulos benigno e maligno segmentados. Fonte: Imagem extraída de (ARMATO et al., 2011).

3.3 Segmentação do Parênquima

O objetivo desta etapa foi selecionar a Região de Interesse (do inglês Region of Interest - ROI) da imagem de TC contendo os componentes de interesse deste trabalho, nódulo com parênquima e apenas o parênquima segmentado.

O processo de segmentação do parênquima já foi realizado em nosso laboratório por Lima Filho (FILHO; MACHADO; OLIVEIRA, 2016), que segmentou o parênquima e armazenou de forma proporcional ao corte central de cada nódulo (D) e armazenou no nosso BNP. A etapa de segmentação foi executada de forma automática e pode ser explicada em quatro etapas (Figura 3.5):

1. Primeiro já contamos com as demarcações das coordenadas dos nódulos no nosso BNP já realizadas pelos radiologistas do LIDC-IDRI (Figura 3.5a);
2. Para cada corte/imagem de cada nódulo, uma ROI foi estabelecida em torno do nódulo (Figura 3.5b), sendo que para se realizar esse passo foram recuperados os valores de mínimos e máximos das respectivas coordenadas x e y demarcadas pelos radiologistas, e a quantidade de parênquima selecionada em cada ROI foi definida de forma proporcional ao tamanho do nódulo em cada corte do respectivo nódulo em questão;
3. Posteriormente, foi criada uma máscara para destacar apenas a informação de interesse, o parênquima, através de um método de limiarização de imagem com base na minimização das medidas de *fuzziness* que tem como objetivo encontrar o valor de *threshold* mais apropriado para separar os objetos nas imagens, neste caso foi possível eliminar o nódulo, a pleura e os vasos presentes no parênquima, deixando apenas o tecido do parênquima em evidência (Figura 3.5c);
4. Por fim, recuperando apenas a região do parênquima em branco de (c) em (b), o resultado final foi uma imagem com apenas o parênquima segmentado em cada um dos cortes que

continham os respectivos nódulos (HUANG; WANG, 1995) (Figura 3.5d).

Neste trabalho a segmentação foi realizada de forma proporcional a cada corte dos nódulos, e não proporcional ao corte principal dos nódulos, como foi feito por Lima Filho.

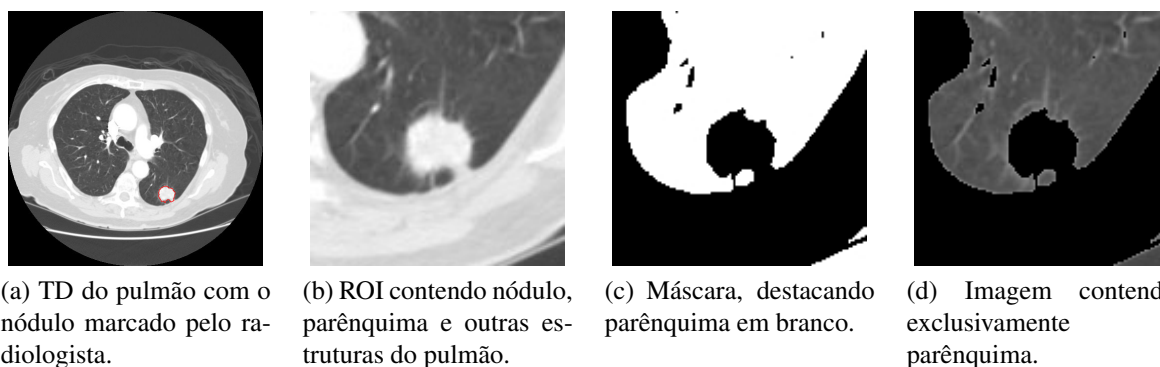


Figura 3.5: Ilustração da etapa de segmentação do parênquima proporcional (D) ao nódulo pulmonar. Fonte: Imagem extraída de (FILHO; MACHADO; OLIVEIRA, 2016)

3.4 Avaliação do Parênquima

Com o objetivo de avaliar um tamanho adequado da ROI no processo de classificação dos nódulos pulmonares, neste trabalho foram definidas três ROIs distintas proporcionais ao tamanho dos nódulos de duas maneiras: a imagem original do nódulo com o parênquima (Figura 3.5b) e do parênquima segmentado (Figura 3.5d). As ROIs foram selecionadas a partir de uma medida que pegasse a menor proporção de parênquima possível do nódulo ($1/8$), ou seja, o tamanho exato do diâmetro dos cortes de cada nódulo, onde uma medida menor do que esta seria selecionado a região interna do nódulo excluindo assim a região do parênquima, até uma proporção um pouco maior ($1/2$), mas não tão grande para evitar de pegar ruídos ou região do parênquima que não já pudesse fornecer mais detalhes de malignância dos nódulos, deixando-nos assim com seis bases para avaliar o parênquima :

- Proporcional a $1/2$ do tamanho de cada corte do nódulo ($D/2$) (Figura 3.6a e 3.6d);
- Proporcional a $1/4$ do tamanho de cada corte do nódulo ($D/4$) (Figura 3.6b e 3.6e);
- Proporcional a $1/8$ do tamanho de cada corte do nódulo ($D/8$) (Figura 3.6c e 3.6f).

Com o objetivo de tirar o máximo de fundo preto nas imagens das seis bases geradas e assim permitir que as CNNs pudessem ser mais objetivas para extrair as informações das imagens, cada base teve uma resolução diferente:

- $D/2$: 74×74 ;

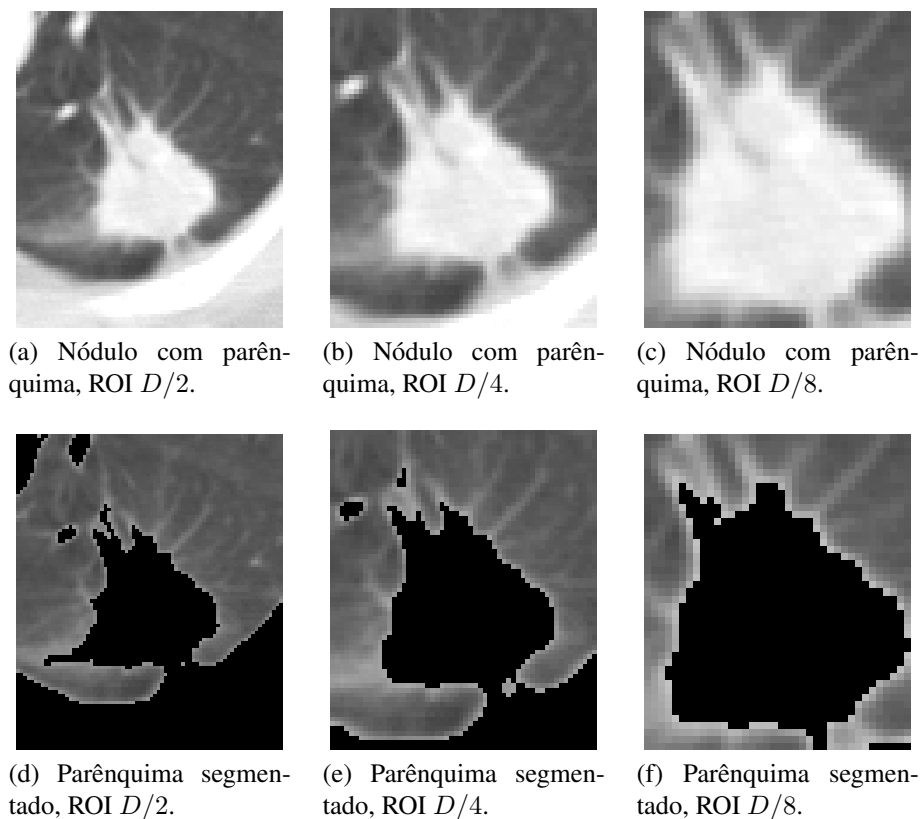


Figura 3.6: Exemplificação dos três ROIs selecionados ($D/2$, $D/4$ e $D/8$) utilizando nódulo com parênquima (linha superior) e parênquima isolado (linha inferior). Fonte: elaborado pelo autor.

- $D/4$: 56x56;
- $D/8$: 38x38.

3.5 Classificação

Neste trabalho, para analisar os melhores valores de cada hiperparâmetro, o número de camadas convolucionais e totalmente conectadas em uma CNN, foi utilizado o algoritmo Tree-of-Parzen-Estimators (TPE) através da ferramenta Hyperas, que pode ser integrado com o framework Keras, onde o Hyperas trabalha em torno da ferramenta Hyperopt, porém com mais praticidade de se trabalhar (BERGSTRA; YAMINS; COX, 2013b; CHOLLET et al., 2015; MAXUMPERIA, 2017).

Para determinar o desempenho em relação as configurações das CNNs, foram avaliados diferentes configurações das CNNs com relação ao número de camadas convolucionais (m) e ao número de camadas totalmente conectadas (n), onde $m = \{1, 2, 3\}$ e $n = \{1, 2\}$, sem mencionar a última camada totalmente conectada que estará sempre fixada para concretizar a classificação, então 6 arquiteturas foram analisadas (Tabela 3.3). Cada camada convolucional foi seguida de

uma camada de subamostragem do tipo *max-pooling*, e devido ao tamanho de nossas imagens em todas as sete bases utilizadas neste trabalho (Seções 3.2 e 3.4), foi limitado a exploração das CNNs em até 3 camadas convolucionais.

Nas camadas convolucionais foram explorados dois hiperparâmetros: o número de filtros e o tamanho do *kernel*. Na camada de *max-pooling* foi explorado o tamanho do kernel. Como foram utilizados imagens com resoluções baixas, foram avaliados apenas três e dois tamanhos do *kernel* nas camadas de convolução e de *max-pooling*, respectivamente, a fim de explorar mais camadas de convolução e de *max-pooling* através das 6 arquiteturas exploradas. Nas camadas totalmente conectadas foi explorado o número de unidades, e depois de cada camada totalmente conectada, exceto a última, foi avaliado o valor do hiperparâmetro DROP. Avaliamos os otimizadores, e para isto foram selecionados todos os otimizadores (Oti.) disponíveis no Keras: RMSprop, Stochastic Gradient Descent (SGD), Adagra, Nadam, Adam, Adamax e Adadelta (CHOLLET et al., 2015; RUDER, 2016) e também foi analisado o melhor valor de *Batch Size* (BS) com o objetivo de se extrair o melhor aproveitamento de treinamento para a CNN correspondente. A fim de se encontrar o número ideal de épocas (hiperparâmetro *epochs*) foi utilizado a função “*Early Stopping*” fornecida pelo framework Keras, em que quando não se há mais ganho de performance ou melhora de acurácia, o treinamento é automaticamente interrompido colaborando também para que o problema de *overfitting* não ocorra. Na tabela 3.4 está resumido a exploração dos hiperparâmetros das 6 arquiteturas (Tabela 3.3) nas sete bases.

| Arquiteturas | Camadas | | |
|--------------|-----------------|-------------|----------------------|
| | Convolutacional | Max-pooling | Totalmente conectada |
| 1 | 1 | 1 | 1 |
| 2 | 2 | 2 | 1 |
| 3 | 2 | 2 | 2 |
| 4 | 3 | 3 | 1 |
| 5 | 3 | 3 | 2 |
| 6 | 3 | 3 | 3 |

Tabela 3.3: Arquiteturas exploradas neste trabalho variando o números de camadas convolucionais, *max-pooling* e totalmente conectadas.

| Camada | Hiperparâmetro | Alcance |
|-----------------------------|--------------------------|-------------------------------|
| Convolutacional | Número de filtros | $\mathbb{Z} \subset [1, 512]$ |
| | Tamanho do <i>kernel</i> | (3, 3), (4, 4) e (5, 5) |
| Max-Pooling | <i>kernel</i> | (2, 2) e (3, 3) |
| Totalmente conectada | Unidades | $\mathbb{Z} \subset [2, 512]$ |
| | <i>Dropout</i> | $\mathbb{R} \subset [0, 1]$ |

| | | |
|--|-------------------|--|
| | Otimizador | RMSprop SGD Adagrad Nadam Adam Adamax Adadelta |
| | <i>Batch size</i> | $\mathbb{Z} \subset [1, 50]$ |

Tabela 3.4: Valores explorados dos hiperparâmetros utilizados nas CNNs.

Devido as baixas resoluções das sete bases utilizadas, os *kernels* precisaram ser fixados ou menos explorados em algumas camadas de convolução e de *max-pooling*. Na base dos nódulos segmentados a resolução foi de 64x64, então na terceira camada de convolução o *kernel* foi explorado apenas nos tamanhos 3x3 e 4x4, e na terceira camada de *max-pooling* o *kernel* foi fixado em 2x2. Na base $D/2$ a resolução foi de 74x74, logo o *kernel* da terceira camada de *max-pooling* precisou ser fixada em 2x2 nas arquiteturas 4, 5 e 6. Na base $D/4$ a resolução foi de 56x56, então os *kernels* da segunda e terceira camadas de *max-pooling* foram fixados em 2x2 quando foi avaliado as arquiteturas com três camadas de convolução, as arquiteturas quatro, cinco e seis. Na base $D/8$ a resolução foi de 38x38, a menor resolução das sete bases utilizadas, então o *kernel* da segunda camada de convolução teve a exploração feita entre os tamanhos 3x3 e 4x4, na terceira camada de convolução o *kernel* foi fixado em 3x3, e todas as camadas de *max-pooling* nas arquiteturas 4, 5 e 6 tiveram os *kernels* fixados em 2x2.

Para determinar o desempenho das arquiteturas CNNs propostas, primeiro os nódulos foram embaralhados e divididos em três conjuntos para a realização dos experimentos de forma *hold-out*: 70% para treinamento, 10% para validação e 20% para teste. Sendo a separação feita de forma balanceada entre os cortes benignos e malignos e foi assegurado que a separação dos cortes para treinamento, validação e teste não fizessem parte do mesmo nódulo, evitando assim viés na etapa de classificação. O modelo de classificação realizou 100 iterações utilizando o algoritmo TPE, buscando-se maximizar as funções objetivo especificidade e sensibilidade, nas bases de treinamento e validação para encontrar os melhores valores dos hiperparâmetros da CNN em análise, e em seguida os resultados foram gerados a partir da avaliação do classificador na base de teste.

Os cortes dos nódulos nos conjuntos de treinamento e validação foram tratados como casos individuais, mas os cortes do conjunto de testes foram avaliados tanto de forma individual (balanceados e desbalanceados) como também pelo volume dos nódulos. Na avaliação por volume, como foi utilizado CNNs 2D na classificação de cada corte do nódulo, os resultados foram gerados considerando o volume dos nódulos de cinco maneiras (Tabela 3.5):

| |
|--|
| 1^a - Classificação majoritária de todas os cortes, se der empate, classificação benigna; |
| 2^a - Classificação majoritária de todas os cortes, se der empate, classificação maligna; |
| 3^a - Classificação baseado na predição máxima das predições de todas os cortes; |
| 4^a - Classificação baseado na média das predições de todas os cortes; |
| 5^a - Caso um corte fosse classificado como maligno, classificação maligna. |

Tabela 3.5: Abordagens avaliadas nas classificações dos nódulos por volume. Majoritária se refere a maioria, ou seja, se a maior quantidade de cortes foi classificada como benigna ou maligna ou se a quantidade para os dois lados foi igual.

As métricas utilizadas para avaliar os resultados deste trabalho foram: taxa de erro, sensibilidade, especificidade, acurácia, F-Measure e área sob a curva ROC (Seção 2.6). Tais métricas são utilizadas na literatura em avaliações de técnicas de aprendizagem para classificação de imagens. Através destas métricas, foi possível avaliar os resultados do trabalho proposto como promissores ou não, identificar pontos positivos e negativos para no futuro serem realizados aprimoramentos.

Na análise por volume, foram avaliados as cinco abordagens (Tabela 3.5). Em todas as abordagens foram calculadas as métricas: taxa de erro (Err), sensibilidade (SEN), especificidade (ESP), *f-measure* (FM), acurácia (ACC) e a área sob a curva ROC (AUC). Porém, a AUC só foi calculada nas abordagens três e quatro, pois como a biblioteca *scikit-learn* só lida com as predições para o cálculo da AUC estas foram as únicas abordagens em que trabalhamos com as predições, e assim foi possível calcular a AUC apenas nessas abordagens.

Nas abordagens 1 e 2, onde foi calculado o número majoritário de cortes de cada nódulo, houve casos em que a quantidade de cortes classificados como benigno e maligno foi igual, e essa quantidade será mostrada em “Número (N^o) de empate”.

4 Resultados e Discussão

Neste capítulo serão apresentados os resultados obtidos através da metodologia apresentada no capítulo anterior para a classificação de nódulos pulmonares em imagens de TC.

A seguir, será apresentado os valores dos hiperparâmetros que foram definidos de forma fixa e que foram escolhidos pelo algoritmo TPE nas seis arquiteturas. A partir destas arquiteturas, os resultados foram calculados na base de teste dos nódulos de duas formas: cada corte foi avaliado de forma individual tanto de forma balanceada como também de forma desbalanceada, e por volume de forma balanceada. Os valores dos hiperparâmetros escolhidos nas camadas de convolucional e de *max-pooling* serão mostrados na primeira camada (1^a), na segunda camada (2^a) e por último na terceira camada (3^a). Já na camada totalmente conectada os valores dos hiperparâmetros serão mostrados de trás pra frente, ou seja, partindo da camada de classificação, a primeira (1^a), depois a segunda (2^a), e por último a terceira camada (3^a), a camada logo após a camada *Flatten*. O símbolo (*) vai indicar que os valores foram previamente fixados e não foram escolhidos pelo TPE.

A seguir, na tabela 4.1 estão os valores dos hiperparâmetros (Seção 3.4) escolhidos pelo TPE nas seis arquiteturas (Seção 3.3) exploradas para a classificação dos nódulos segmentados.

| | Camadas | | | | | Oti. | BS |
|----------|----------------------|--------|---------|----------------------|------|----------|----|
| | Convolução | | MP | TC | | | |
| Arqui. | Filtro | kernel | kernel | Unidades | DROP | | |
| 1 | 1 ^a : 433 | (4, 4) | (2, 2) | 1 ^a : 2* | - | Adamax | 27 |
| 2 | 1 ^a : 210 | (4, 4) | (2, 2) | 1 ^a : 2* | - | RMSProp | 47 |
| | 2 ^a : 495 | (3, 3) | (2, 2) | | | | |
| 3 | 1 ^a : 317 | (4, 4) | (2, 2) | 1 ^a : 2* | - | Adam | 19 |
| | 2 ^a : 297 | (4, 4) | (3, 3) | 2 ^a : 255 | 0.52 | | |
| 4 | 1 ^a : 294 | (5, 5) | (3, 3) | 1 ^a : 2* | - | SGD | 33 |
| | 2 ^a : 279 | (5, 5) | (3, 3) | | | | |
| | 3 ^a : 114 | (3, 3) | (2, 2)* | | | | |
| 5 | 1 ^a : 177 | (5, 5) | (3, 3) | 1 ^a : 2* | - | Adadelta | 11 |
| | 2 ^a : 409 | (3, 3) | (3, 3) | 2 ^a : 476 | 0.06 | | |
| | 3 ^a : 228 | (3, 3) | (2, 2)* | | | | |
| 6 | 1 ^a : 301 | (5, 5) | (2, 2) | 1 ^a : 2* | - | Adamax | 4 |

| | | | | | | | |
|--|----------------------|--------|---------|----------------------|------|--|--|
| | 2 ^a : 264 | (4, 4) | (2, 2) | 2 ^a : 351 | 0.51 | | |
| | 3 ^a : 287 | (3, 3) | (2, 2)* | 3 ^a : 394 | 0.26 | | |

Tabela 4.1: Valores dos hiperparâmetros escolhidos nas seis arquiteturas para a classificação dos nódulos pulmonares segmentados.

A partir de agora serão mostrados os valores dos hiperparâmetros escolhidos na avaliação dos nódulos com o parênquima nas proporções $D/2$, $D/4$ e $D/8$.

A seguir, na tabela 4.2 estão os valores dos hiperparâmetros (Seção 3.4) escolhidos pelo TPE nas seis arquiteturas (Seção 3.3) exploradas para a classificação dos nódulos com o parênquima na proporção $D/2$.

| Arqui. | Camadas | | | | | Oti. | BS |
|----------|--|----------------------------|-----------------------------|--|-------------------|----------|----|
| | Convolução | | MP | TC | | | |
| | Filtro | kernel | kernel | Unidades | DROP | | |
| 1 | 1 ^a : 228 | (5, 5) | (3, 3) | 1 ^a : 2* | - | Adadelta | 28 |
| 2 | 1 ^a : 33 2 ^a : 506 | (5, 5) (4, 4) | (3, 3) (2, 2) | 1 ^a : 2* | - | Nadam | 34 |
| 3 | 1 ^a : 360 2 ^a : 328 | (4, 4) (3, 3) | (2, 2) (3, 3) | 1: 2* 2 ^a : 102 | - 0.43 | Nadam | 29 |
| 4 | 1 ^a : 407 2 ^a : 428 3 ^a : 488 | (4, 4) (5, 5) (5, 5) | (2, 2) (3, 3) (2, 2)* | 1 ^a : 2* | - | Adadelta | 8 |
| 5 | 1 ^a : 398 2 ^a : 397 3 ^a : 381 | (5, 5) (5, 5) (5, 5) | (3, 3) (3, 3) (2, 2)* | 1 ^a : 2* 2 ^a : 130 | - 0.20 | Nadam | 11 |
| 6 | 1 ^a : 101 2 ^a : 4 3 ^a : 238 | (5, 5) (4, 4) (4, 4) | (2, 2) (2, 2) (2, 2)* | 1: 2* 2 ^a : 39 3 ^a : 112 | - 0.25 0.02 | Adagrad | 42 |

Tabela 4.2: Valores dos hiperparâmetros utilizados nas seis arquiteturas para a classificação dos nódulos pulmonares com o parênquima na proporção $D/2$.

A seguir, na tabela 4.3 estão os valores dos hiperparâmetros (Seção 3.4) escolhidos pelo TPE nas seis arquiteturas (Seção 3.3) exploradas para a classificação dos nódulos com o parênquima na proporção $D/4$.

| Arqui. | Camadas | | | | | Oti. | BS |
|----------|----------------------|--------|--------|---------------------|------|---------|----|
| | Convolução | | MP | TC | | | |
| | Filtro | kernel | kernel | Unidades | DROP | | |
| 1 | 1 ^a : 276 | (5, 5) | (2, 2) | 1 ^a : 2* | - | Adagrad | 7 |

| | | | | | | | | | |
|----------|----------------------|--------|---------|---------------------|---|---------|----|----------------------|------|
| 2 | 1 ^a : 224 | (5, 5) | (2, 2) | 1 ^a : 2* | - | SGD | 38 | | |
| | 2 ^a : 276 | (4, 4) | (3, 3) | | | | | | |
| 3 | 1 ^a : 267 | (4, 4) | (3, 3) | 1 ^a : 2* | - | Nadam | 31 | | |
| | 2 ^a : 305 | (4, 4) | (2, 2) | | | | | 2 ^a : 510 | 0.85 |
| 4 | 1 ^a : 63 | (4, 4) | (3, 3) | 1 ^a : 2* | - | Adamax | 42 | | |
| | 2 ^a : 83 | (4, 4) | (2, 2)* | | | | | | |
| | 3 ^a : 61 | (4, 4) | (2, 2)* | | | | | | |
| 5 | 1 ^a : 36 | (3, 3) | (3, 3) | 1 ^a : 2* | - | RMSprop | 13 | | |
| | 2 ^a : 276 | (4, 4) | (2, 2)* | | | | | 2 ^a : 75 | 0.20 |
| | 3 ^a : 77 | (3, 3) | (2, 2)* | | | | | | |
| 6 | 1 ^a : 504 | (3, 3) | (3, 3) | 1 ^a : 2* | - | RMSprop | 6 | | |
| | 2 ^a : 511 | (5, 5) | (2, 2)* | | | | | 2 ^a : 157 | 0.49 |
| | 3 ^a : 405 | (5, 5) | (2, 2)* | | | | | 3 ^a : 372 | 0.64 |

Tabela 4.3: Valores dos hiperparâmetros escolhidos nas seis arquiteturas para a classificação dos nódulos pulmonares com o parênquima na proporção $D/4$.

A seguir, na tabela 4.4 estão os valores dos hiperparâmetros (Seção 3.4) escolhidos pelo TPE nas seis arquiteturas (Seção 3.3) exploradas para a classificação dos nódulos com o parênquima na proporção $D/8$.

| Arqui. | Camadas | | | | | Oti. | BS | | |
|----------|----------------------|--------|---------|---------------------|------|---------|----|----------------------|------|
| | Convolução | | MP | TC | | | | | |
| | Filtro | kernel | kernel | Unidades | DROP | | | | |
| 1 | 1 ^a : 311 | (3, 3) | (2, 2) | 1 ^a : 2* | - | Adagrad | 1 | | |
| 2 | 1 ^a : 306 | (5, 5) | (2, 2) | 1 ^a : 2* | - | Adam | 30 | | |
| | 2 ^a : 408 | (3, 3) | (2, 2) | | | | | | |
| 3 | 1 ^a : 303 | (5, 5) | (3, 3) | 1 ^a : 2* | - | Nadam | 18 | | |
| | 2 ^a : 188 | (3, 3) | (2, 2) | | | | | 2 ^a : 370 | 0.69 |
| 4 | 1 ^a : 316 | (5, 5) | (2, 2)* | 1: 2* | - | Adam | 7 | | |
| | 2 ^a : 441 | (5, 5) | (2, 2)* | | | | | | |
| | 3 ^a : 305 | (5, 5) | (2, 2)* | | | | | | |
| 5 | 1 ^a : 113 | (5, 5) | (2, 2)* | 1: 2* | - | Adam | 25 | | |
| | 2 ^a : 475 | (3, 3) | (2, 2)* | | | | | 2 ^a : 512 | 0.48 |
| | 3 ^a : 511 | (5, 5) | (2, 2)* | | | | | | |
| 6 | 1 ^a : 433 | (5, 5) | (2, 2)* | 1 ^a : 2* | - | RMSprop | 50 | | |
| | 2 ^a : 441 | (4, 4) | (2, 2)* | | | | | 2 ^a : 174 | 0.24 |
| | 3 ^a : 230 | (4, 4) | (2, 2)* | | | | | 3 ^a : 294 | 0.23 |

Tabela 4.4: Valores dos hiperparâmetros escolhidos nas seis arquiteturas para a classificação dos nódulos pulmonares com o parênquima na proporção $D/8$.

A partir de agora serão mostrados os valores dos hiperparâmetros escolhidos na avaliação do parênquima segmentado nas proporções $D/2$, $D/4$ e $D/8$.

A seguir na tabela 4.5 estão os valores dos hiperparâmetros (Seção 3.4) escolhidos pelo TPE nas seis arquiteturas (Seção 3.3) exploradas para a classificação dos parênquimas na proporção $D/2$.

| Arqui. | Camadas | | | | | OTI. | BS |
|----------|----------------------|---------------|---------------|----------------------|-------------|------|----|
| | Convolução | | MP | TC | | | |
| | Filtro | <i>kernel</i> | <i>kernel</i> | Unidades | <i>DROP</i> | | |
| 1 | 1 ^a : 204 | (5, 5) | (2, 2) | 1 ^a : 2* | - | SGD | 24 |
| 2 | 1 ^a : 264 | (4, 4) | (2, 2) | 1 ^a : 2* | - | SGD | 20 |
| | 2 ^a : 408 | (3, 3) | (2, 2) | | | | |
| 3 | 1 ^a : 47 | (4, 4) | (2, 2) | 1 ^a : 2* | - | SGD | 34 |
| | 2 ^a : 331 | (3, 3) | (2, 2) | 2 ^a : 84 | 0.06 | | |
| 4 | 1 ^a : 271 | (3, 3) | (2, 2) | 1 ^a : 2* | - | SGD | 40 |
| | 2 ^a : 213 | (3, 3) | (2, 2) | | | | |
| | 3 ^a : 167 | (5, 5) | (2, 2)* | | | | |
| 5 | 1 ^a : 102 | (3, 3) | (2, 2) | 1 ^a : 2* | - | SGD | 45 |
| | 2 ^a : 45 | (3, 3) | (2, 2) | 2 ^a : 162 | 0.65 | | |
| | 3 ^a : 140 | (5, 5) | (2, 2)* | | | | |
| 6 | 1 ^a : 460 | (5, 5) | (2, 2) | 1 ^a : 2* | - | SGD | 26 |
| | 2 ^a : 152 | (5, 5) | (2, 2) | 2 ^a : 171 | 0.05 | | |
| | 3 ^a : 238 | (5, 5) | (2, 2)* | 3 ^a : 256 | 0.42 | | |

Tabela 4.5: Valores dos hiperparâmetros escolhidos nas seis arquiteturas para a classificação do parênquima segmentado na proporção $D/2$.

A seguir na tabela 4.6 estão os valores dos hiperparâmetros (Seção 3.4) escolhidos pelo TPE nas seis arquiteturas (Seção 3.3) utilizadas para a classificação do parênquima segmentado na proporção $D/4$:

| Arqui. | Camadas | | | | | OTI. | BS |
|----------|----------------------|---------------|---------------|---------------------|-------------|---------|----|
| | Convolução | | MP | TC | | | |
| | Filtro | <i>kernel</i> | <i>kernel</i> | Unidades | <i>DROP</i> | | |
| 1 | 1 ^a : 330 | (4, 4) | (3, 3) | 1 ^a : 2* | - | SGD | 44 |
| 2 | 1 ^a : 184 | (3, 3) | (3, 3) | 1 ^a : 2* | - | Adagrad | 25 |

| | | | | | | | |
|----------|----------------------|--------|---------|---------------------|------|--------|----|
| | 2 ^a : 313 | (4, 4) | (2, 2) | | | | |
| 3 | 1 ^a : 182 | (5, 5) | (3, 3) | 1 ^a : 2* | - | SGD | 10 |
| | 2 ^a : 509 | (3, 3) | (3, 3) | 2 ^a : 24 | 0.76 | | |
| 4 | 1 ^a : 56 | (5, 5) | (3, 3) | 1 ^a : 2* | - | Adamax | 11 |
| | 2 ^a : 272 | (5, 5) | (2, 2)* | | | | |
| | 3 ^a : 190 | (5, 5) | (2, 2)* | | | | |
| 5 | 1 ^a : 457 | (4, 4) | (2, 2) | 1 ^a : 2* | - | SGD | 28 |
| | 2 ^a : 34 | (3, 3) | (2, 2)* | | | | |
| | 3 ^a : 287 | (4, 4) | (2, 2)* | | | | |
| 6 | 1 ^a : 170 | (5, 5) | (3, 3) | 1 ^a : 2* | - | SGD | 34 |
| | 2 ^a : 180 | (4, 4) | (2, 2)* | | | | |
| | 3 ^a : 444 | (4, 4) | (2, 2)* | | | | |

Tabela 4.6: Valores dos hiperparâmetros selecionados nas seis arquiteturas para a classificação do parênquima segmentado na proporção $D/4$.

A seguir na tabela 4.7 estão os valores dos hiperparâmetros (Seção 3.4) escolhidos pelo TPE nas seis arquiteturas (Seção 3.3) utilizadas para a classificação do parênquima segmentado na proporção $D/8$.

| Arqui. | Camadas | | | | | OTI. | BS |
|----------|-----------------------|--------|---------|----------------------|------|----------|----|
| | Convolução | | MP | TC | | | |
| | Filtro | kernel | kernel | Unidades | DROP | | |
| 1 | 1 ^a : 410 | (4, 4) | (2, 2) | 1 ^a : 2* | - | Adadelta | 24 |
| 2 | 1 ^a : 32 | (5, 5) | (3, 3) | 1 ^a : 2* | - | Adadelta | 4 |
| | 2 ^a : 260* | (4, 4) | (2, 2) | | | | |
| 3 | 1 ^a : 491 | (4, 4) | (2, 2) | 1 ^a : 2* | - | SGD | 2 |
| | 2 ^a : 491 | (4, 4) | (3, 3) | | | | |
| 4 | 1 ^a : 71 | (4, 4) | (2, 2)* | 1 ^a : 2* | - | Nadam | 11 |
| | 2 ^a : 460 | (5, 5) | (2, 2)* | | | | |
| | 3 ^a : 492 | (3, 3) | (2, 2)* | | | | |
| 5 | 1 ^a : 357 | (3, 3) | (2, 2)* | 1 ^a : 2* | - | Adamax | 21 |
| | 2 ^a : 312 | (4, 4) | (2, 2)* | | | | |
| | 3 ^a : 193 | (3, 3) | (2, 2)* | | | | |
| 6 | 1 ^a : 356 | (4, 4) | (2, 2)* | 1 ^a : 2* | - | Adamax | 23 |
| | 2 ^a : 202 | (4, 4) | (2, 2)* | | | | |
| | 3 ^a : 356 | (3, 3) | (2, 2)* | | | | |
| | | | | 3 ^a : 151 | 0.51 | | |

Tabela 4.7: Valores dos hiperparâmetros selecionados nas seis arquiteturas para a classificação do parênquima segmentado na proporção $D/8$.

A seguir serão apresentados os resultados obtidos utilizando as seis arquiteturas apresentadas nos nódulos por cortes (balanceados e desbalanceados) e por volume (utilizando as cinco abordagens).

4.1 Resultados por Corte

A seguir na tabela 4.8, estão os resultados obtidos pelas seis arquiteturas (Seção 3.3) analisando os cortes dos nódulos segmentados de forma individual tanto balanceados como desbalanceados.

| Cortes Balanceados | | | | | | |
|--------------------|--------------|--------------|--------------|--------------|--------------|-------------|
| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
| 1 | 20.92 | 72.82 | 85.37 | 77.72 | 79.08 | 0.81 |
| 2 | 20.44 | 74.27 | 84.88 | 78.46 | 79.56 | 0.81 |
| 3 | 20.68 | 78.64 | 80.00 | 79.22 | 79.32 | 0.84 |
| 4 | 23.11 | 89.32 | 64.39 | 79.48 | 76.89 | 0.79 |
| 5 | 21.41 | 70.39 | 86.83 | 76.72 | 78.59 | 0.82 |
| 6 | 21.41 | 70.39 | 86.83 | 76.72 | 78.56 | 0.79 |

| Cortes desbalanceados | | | | | | |
|-----------------------|--------------|--------------|--------------|--------------|--------------|-------------|
| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
| 1 | 20.98 | 76.37 | 85.37 | 83.11 | 79.02 | 0.80 |
| 2 | 20.98 | 76.58 | 84.88 | 83.74 | 79.02 | 0.81 |
| 3 | 19.68 | 80.45 | 80.00 | 85.22 | 80.32 | 0.83 |
| 4 | 18.25 | 89.00 | 64.39 | 87.31 | 81.75 | 0.79 |
| 5 | 23.85 | 71.69 | 86.83 | 80.92 | 76.15 | 0.82 |
| 6 | 23.28 | 72.51 | 86.83 | 81.46 | 76.72 | 0.79 |

Tabela 4.8: Resultados obtidos analisando os nódulos segmentados pelos cortes de formas balanceadas e desbalanceadas.

Ao analisarmos os resultados dos nódulos segmentados, a maior AUC obtida melhor foi usando a arquitetura três nos cortes balanceados, que atingiu AUC de 0.84, e ainda uma taxa de erro de 20%, sensibilidade de 78%, especificidade de 80%, *f-measure* de 79% e acurácia de 79%. Apesar dos resultados terem sido próximos, a menor AUC foi de 0.79, ou seja, a arquitetura três ainda teve um ganho de 0.05 na AUC. Este cenário mostrou que, uma arquitetura com duas camadas convolucionais e duas camadas totalmente conectadas contribuiu um pouco mais do que as outras arquiteturas para obter o melhor valor da AUC de 0.84, mesma arquitetura que conseguiu a melhor AUC no caso dos cortes desbalanceados, de 0.83.

A maior taxa de sensibilidade foi conseguida pela arquitetura 4, com 89%, tanto nos cortes

balanceados como nos desbalanceados; a maior taxa de especificidade foi conseguida pelas arquiteturas 5 e 6, com 86% nas duas formas dos cortes, balanceados e desbalanceados; a maior taxa de *f-measure* foi alcançada pela arquitetura 4 nos cortes desbalanceados com 87% e nos cortes balanceados com 79%; e a maior acurácia foi conseguida pela arquitetura 4 com 81% nos cortes desbalanceados, consequentemente a menor taxa de erro com 18%, e nos cortes balanceados a maior acurácia foi conseguida pela arquitetura 2 com 79%.

A partir de agora serão mostrados os resultados obtidos na análise dos nódulos junto com o parênquima por corte na mesma ROI nas proporções $D/2$, $D/4$ e $D/8$.

A seguir na tabela 4.9, estão os resultados obtidos pelas seis arquiteturas (Seção 3.3) analisando os cortes dos nódulos junto com o parênquima na proporção $D/2$ tanto balanceados como desbalanceados.

Cortes balanceados

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
|--------------|--------------|--------------|--------------|--------------|--------------|-------------|
| 1 | 24.57 | 69.90 | 80.98 | 74.04 | 75.43 | 0.81 |
| 2 | 21.17 | 77.18 | 80.49 | 78.52 | 78.83 | 0.80 |
| 3 | 20.68 | 76.70 | 81.95 | 78.80 | 79.32 | 0.82 |
| 4 | 30.66 | 62.14 | 76.59 | 67.02 | 69.34 | 0.74 |
| 5 | 19.95 | 81.07 | 79.02 | 80.29 | 80.05 | 0.83 |
| 6 | 20.68 | 81.07 | 77.56 | 79.71 | 79.32 | 0.82 |

Cortes desbalanceados

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
|--------------|--------------|--------------|--------------|--------------|--------------|-------------|
| 1 | 23.28 | 74.95 | 80.98 | 81.26 | 76.72 | 0.80 |
| 2 | 24.14 | 73.93 | 80.49 | 81.21 | 75.86 | 0.81 |
| 3 | 18.25 | 81.67 | 81.95 | 86.33 | 81.75 | 0.79 |
| 4 | 30.60 | 66.40 | 76.59 | 75.38 | 69.40 | 0.72 |
| 5 | 17.96 | 83.20 | 79.02 | 86.74 | 82.04 | 0.82 |
| 6 | 18.68 | 82.89 | 77.56 | 86.23 | 81.32 | 0.82 |

Tabela 4.9: Resultados obtidos analisando os nódulos e parênquima (caso $D/2$) de formas balanceadas e desbalanceadas.

Ao analisarmos os resultados dos nódulos e parênquima (caso $D/2$), a maior AUC foi obtida utilizando a arquitetura 5 (cortes balanceados), que obteve a maior AUC de 0.83. Neste arquitetura ainda tivemos taxa de erro de 19%, sensibilidade de 81%, especificidade de 79% e acurácia de 80%. Neste cenário vemos que, uma arquitetura com mais uma camada de convolução, em comparação a arquitetura 3 no cenário da classificação dos nódulos segmentados por cortes, classificou melhor, visto que nas imagens tivemos o parênquima acrescentado, a RNC precisou extrair mais características abstratas para aprender melhor a diferenciar os nódulos e de duas camadas totalmente conectada. As arquiteturas 5 e 6 tanto nos cortes balanceados como desba-

lançados atingiram AUC praticamente idênticas, mas a arquitetura 5 nos cortes balanceados, por 0.01, atingiu a AUC levemente maior.

A maior taxa de sensibilidade foi conseguida pela arquitetura 5, com 83%, nos cortes desbalanceados, já nos cortes desbalanceados a maior taxa de sensibilidade foi de 81% nas arquiteturas 5 e 6; a maior taxa de especificidade foi conseguida pela arquitetura 3, com 81% nas duas formas dos cortes, balanceados e desbalanceados; a maior taxa de *f-measure* foi alcançada pela arquitetura 5 nos cortes desbalanceados com 86% e nos cortes balanceados com 80%; e a maior acurácia foi conseguida pela arquitetura 5 com 86% nos cortes desbalanceados, consequentemente a menor taxa de erro com 17%, e nos cortes balanceados a maior acurácia foi de 80%.

A seguir na tabela 4.10, estão os resultados obtidos pelas seis arquiteturas (Seção 3.3) analisando os cortes dos nódulos junto com o parênquima na proporção $D/4$ tanto balanceados como desbalanceados.

Cortes balanceados

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
|--------------|--------------|--------------|--------------|--------------|--------------|-------------|
| 1 | 21.65 | 70.87 | 85.85 | 76.64 | 78.35 | 0.81 |
| 2 | 20.68 | 77.67 | 80.98 | 79.01 | 79.32 | 0.84 |
| 3 | 20.92 | 76.70 | 81.46 | 78.61 | 79.08 | 0.84 |
| 4 | 20.92 | 76.21 | 81.95 | 78.50 | 79.08 | 0.82 |
| 5 | 21.65 | 77.18 | 79.51 | 78.13 | 78.35 | 0.82 |
| 6 | 21.41 | 81.07 | 76.10 | 79.15 | 79.24 | 0.81 |

Cortes desbalanceados

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
|--------------|--------------|--------------|--------------|--------------|--------------|-------------|
| 1 | 22.99 | 73.32 | 85.85 | 81.82 | 77.01 | 0.79 |
| 2 | 18.25 | 82.08 | 80.98 | 86.39 | 81.78 | 0.82 |
| 3 | 18.68 | 81.26 | 81.32 | 85.99 | 81.32 | 0.82 |
| 4 | 19.25 | 80.24 | 81.95 | 85.47 | 80.75 | 0.80 |
| 5 | 18.53 | 82.28 | 79.51 | 86.23 | 81.47 | 0.80 |
| 6 | 17.39 | 85.34 | 76.10 | 87.38 | 81.13 | 0.79 |

Tabela 4.10: Resultados obtidos analisando os nódulos e parênquima (caso $D/4$) por corte de formas balanceada e desbalanceada.

Ao analisarmos os resultados do nódulo e parênquima (caso $D/4$), a maior AUC foi obtida utilizando a arquitetura 2 nos cortes balanceados, que atingiu AUC de 0.84, e ainda, taxa de erro de 20%, sensibilidade de 80%, especificidade de 79%, *f-measure* de 79% e acurácia de 79%. Neste cenário de nódulo e parênquima em uma proporção menor do a anterior, $D/4$ em vez de $D/2$, uma arquitetura menor contribuiu mais. Uma arquitetura com duas camadas convolucionais e apenas uma camada totalmente conectada conseguiu a melhor AUC nos cortes

balanceadas e nos cortes desbalanceados obteve AUC de 0.82 que também foi obtida por outras três arquiteturas. O fato de a arquitetura 2 ter sido melhor neste cenário pode ser justificado pelo fato de que, com uma proporção menor, uma arquitetura com menos camadas de convolução e totalmente conectada do que na proporção $D/2$ já conseguiu extrair os atributos mais relevantes para realizar a classificação com mais sucesso.

A maior taxa de sensibilidade foi conseguida pela arquitetura 5, com 85% nos cortes desbalanceados, e com 81% nos cortes balanceados; a maior taxa de especificidade foi conseguida pela arquitetura 1, com 85% nas duas formas dos cortes, balanceados e desbalanceados; a maior taxa de f -measure foi alcançada pela arquitetura 6 nos cortes desbalanceados com 87% e nos cortes balanceados com 79%; e a maior acurácia foi conseguida pela arquitetura 2 com 81% nos cortes desbalanceados, conseqüentemente a menor taxa de erro com 18%, e nos cortes balanceados a maior acurácia foi conseguida pela arquitetura 6 com 79%.

A seguir na tabela 4.11, estão os resultados obtidos pelas seis arquiteturas (Seção 3.3) analisando os cortes dos nódulos junto com o parênquima na proporção $D/8$ tanto balanceados como desbalanceados.

Cortes balanceados

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
|--------------|--------------|--------------|--------------|--------------|--------------|-------------|
| 1 | 24.33 | 60.08 | 90.73 | 71.43 | 75.67 | 0.79 |
| 2 | 21.17 | 79.61 | 78.05 | 79.04 | 78.83 | 0.82 |
| 3 | 23.36 | 75.73 | 77.56 | 76.47 | 76.64 | 0.83 |
| 4 | 26.52 | 64.56 | 82.44 | 70.93 | 73.48 | 0.78 |
| 5 | 22.38 | 73.30 | 81.95 | 76.65 | 77.62 | 0.83 |
| 6 | 21.90 | 75.73 | 80.49 | 77.61 | 78.10 | 0.83 |

Cortes desbalanceados

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
|--------------|--------------|--------------|--------------|--------------|--------------|-------------|
| 1 | 28.78 | 63.14 | 90.73 | 75.61 | 71.26 | 0.79 |
| 2 | 18.25 | 83.30 | 78.05 | 86.56 | 81.75 | 0.81 |
| 3 | 20.98 | 79.63 | 77.56 | 84.27 | 79.02 | 0.82 |
| 4 | 29.89 | 64.97 | 82.44 | 75.41 | 70.11 | 0.77 |
| 5 | 21.12 | 77.60 | 81.95 | 83.83 | 78.88 | 0.82 |
| 6 | 22.27 | 76.58 | 80.49 | 82.91 | 77.73 | 0.84 |

Tabela 4.11: Resultados obtidos analisando os nódulos e parênquima (caso $D/8$) por corte de formas balanceada e desbalanceada.

Ao analisarmos os resultados do nódulo e parênquima (caso $D/8$), a maior AUC foi utilizando a arquitetura quatro nos cortes balanceados, que atingiu AUC de 0.86, e ainda, taxa de erro de 20%, sensibilidade de 76%, especificidade de 81%, f -measure de 78% e acurácia de 79%. Esta mesma arquitetura também conseguiu a maior AUC nos cortes desbalanceadas

de 0.85. Neste cenário vemos que, a mesma arquitetura atingiu a maior AUC nos dois casos dos cortes (balanceados e desbalanceados), porém o caso balanceado atingiu uma AUC de 0.01 maior. Diferente da proporção $D/4$, que conseguiu a melhor AUC com uma arquitetura menor, a arquitetura 2, aqui foi preciso de mais uma camada de convolução (3 em vez 2) e também de uma camada totalmente camada. Ou seja, apesar da proporção ter sido menor aqui, a arquitetura precisou de mais uma camada de convolução para atingir a maior AUC. Logo, nem sempre uma proporção menor de parênquima influenciará uma arquitetura vim a ser menor.

A maior taxa de sensibilidade foi conseguida pela arquitetura 2, com 83% nos cortes desbalanceados, e com 79% nos cortes balanceados; a maior taxa de especificidade foi conseguida pela arquitetura 1, com 90% nas duas formas dos cortes, balanceados e desbalanceados; a maior taxa de f -measure foi alcançada pela arquitetura 2 nos cortes desbalanceados com 86% e nos cortes balanceados com 79%; e a maior acurácia foi conseguida pela arquitetura 2 com 81% nos cortes desbalanceados, consequentemente a menor taxa de erro com 18%, e nos cortes balanceados com 78%.

A maior AUC atingida na análise dos nódulos com o parênquima foi de 0.84, mesmo valor atingido nos nódulos segmentados, porém a arquitetura 6 aqui ainda atingiu uma f -measure de 82%, enquanto que a arquitetura 3 nos nódulos segmentados atingiu uma f -measure de 79%, ou seja, analisar nódulos com parênquima utilizando uma RNC contribui mais para a classificação dos nódulos pulmonares por corte do que analisar apenas os nódulos segmentados.

A partir de agora serão mostrados os resultados obtidos na análise do parênquima segmentado por corte na mesma ROI nas proporções $D/2$, $D/4$ e $D/8$.

A seguir na tabela 4.12, estão os resultados obtidos pelas seis arquiteturas (Seção 3.3) analisando os cortes dos nódulos junto com o parênquima na proporção $D/2$ tanto balanceados como desbalanceados.

Cortes balanceados

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
|--------------|--------------|--------------|--------------|--------------|--------------|-------------|
| 1 | 21.41 | 89.81 | 67.32 | 80.79 | 78.59 | 0.81 |
| 2 | 21.65 | 87.86 | 68.78 | 80.27 | 78.35 | 0.82 |
| 3 | 21.17 | 92.72 | 64.88 | 81.45 | 78.83 | 0.81 |
| 4 | 21.41 | 93.20 | 63.90 | 81.36 | 78.59 | 0.80 |
| 5 | 21.41 | 92.23 | 92.23 | 64.88 | 81.20 | 0.82 |
| 6 | 21.17 | 86.89 | 70.73 | 80.45 | 78.83 | 0.84 |

Cortes desbalanceados

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
|--------------|--------------|--------------|--------|--------------|--------------|------|
| 1 | 15.09 | 92.26 | 67.32 | 89.61 | 81.27 | 0.81 |
| 2 | 15.80 | 90.63 | 68.78 | 89.00 | 84.20 | 0.81 |
| 3 | 14.37 | 94.30 | 64.88 | 90.25 | 85.63 | 0.81 |
| 4 | 14.22 | 94.97 | 63.90 | 90.40 | 85.78 | 0.80 |

| | | | | | | |
|----------|-------|-------|--------------|-------|-------|-------------|
| 5 | 14.66 | 93.89 | 64.88 | 90.04 | 85.34 | 0.81 |
| 6 | 15.95 | 89.61 | 70.73 | 88.80 | 84.05 | 0.83 |

Tabela 4.12: Resultados obtidos analisando o parênquima segmentado (caso $D/2$) por corte de formas balanceada e desbalanceada.

Ao analisarmos os resultados do parênquima segmentado (caso $D/2$), a maior AUC foi utilizando a arquitetura 6 nos cortes balanceados que, atingiu uma AUC de 0.84, e esta mesma arquitetura atingiu a maior AUC no caso dos cortes desbalanceados, AUC de 0.83. Pela primeira vez a maior arquitetura foi a que atingiu a maior AUC, porém num cenário diferente, do parênquima segmentado. Isso mostra que, com apenas o parênquima segmentado nas imagens, há a necessidade da utilização de uma arquitetura maior, em comparação aos casos dos nódulos segmentados e dos nódulos com o parênquima, ou seja, a RNC precisou extrair mais informações e depois precisou selecionar melhor os atributos mais relevantes nas camadas totalmente conectadas. Esses resultados também sugerem que uma arquitetura maior do que a arquitetura 6 poderia conseguir uma AUC ainda maior.

A maior taxa de sensibilidade foi conseguida pela arquitetura 4 com 94% nos cortes desbalanceados, e com 93% nos cortes balanceados; a maior taxa de especificidade foi conseguida pela arquitetura 5 com 92% nos cortes balanceados, já nos cortes desbalanceados a maior taxa de especificidade foi conseguida pela arquitetura 6 com 70%; a maior taxa de f -measure foi alcançada pela arquitetura 4 nos cortes desbalanceados com 90% e pela arquitetura 3 nos cortes balanceados com 81%; e a maior acurácia foi conseguida pela arquitetura 4 com 85% nos cortes desbalanceados, conseqüentemente a menor taxa de erro com 14%, e nos cortes balanceados a maior acurácia foi conseguida pela arquitetura 5 com 81%.

A partir de agora serão mostrados os resultados obtidos dos parênquimas segmentados na proporção $D/4$. A seguir na tabela 4.13, estão os resultados obtidos pelas seis arquiteturas (Seção 3.3) analisando os cortes dos nódulos junto com o parênquima na proporção $D/4$ de forma tanto balanceados como desbalanceados.

Cortes balanceados

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
|--------------|--------------|--------------|--------------|--------------|--------------|-------------|
| 1 | 21.17 | 87.38 | 70.24 | 80.54 | 78.83 | 0.83 |
| 2 | 20.44 | 80.10 | 79.02 | 79.71 | 79.56 | 0.84 |
| 3 | 20.68 | 90.78 | 67.80 | 81.48 | 79.32 | 0.83 |
| 4 | 20.44 | 87.38 | 71.71 | 81.08 | 81.70 | 0.83 |
| 5 | 20.44 | 86.89 | 72.20 | 81.00 | 79.56 | 0.82 |
| 6 | 21.17 | 87.38 | 70.24 | 80.54 | 78.83 | 0.83 |

Cortes desbalanceados

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
|--------------|--------|--------|--------|-------|--------|-----|
|--------------|--------|--------|--------|-------|--------|-----|

| | | | | | | |
|----------|--------------|--------------|--------------|--------------|--------------|-------------|
| 1 | 15.80 | 90.02 | 70.24 | 88.93 | 84.20 | 0.82 |
| 2 | 17.10 | 84.52 | 79.02 | 87.46 | 82.90 | 0.82 |
| 3 | 14.80 | 92.46 | 67.80 | 89.81 | 85.20 | 0.82 |
| 4 | 15.66 | 89.61 | 71.71 | 88.98 | 84.74 | 0.82 |
| 5 | 15.95 | 89.00 | 72.20 | 88.73 | 84.05 | 0.81 |
| 6 | 15.95 | 89.82 | 70.24 | 88.82 | 84.05 | 0.82 |

Tabela 4.13: Resultados obtidos analisando o parênquima segmentado (caso $D/4$) por corte de formas balanceada e desbalanceada.

Ao analisarmos os resultados do parênquima segmentado (caso $D/4$), o melhor resultado foi obtido utilizando a arquitetura 2 nos cortes balanceados, que atingiu a maior AUC de 0.84. Já no caso dos cortes desbalanceados, a arquitetura 2 ficou entre as arquiteturas que atingiram AUC de 0.82, porém, a arquitetura 3 atingiu a maior f -measure de 89%, sendo a melhor arquitetura no caso dos cortes desbalanceados aqui. Diferente do caso ($D/2$) que atingiu a melhor AUC com a maior arquitetura (arquitetura 6), aqui a maior AUC foi conseguida pela arquitetura 2, com apenas duas camadas de convolução e uma camada totalmente conectada. Ou seja, em uma proporção menor, uma RNC com uma arquitetura menor já conseguiu extrair os atributos mais importantes para o seu treinamento.

A maior taxa de sensibilidade foi conseguida pela arquitetura 3, com 92% nos cortes desbalanceados, e com 90% nos cortes balanceados; a maior taxa de especificidade foi conseguida pela arquitetura 2, com 79% nas duas formas dos cortes, balanceados e desbalanceados; a maior taxa de f -measure foi alcançada pela arquitetura 3 nos cortes desbalanceados com 89% e nos cortes balanceados com 81%; e a maior acurácia foi conseguida pela arquitetura 3 com 89% nos cortes desbalanceados, conseqüentemente a menor taxa de erro com 14%, e nos cortes balanceados a maior acurácia foi conseguida pela arquitetura 4 com 81%.

A partir de agora serão mostrados os resultados obtidos dos parênquimas segmentados na proporção $D/8$. A seguir na tabela 4.14, estão os resultados obtidos pelas seis arquiteturas (Seção 3.3) analisando os cortes dos nódulos junto com o parênquima na proporção $D/8$ de forma individuais tanto balanceados como desbalanceados (3.5).

Cortes balanceados

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
|--------------|--------------|--------------|--------------|--------------|--------------|-------------|
| 1 | 20.19 | 89.32 | 70.24 | 81.60 | 79.81 | 0.82 |
| 2 | 19.46 | 75.24 | 85.85 | 79.49 | 80.54 | 0.84 |
| 3 | 22.38 | 95.15 | 60.00 | 80.99 | 77.62 | 0.81 |
| 4 | 20.68 | 76.70 | 81.95 | 78.80 | 79.32 | 0.86 |
| 5 | 20.68 | 73.79 | 84.88 | 78.15 | 79.32 | 0.84 |
| 6 | 20.44 | 72.82 | 86.34 | 78.12 | 79.56 | 0.85 |

| Cortes desbalanceados | | | | | | |
|-----------------------|--------------|--------------|--------------|--------------|--------------|-------------|
| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
| 1 | 15.23 | 90.84 | 70.24 | 89.38 | 84.77 | 0.81 |
| 2 | 21.12 | 75.97 | 85.85 | 83.54 | 78.88 | 0.83 |
| 3 | 14.80 | 95.72 | 60.00 | 90.12 | 85.20 | 0.80 |
| 4 | 19.11 | 80.45 | 81.95 | 85.59 | 80.89 | 0.85 |
| 5 | 20.83 | 76.78 | 84.88 | 83.87 | 79.17 | 0.83 |
| 6 | 20.26 | 76.99 | 86.34 | 84.28 | 79.74 | 0.83 |

Tabela 4.14: Resultados obtidos analisando o parênquima segmentado (caso $D/8$) por corte de formas balanceada e desbalanceada.

Ao analisarmos os resultados do parênquima segmentado (caso $D/8$), a maior AUC foi obtida utilizando a arquitetura 4 no caso dos cortes balanceados, que atingiu a AUC de 0.86. Esta arquitetura também atingiu a maior AUC no caso dos cortes desbalanceados, porém, de 0.83. Neste cenário foi atingido a maior AUC analisando os nódulos por corte, mostrando que, o parênquima segmentado na menor proporção contribuiu para uma classificação melhor em comparação as análises dos nódulos segmentados e dos nódulos com parênquima. Enquanto as análises de nódulos segmentados e nódulos com parênquima atingiram a maior AUC de 0.84, aqui a AUC foi maior por 0.02.

A maior taxa de sensibilidade foi conseguida pela arquitetura 3 com 95% nos cortes balanceados e desbalanceados; a maior taxa de especificidade foi conseguida pela arquitetura 6 com 86% nas duas formas dos cortes, balanceados e desbalanceados; a maior taxa de f -measure foi alcançada pela arquitetura 3 nos cortes desbalanceados com 90% e pela arquitetura 1 nos cortes balanceados com 81%; e a maior acurácia foi conseguida pela arquitetura 3 com 85% nos cortes desbalanceados, conseqüentemente a menor taxa de erro com 14%, e nos cortes balanceados a maior acurácia foi conseguida pela arquitetura 2 com 80%.

Tanto no cenário dos nódulos com parênquima, como no cenário do parênquima segmentado, a menor proporção de parênquima ($D/8$) melhor a resultados menores, mostrando que, na classificação de nódulos pulmonares por corte abordando a região do parênquima, a menor proporção indicada neste trabalho deve ser a considerada.

Logo, ao classificar os nódulos pulmonares por corte utilizando uma RNC, utilizar como base o parênquima segmentado na proporção $D/8$ contribui mais para os resultados da AUC do que utilizar os nódulos com o parênquima na proporção $D/8$.

A seguir na tabela 4.15, está resumido os melhores resultados alcançados na avaliação dos nódulos por corte em cada métrica, a arquitetura e a abordagem (cortes balanceados ou desbalanceados) utilizadas:

| Métrica | Resultado | Arquitetura - abordagem |
|---------|-----------|-------------------------|
|---------|-----------|-------------------------|

| | | |
|----------------|-------|---|
| Err.(%) | 14.22 | 4 - parênquima segmentado (caso $D/2$), cortes desbalanceados. |
| SEN(%) | 95.72 | 3 - parênquima segmentado (caso $D/8$), cortes desbalanceados. |
| ESP(%) | 92.33 | 5 - parênquima segmentado (caso $D/2$), cortes desbalanceados. |
| FM(%) | 90.40 | 4 - parênquima segmentado (caso $D/2$), cortes desbalanceados. |
| ACC(%) | 85.78 | 4 - parênquima segmentado (caso $D/2$), cortes desbalanceados. |
| AUC | 0.86 | 4 - parênquima segmentado (caso $D/8$), cortes balanceados. |

Tabela 4.15: Melhores resultados na avaliação por corte.

Pela tabela 4.15 conclui-se que, os melhores resultados alcançados foram abordando o parênquima segmentado, e nas proporções $D/2$ e $D/8$. Ou seja, na classificação dos cortes de nódulos pulmonares sólidos é melhor utilizar uma base com o parênquima segmentado dos respectivos nódulos do que uma base com os nódulos segmentados ou dos nódulos com o parênquima. Por fim, a melhor configuração foi arquitetura 5 com a base do parênquima segmentado (caso $D/8$) nos cortes balanceados, visto que a AUC foi a maior, e AUC é a métrica mais importante dentre as utilizadas neste trabalho.

4.2 Resultados por Volume

A seguir na tabela 4.16, estão os resultados obtidos pelas seis arquiteturas (Seção 3.3) analisando os nódulos segmentados por volume nas abordagens 1, 2, 3, 4, e 5 (Seção 3.5).

Abordagem 1

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | Nº de Empates |
|--------------|--------------|--------------|--------------|--------------|--------------|---------------|
| 1 | 19.57 | 69.57 | 91.30 | 78.05 | 80.43 | 1 |
| 2 | 20.65 | 67.39 | 91.30 | 76.54 | 79.35 | 1 |
| 3 | 20.65 | 69.57 | 89.13 | 77.11 | 79.35 | 1 |
| 4 | 21.74 | 65.22 | 91.30 | 75.00 | 78.26 | 2 |
| 5 | 19.57 | 67.39 | 93.48 | 77.50 | 80.43 | 1 |
| 6 | 25.00 | 54.35 | 95.65 | 68.49 | 75.00 | 1 |

Abordagem 2

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | Nº de Empates |
|--------------|--------------|--------------|--------------|--------------|--------------|---------------|
| 1 | 20.65 | 69.57 | 89.13 | 77.11 | 79.35 | 1 |
| 2 | 21.74 | 67.39 | 89.13 | 75.61 | 78.26 | 1 |
| 3 | 19.57 | 71.74 | 89.13 | 78.57 | 80.43 | 1 |
| 4 | 21.74 | 67.39 | 89.13 | 75.61 | 78.26 | 2 |
| 5 | 20.65 | 67.39 | 91.30 | 76.54 | 79.35 | 1 |
| 6 | 21.74 | 60.87 | 95.65 | 73.68 | 78.26 | 3 |

| Abordagem 3 | | | | | | |
|--------------|--------------|--------------|--------------|--------------|--------------|-------------|
| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
| 1 | 20.65 | 63.04 | 95.65 | 75.32 | 79.35 | 0.91 |
| 2 | 21.74 | 63.04 | 93.48 | 74.36 | 78.26 | 0.91 |
| 3 | 22.83 | 60.87 | 93.48 | 72.73 | 77.17 | 0.90 |
| 4 | 21.74 | 60.87 | 95.65 | 73.68 | 78.26 | 0.90 |
| 5 | 22.83 | 56.52 | 97.83 | 71.23 | 77.17 | 0.92 |
| 6 | 26.09 | 50.00 | 97.83 | 65.71 | 73.91 | 0.90 |

| Abordagem 4 | | | | | | |
|--------------|--------------|--------------|--------------|--------------|--------------|-------------|
| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
| 1 | 21.74 | 65.22 | 91.30 | 75.00 | 78.26 | 0.91 |
| 2 | 19.57 | 67.39 | 93.48 | 77.50 | 80.43 | 0.89 |
| 3 | 20.65 | 67.39 | 91.30 | 76.54 | 79.35 | 0.89 |
| 4 | 20.65 | 65.22 | 93.48 | 75.95 | 79.35 | 0.89 |
| 5 | 18.48 | 67.39 | 95.65 | 78.48 | 81.52 | 0.92 |
| 6 | 22.83 | 56.52 | 97.83 | 71.23 | 77.17 | 0.89 |

| Abordagem 5 | | | | | | |
|--------------|--------------|--------------|--------------|--------------|--------------|---|
| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | - |
| 1 | 17.39 | 82.61 | 82.61 | 82.61 | 82.61 | - |
| 2 | 17.39 | 82.61 | 82.61 | 82.61 | 82.61 | - |
| 3 | 16.30 | 82.61 | 84.78 | 83.52 | 83.70 | - |
| 4 | 21.74 | 80.43 | 76.09 | 78.72 | 78.26 | - |
| 5 | 17.39 | 82.61 | 82.61 | 82.61 | 82.61 | - |
| 6 | 17.39 | 78.26 | 86.96 | 81.82 | 82.61 | - |

Tabela 4.16: Resultados obtidos analisando os nódulos segmentados por volume nas abordagens 1, 2, 3, 4 e 5.

Ao analisarmos os resultados dos nódulos segmentados por volume, a AUC de maior valor foi obtida pela arquitetura 5 com 0.92 nas abordagens três e quatro; a maior taxa de sensibilidade foi conseguida pelas arquiteturas 1, 2 e 3 com 82% na abordagem cinco; a maior taxa de especificidade foi conseguida pela arquitetura 6 com 97% na abordagem quatro, e pelas arquiteturas 5 e 6 com 97% na abordagem três; a maior taxa de *f-measure* foi alcançada pela arquitetura 3 na abordagem cinco com 83%; e a maior acurácia foi conseguida pela arquitetura 3 na abordagem cinco com 83%, e conseqüentemente a menor taxa de erro com 16%. Comparando os resultados aqui com a avaliação dos nódulos segmentados por corte, a AUC foi maior por 0.08, 0.92 contra 0.84.

A partir de agora serão mostrados os resultados obtidos na análise dos nódulos junto com o

parênquima por volume nas proporções $D/2$, $D/4$ e $D/8$.

A seguir na tabela 4.17, estão os resultados obtidos pelas seis arquiteturas (Seção 3.3) analisando os nódulos juntos com o parênquima na proporção $D/2$ por volume nas abordagens 1, 2, 3, 4, e 5 (Seção 3.5).

Abordagem 1

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | Nº de Empates |
|--------------|--------------|--------------|--------------|--------------|--------------|---------------|
| 1 | 21.74 | 71.74 | 84.78 | 76.74 | 78.26 | 2 |
| 2 | 19.57 | 80.43 | 80.43 | 80.43 | 80.43 | 2 |
| 3 | 20.65 | 71.74 | 86.96 | 77.65 | 79.35 | 0 |
| 4 | 21.74 | 76.09 | 80.43 | 77.78 | 78.26 | 2 |
| 5 | 16.30 | 80.43 | 86.96 | 83.15 | 83.70 | 1 |
| 6 | 19.57 | 78.26 | 82.61 | 80.00 | 80.43 | 0 |

Abordagem 2

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | Nº de Empates |
|--------------|--------------|--------------|--------------|--------------|--------------|---------------|
| 1 | 19.57 | 76.09 | 84.78 | 79.55 | 80.43 | 2 |
| 2 | 19.57 | 82.61 | 78.26 | 80.85 | 80.43 | 2 |
| 3 | 20.65 | 71.74 | 86.96 | 77.65 | 79.35 | 0 |
| 4 | 19.57 | 80.43 | 80.43 | 80.43 | 80.43 | 2 |
| 5 | 17.39 | 80.43 | 84.78 | 82.22 | 82.61 | 1 |
| 6 | 19.57 | 78.26 | 82.61 | 80.00 | 80.43 | 0 |

Abordagem 3

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
|--------------|--------------|--------------|--------------|--------------|--------------|-------------|
| 1 | 22.83 | 65.22 | 89.13 | 74.07 | 77.17 | 0.88 |
| 2 | 19.57 | 71.74 | 89.13 | 78.57 | 80.43 | 0.88 |
| 3 | 20.65 | 69.57 | 89.13 | 77.11 | 79.35 | 0.88 |
| 4 | 20.65 | 73.91 | 84.78 | 78.16 | 79.35 | 0.86 |
| 5 | 19.57 | 71.74 | 89.13 | 78.57 | 80.43 | 0.88 |
| 6 | 20.65 | 69.57 | 79.35 | 77.11 | 79.35 | 0.89 |

Abordagem 4

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
|--------------|--------------|--------------|--------------|--------------|--------------|-------------|
| 1 | 17.39 | 76.09 | 89.13 | 81.40 | 82.61 | 0.88 |
| 2 | 17.39 | 80.43 | 54.78 | 82.22 | 82.61 | 0.88 |
| 3 | 19.57 | 71.74 | 89.13 | 78.57 | 80.43 | 0.88 |
| 4 | 19.57 | 80.43 | 80.43 | 80.43 | 80.43 | 0.87 |
| 5 | 18.48 | 78.26 | 84.78 | 80.90 | 81.52 | 0.87 |
| 6 | 17.39 | 76.09 | 89.13 | 81.40 | 82.61 | 0.88 |

Abordagem 5

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | - |
|--------------|--------------|--------------|--------------|--------------|--------------|---|
| 1 | 18.48 | 82.61 | 80.43 | 81.72 | 81.52 | - |
| 2 | 18.48 | 86.96 | 76.09 | 82.47 | 81.52 | - |
| 3 | 19.57 | 80.43 | 80.43 | 80.43 | 80.43 | - |
| 4 | 18.48 | 89.13 | 73.91 | 82.83 | 81.52 | - |
| 5 | 16.30 | 86.96 | 80.43 | 84.21 | 83.70 | - |
| 6 | 19.57 | 84.78 | 76.09 | 81.25 | 80.43 | - |

Tabela 4.17: Resultados obtidos analisando os nódulos com o parênquima na proporção $D/2$ por volume nas abordagens 1, 2, 3, 4 e 5.

Ao analisarmos os resultados dos nódulos com parênquima por volume (caso $D/2$), a AUC de maior valor foi obtida pela arquitetura 6 com 0.89 na abordagem três; a maior taxa de sensibilidade foi conseguida pela arquitetura 4 com 89% na abordagem cinco; a maior taxa de especificidade foi conseguida pelas arquiteturas 1, 2, 3 e 5 com 89% na abordagem três, e pelas arquiteturas 1, 3 e 6 na abordagem quatro; a maior taxa de f -measure foi alcançada pela arquitetura 5 na abordagem cinco com 84%; e a maior acurácia foi conseguida pela arquitetura 5 nas abordagens 1 e 5 com 83%, e conseqüentemente a menor taxa de erro com 16%. Comparando com os resultados aqui com a avaliação dos nódulos com parênquima (caso $D/2$) por corte, a AUC por volume também foi maior, diferença de 0.06, 0.89 contra 0.83.

A seguir na tabela 4.18, estão os resultados obtidos pelas seis arquiteturas (Seção 3.3) analisando os nódulos juntos com o parênquima na proporção $D/4$ por volume nas abordagens 1, 2, 3, 4, e 5 (Seção 3.5).

Abordagem 1

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | Nº de Empates |
|--------------|--------------|--------------|--------------|--------------|--------------|---------------|
| 1 | 19.57 | 73.91 | 86.96 | 79.07 | 80.43 | 1 |
| 2 | 21.74 | 65.22 | 91.30 | 75.00 | 78.26 | 1 |
| 3 | 20.65 | 71.74 | 86.96 | 77.65 | 79.35 | 0 |
| 4 | 20.65 | 71.74 | 86.96 | 77.65 | 79.35 | 1 |
| 5 | 21.74 | 76.09 | 80.43 | 77.78 | 78.26 | 1 |
| 6 | 20.65 | 78.26 | 80.43 | 79.12 | 79.35 | 0 |

Abordagem 2

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | Nº de Empates |
|--------------|--------------|--------------|--------------|--------------|--------------|---------------|
| 1 | 20.65 | 73.91 | 84.78 | 78.16 | 79.35 | 1 |
| 2 | 20.65 | 67.39 | 91.30 | 76.54 | 79.85 | 1 |
| 3 | 20.65 | 71.74 | 86.96 | 77.65 | 79.35 | 0 |
| 4 | 21.74 | 71.74 | 84.78 | 76.74 | 78.26 | 1 |
| 5 | 20.65 | 78.26 | 80.43 | 79.12 | 79.35 | 1 |

| 6 | 20.65 | 78.26 | 80.43 | 79.12 | 79.35 | 0 |
|--------------------|--------------|--------------|--------------|--------------|--------------|-------------|
| Abordagem 3 | | | | | | |
| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
| 1 | 21.74 | 67.39 | 89.13 | 75.61 | 78.26 | 0.88 |
| 2 | 20.65 | 63.04 | 95.05 | 75.32 | 79.35 | 0.91 |
| 3 | 22.83 | 65.22 | 89.13 | 74.17 | 77.17 | 0.87 |
| 4 | 23.91 | 60.87 | 91.30 | 71.79 | 76.09 | 0.88 |
| 5 | 21.74 | 69.57 | 86.96 | 76.19 | 78.26 | 0.88 |
| 6 | 21.74 | 69.57 | 86.96 | 76.19 | 78.26 | 0.88 |
| Abordagem 4 | | | | | | |
| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
| 1 | 19.57 | 71.74 | 89.13 | 78.57 | 80.43 | 0.88 |
| 2 | 20.65 | 65.22 | 93.48 | 75.95 | 79.35 | 0.90 |
| 3 | 21.71 | 67.39 | 89.13 | 75.61 | 78.26 | 0.87 |
| 4 | 21.74 | 67.39 | 89.13 | 75.61 | 78.26 | 0.88 |
| 5 | 20.65 | 73.91 | 84.78 | 78.16 | 79.35 | 0.88 |
| 6 | 20.65 | 76.09 | 82.61 | 78.65 | 79.35 | 0.88 |
| Abordagem 5 | | | | | | |
| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | - |
| 1 | 19.57 | 78.26 | 82.61 | 80.00 | 80.43 | - |
| 2 | 17.39 | 76.09 | 89.13 | 81.40 | 82.61 | - |
| 3 | 19.57 | 80.43 | 80.43 | 80.43 | 80.43 | - |
| 4 | 19.57 | 78.26 | 82.61 | 80.00 | 80.43 | - |
| 5 | 22.83 | 82.61 | 71.74 | 78.35 | 77.17 | - |
| 6 | 20.65 | 82.61 | 76.09 | 80.00 | 79.35 | - |

Tabela 4.18: Resultados obtidos analisando os nódulos com o parênquima na proporção $D/4$ por volume nas abordagens 1, 2, 3, 4 e 5.

Ao analisarmos os resultados dos nódulos com parênquima por volume (caso $D/4$), a AUC de maior valor foi obtida pela arquitetura 2 com 0.91 na abordagem três; a maior taxa de sensibilidade foi conseguida pelas arquiteturas 5 e 6 com 82% na abordagem cinco; a maior taxa de especificidade foi conseguida pelas arquitetura 2 com 95% na abordagem três; a maior taxa de *f-measure* foi alcançada pela arquitetura 2 na abordagem cinco com 81%; e a maior acurácia foi conseguida pela arquitetura 2 na abordagem 5 com 82%, e conseqüentemente a menor taxa de erro com 17%. Comparando com os resultados aqui com a avaliação dos nódulos com parênquima (caso $D/4$) por corte, a AUC por volume também foi maior, diferença de 0.1, 0.92 contra 0.82.

A seguir na tabela 4.19, estão os resultados obtidos pelas seis arquiteturas (Seção 3.3) analisando os nódulos juntos com o parênquima na proporção $D/8$ por volume nas abordagens 1, 2, 3, 4, e 5 (Seção 3.5).

Abordagem 1

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | Nº de Empates |
|--------------|--------------|--------------|--------------|--------------|--------------|---------------|
| 1 | 20.65 | 73.91 | 84.78 | 78.16 | 79.35 | 0 |
| 2 | 20.65 | 78.26 | 80.43 | 79.12 | 79.35 | 0 |
| 3 | 19.57 | 71.74 | 89.13 | 78.57 | 80.43 | 2 |
| 4 | 21.74 | 69.57 | 86.96 | 76.19 | 78.26 | 2 |
| 5 | 22.83 | 71.74 | 82.61 | 75.86 | 77.17 | 2 |
| 6 | 21.74 | 69.57 | 86.96 | 76.19 | 78.26 | 3 |

Abordagem 2

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | Nº de Empates |
|--------------|--------------|--------------|--------------|--------------|--------------|---------------|
| 1 | 20.65 | 73.91 | 84.78 | 78.16 | 79.35 | 0 |
| 2 | 20.65 | 78.26 | 80.43 | 79.12 | 79.35 | 0 |
| 3 | 21.74 | 71.74 | 84.78 | 76.74 | 78.26 | 2 |
| 4 | 21.74 | 71.74 | 84.78 | 76.74 | 78.26 | 2 |
| 5 | 20.65 | 76.09 | 82.61 | 78.65 | 79.35 | 2 |
| 6 | 22.83 | 71.74 | 82.61 | 75.86 | 77.17 | 3 |

Abordagem 3

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
|--------------|--------------|--------------|--------------|--------------|--------------|-------------|
| 1 | 22.83 | 65.22 | 89.13 | 74.07 | 77.17 | 0.88 |
| 2 | 21.74 | 69.57 | 86.96 | 76.19 | 78.26 | 0.88 |
| 3 | 22.83 | 66.04 | 91.30 | 73.42 | 77.17 | 0.88 |
| 4 | 21.74 | 67.39 | 89.13 | 75.61 | 78.26 | 0.84 |
| 5 | 21.74 | 67.39 | 89.13 | 75.61 | 78.26 | 0.88 |
| 6 | 30.43 | 45.65 | 93.48 | 60.00 | 69.57 | 0.88 |

Abordagem 4

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
|--------------|--------------|--------------|--------------|--------------|--------------|-------------|
| 1 | 18.48 | 73.91 | 89.13 | 80.00 | 81.52 | 0.87 |
| 2 | 19.57 | 80.43 | 80.43 | 80.43 | 80.43 | 0.88 |
| 3 | 19.57 | 69.57 | 91.30 | 78.05 | 80.43 | 0.87 |
| 4 | 20.65 | 71.74 | 86.96 | 77.65 | 79.35 | 0.87 |
| 5 | 21.74 | 69.57 | 86.96 | 76.19 | 78.26 | 0.88 |
| 6 | 22.83 | 65.22 | 89.13 | 74.07 | 77.17 | 0.89 |

Abordagem 5

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | - |
|--------------|--------|--------|--------|-------|--------|---|
|--------------|--------|--------|--------|-------|--------|---|

| | | | | | | |
|----------|--------------|--------------|--------------|--------------|--------------|---|
| 1 | 21.74 | 80.43 | 76.09 | 78.72 | 78.26 | - |
| 2 | 20.65 | 84.78 | 73.91 | 80.41 | 79.35 | - |
| 3 | 18.48 | 82.61 | 80.43 | 81.72 | 81.52 | - |
| 4 | 23.91 | 76.09 | 76.09 | 76.09 | 76.09 | - |
| 5 | 21.74 | 78.26 | 78.26 | 78.26 | 78.26 | - |
| 6 | 19.57 | 78.26 | 82.61 | 80.00 | 80.43 | - |

Tabela 4.19: Resultados obtidos analisando os nódulos com o parênquima na proporção $D/8$ por volume nas abordagens 1, 2, 3, 4 e 5.

Ao analisarmos os resultados dos nódulos com parênquima por volume (caso $D/8$), a AUC de maior valor foi obtida pela arquitetura 6 com 0.89 na abordagem quatro; a maior taxa de sensibilidade foi conseguida pela arquitetura 2 com 84% na abordagem cinco; a maior taxa de especificidade foi conseguida pela arquitetura 6 com 91% na abordagem quatro; a maior taxa de f -measure foi alcançada pela arquitetura 3 na abordagem cinco com 81%; e a maior acurácia foi conseguida pela arquitetura 1 na abordagem quatro, e pela arquitetura 3 na abordagem cinco com 81%, e consequentemente a menor taxa de erro com 16%. Comparando com os resultados aqui com a avaliação dos nódulos com parênquima (caso $D/8$) por corte, a AUC por volume também foi maior, diferença de 0.07, 0.91 contra 0.84.

A partir de agora serão mostrados os resultados obtidos na análise do parênquima segmentado na análise por volume nas proporções $D/2$, $D/4$ e $D/8$.

A seguir na tabela 4.20, estão os resultados obtidos pelas seis arquiteturas (3.3) analisando o parênquima segmentado na proporção $D/2$ por volume nas abordagens 1, 2, 3, 4, e 5 (3.5).

Abordagem 1

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | Nº de Empates |
|--------------|--------------|--------------|--------------|--------------|--------------|---------------|
| 1 | 22.83 | 60.87 | 93.48 | 72.73 | 77.17 | 2 |
| 2 | 23.91 | 60.87 | 91.30 | 71.79 | 76.09 | 2 |
| 3 | 23.91 | 63.04 | 89.13 | 72.50 | 76.09 | 2 |
| 4 | 22.83 | 65.22 | 89.13 | 74.07 | 77.17 | 3 |
| 5 | 21.74 | 65.22 | 91.30 | 75.00 | 78.26 | 2 |
| 6 | 22.83 | 60.87 | 93.48 | 72.73 | 77.17 | 4 |

Abordagem 2

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | Nº de Empates |
|--------------|--------|--------------|--------------|--------------|--------|---------------|
| 1 | 22.83 | 63.04 | 91.30 | 73.42 | 77.17 | 2 |
| 2 | 23.91 | 63.04 | 89.13 | 72.50 | 76.09 | 2 |
| 3 | 23.91 | 65.22 | 86.96 | 73.17 | 76.09 | 2 |
| 4 | 21.74 | 69.57 | 86.96 | 76.19 | 78.26 | 3 |
| 5 | 21.74 | 67.38 | 89.13 | 75.61 | 78.26 | 2 |

| | | | | | | |
|----------|--------------|-------|--------------|-------|--------------|---|
| 6 | 20.65 | 67.39 | 91.30 | 76.54 | 79.35 | 4 |
|----------|--------------|-------|--------------|-------|--------------|---|

Abordagem 3

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
|--------------|--------------|--------------|---------------|--------------|--------------|-------------|
| 1 | 27.17 | 47.83 | 97.83 | 63.77 | 72.83 | 0.89 |
| 2 | 28.26 | 47.83 | 95.65 | 62.86 | 71.74 | 0.90 |
| 3 | 27.17 | 50.00 | 95.65 | 64.79 | 72.83 | 0.90 |
| 4 | 27.17 | 50.00 | 95.65 | 64.79 | 72.83 | 0.90 |
| 5 | 29.35 | 45.65 | 95.65 | 60.87 | 70.65 | 0.90 |
| 6 | 29.35 | 41.30 | 100.00 | 58.46 | 70.65 | 0.91 |

Abordagem 4

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
|--------------|--------------|--------------|--------------|--------------|--------------|-------------|
| 1 | 27.17 | 47.83 | 97.83 | 63.77 | 72.83 | 0.89 |
| 2 | 25.00 | 58.70 | 91.30 | 70.13 | 75.00 | 0.90 |
| 3 | 23.91 | 60.87 | 91.30 | 71.79 | 76.09 | 0.90 |
| 4 | 21.74 | 65.22 | 91.30 | 75.00 | 78.26 | 0.89 |
| 5 | 25.00 | 58.70 | 91.30 | 70.13 | 75.00 | 0.90 |
| 6 | 25.00 | 56.52 | 93.48 | 69.33 | 75.00 | 0.91 |

Abordagem 5

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | - |
|--------------|--------------|--------------|--------------|--------------|--------------|---|
| 1 | 22.83 | 67.39 | 86.96 | 74.70 | 77.17 | - |
| 2 | 22.83 | 69.57 | 84.78 | 75.29 | 77.17 | - |
| 3 | 22.83 | 69.57 | 84.78 | 75.29 | 77.17 | - |
| 4 | 21.74 | 71.74 | 84.78 | 76.74 | 78.26 | - |
| 5 | 22.83 | 69.57 | 84.78 | 75.29 | 77.17 | - |
| 6 | 19.57 | 69.57 | 91.30 | 78.05 | 80.43 | - |

Tabela 4.20: Resultados obtidos analisando o parênquima segmentado por volume na proporção $D/2$ nas abordagens 1, 2, 3, 4 e 5.

Ao analisarmos os resultados do parênquima segmentado por volume ($D/2$), a AUC de maior valor foi obtida pela arquitetura 6 com 0.91 nas abordagens três e quatro; a maior taxa de sensibilidade foi conseguida pela arquitetura 4 com 71% na abordagem cinco; a maior taxa de especificidade foi conseguida pela arquitetura 6 com 100% na abordagem três; a maior taxa de *f-measure* foi alcançada pela arquitetura 6 na abordagem cinco com 78%; e a maior acurácia foi conseguida pela arquitetura 6 na abordagem cinco com 80%, e consequentemente a menor taxa de erro com 19%. Comparando com os resultados aqui com a avaliação do parênquima segmentador (caso $D/2$) por corte, a AUC por volume também foi maior, diferença de 0.07, 0.91 contra 0.84.

A seguir na tabela 4.21, estão os resultados obtidos pelas seis arquiteturas (Seção 3.3) analisando o parênquima segmentado na proporção $D/4$ por volume nas abordagens 1, 2, 3, 4, e 5 (Seção 3.5).

Abordagem 1

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | Nº de Empates |
|--------------|--------------|--------------|--------------|--------------|--------------|---------------|
| 1 | 21.74 | 65.22 | 91.30 | 75.00 | 78.26 | 2 |
| 2 | 26.09 | 58.70 | 89.13 | 69.23 | 73.91 | 3 |
| 3 | 23.91 | 63.04 | 89.13 | 72.50 | 76.09 | 3 |
| 4 | 22.83 | 63.04 | 91.30 | 73.42 | 77.17 | 4 |
| 5 | 23.91 | 60.87 | 91.30 | 71.79 | 76.09 | 4 |
| 6 | 22.83 | 63.04 | 91.30 | 73.42 | 77.17 | 4 |

Abordagem 2

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | Nº de Empates |
|--------------|--------------|--------------|--------------|--------------|--------------|---------------|
| 1 | 21.74 | 67.39 | 89.13 | 75.61 | 78.26 | 2 |
| 2 | 25.00 | 63.04 | 86.96 | 71.60 | 75.00 | 3 |
| 3 | 22.83 | 67.39 | 86.96 | 74.70 | 77.17 | 3 |
| 4 | 28.26 | 47.83 | 95.65 | 62.86 | 71.74 | 4 |
| 5 | 21.74 | 67.39 | 89.13 | 75.61 | 78.26 | 4 |
| 6 | 22.83 | 67.39 | 86.96 | 74.70 | 77.17 | 4 |

Abordagem 3

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
|--------------|--------------|--------------|--------------|--------------|--------------|-------------|
| 1 | 25.00 | 52.17 | 97.83 | 67.61 | 75.00 | 0.91 |
| 2 | 27.17 | 50.00 | 95.65 | 64.79 | 72.83 | 0.91 |
| 3 | 28.26 | 47.83 | 95.65 | 62.86 | 71.74 | 0.91 |
| 4 | 28.26 | 47.83 | 95.65 | 62.86 | 71.74 | 0.91 |
| 5 | 28.26 | 45.65 | 97.83 | 61.76 | 71.74 | 0.90 |
| 6 | 26.09 | 50.00 | 97.83 | 65.71 | 73.91 | 0.91 |

Abordagem 4

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
|--------------|--------------|--------------|--------------|--------------|--------------|-------------|
| 1 | 20.65 | 65.22 | 93.48 | 75.95 | 79.35 | 0.91 |
| 2 | 25.00 | 58.70 | 91.30 | 70.13 | 75.00 | 0.91 |
| 3 | 22.83 | 63.04 | 91.30 | 73.42 | 77.17 | 0.91 |
| 4 | 23.91 | 60.87 | 91.30 | 71.79 | 76.09 | 0.91 |
| 5 | 22.83 | 63.04 | 91.30 | 73.42 | 77.17 | 0.91 |
| 6 | 22.83 | 63.04 | 91.30 | 73.42 | 77.17 | 0.91 |

Abordagem 5

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | - |
|--------------|--------|--------|--------|-------|--------|---|
|--------------|--------|--------|--------|-------|--------|---|

| | | | | | | |
|----------|--------------|--------------|--------------|--------------|--------------|---|
| 1 | 19.57 | 71.74 | 89.13 | 78.57 | 80.43 | - |
| 2 | 22.83 | 67.39 | 86.96 | 74.70 | 77.17 | - |
| 3 | 21.74 | 69.57 | 86.96 | 76.19 | 78.26 | - |
| 4 | 21.74 | 69.57 | 86.96 | 76.19 | 78.26 | - |
| 5 | 20.65 | 69.57 | 89.13 | 77.11 | 79.35 | - |
| 6 | 21.74 | 69.57 | 86.96 | 76.19 | 78.26 | - |

Tabela 4.21: Resultados obtidos analisando o parênquima segmentado por volume na proporção $D/4$ nas abordagens 1, 2, 3, 4 e 5.

Ao analisarmos os resultados do parênquima segmentado por volume (caso $D/4$), a AUC de maior valor foi obtida por todas as arquiteturas com 0.91 nas abordagens três e quatro, com exceção da arquitetura 5 na abordagem três; a maior taxa de sensibilidade foi conseguida pelas arquitetura 1 com 71% na abordagem cinco; a maior taxa de especificidade foi conseguida pela arquitetura 1 com 97% na abordagem três; a maior taxa de f -measure foi alcançada pela arquitetura 1 na abordagem cinco com 78%; e a maior acurácia foi conseguida pela arquitetura 1 na abordagem cinco com 80%, e consequentemente a menor taxa de erro com 19%. Comparando com os resultados aqui com a avaliação do parênquima segmentado (caso $D/4$) por corte, a AUC por volume também foi maior, diferença de 0.07, 0.91 contra 0.84.

A seguir na tabela 4.22, estão os resultados obtidos pelas seis arquiteturas (Seção 3.3) analisando o parênquima segmentado na proporção $D/8$ por volume nas abordagens 1, 2, 3, 4, e 5 (Seção 3.5).

Abordagem 1

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | Nº de Empates |
|--------------|--------------|--------------|--------------|--------------|--------------|---------------|
| 1 | 20.65 | 76.09 | 82.61 | 78.65 | 79.35 | 3 |
| 2 | 20.65 | 63.04 | 95.65 | 75.32 | 79.35 | 1 |
| 3 | 20.65 | 89.13 | 69.57 | 81.19 | 79.35 | 0 |
| 4 | 21.74 | 65.22 | 91.30 | 78.26 | 78.26 | 2 |
| 5 | 23.91 | 60.87 | 91.30 | 71.79 | 76.09 | 2 |
| 6 | 22.83 | 60.87 | 93.48 | 72.73 | 77.17 | 4 |

Abordagem 2

| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | Nº de Empates |
|--------------|--------------|--------------|--------------|--------------|--------------|---------------|
| 1 | 19.57 | 80.43 | 80.43 | 80.43 | 80.43 | 3 |
| 2 | 21.74 | 63.04 | 93.48 | 74.36 | 78.26 | 1 |
| 3 | 20.65 | 89.13 | 69.57 | 81.19 | 79.35 | 0 |
| 4 | 19.57 | 69.57 | 91.30 | 78.05 | 80.43 | 2 |
| 5 | 23.91 | 63.04 | 89.13 | 72.50 | 76.09 | 2 |
| 6 | 20.65 | 67.39 | 91.30 | 76.54 | 79.35 | 4 |

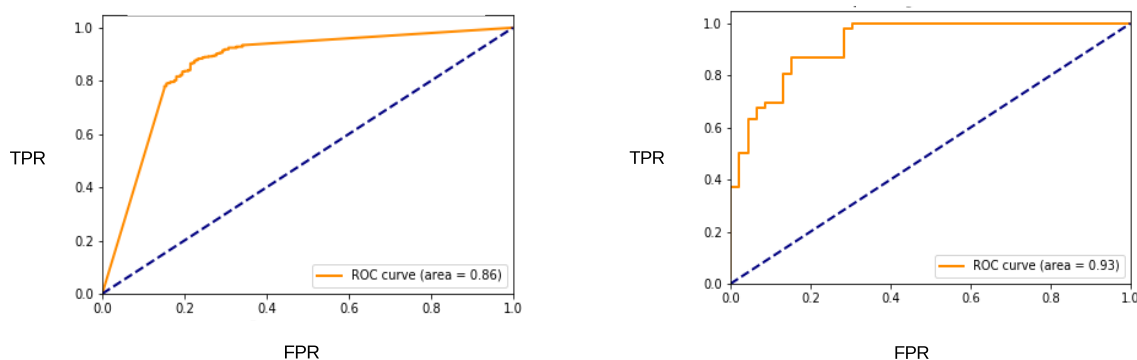
| Abordagem 3 | | | | | | |
|---------------------|---------------|---------------|---------------|--------------|---------------|-------------|
| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
| 1 | 22.83 | 67.39 | 86.96 | 74.70 | 77.17 | 0.91 |
| 2 | 28.26 | 43.48 | 100.00 | 60.61 | 71.74 | 0.92 |
| 3 | 21.74 | 78.26 | 78.26 | 78.26 | 78.26 | 0.90 |
| 4 | 23.91 | 55.52 | 95.65 | 70.27 | 76.09 | 0.92 |
| 5 | 27.17 | 50.00 | 95.65 | 64.79 | 72.83 | 0.92 |
| 6 | 25.00 | 54.35 | 95.65 | 68.49 | 75.00 | 0.92 |

| Abordagem 4 | | | | | | |
|---------------------|---------------|---------------|---------------|--------------|---------------|-------------|
| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | AUC |
| 1 | 19.57 | 76.09 | 84.78 | 79.55 | 80.43 | 0.91 |
| 2 | 26.09 | 52.17 | 95.65 | 66.67 | 73.91 | 0.93 |
| 3 | 17.39 | 89.13 | 76.09 | 83.67 | 82.61 | 0.90 |
| 4 | 20.65 | 65.22 | 93.48 | 75.95 | 79.35 | 0.92 |
| 5 | 21.74 | 63.04 | 93.48 | 74.36 | 78.26 | 0.92 |
| 6 | 20.65 | 65.22 | 93.48 | 75.95 | 79.35 | 0.92 |

| Abordagem 5 | | | | | | |
|---------------------|---------------|---------------|---------------|--------------|---------------|----------|
| Arquiteturas | Err(%) | SEN(%) | ESP(%) | FM(%) | ACC(%) | - |
| 1 | 18.48 | 84.78 | 78.26 | 82.11 | 81.52 | - |
| 2 | 20.65 | 69.57 | 89.13 | 77.11 | 79.35 | - |
| 3 | 20.65 | 91.30 | 67.39 | 81.55 | 79.35 | - |
| 4 | 20.65 | 71.74 | 86.96 | 77.65 | 79.35 | - |
| 5 | 20.65 | 69.57 | 89.13 | 77.11 | 79.35 | - |
| 6 | 19.57 | 71.74 | 89.13 | 78.57 | 80.43 | - |

Tabela 4.22: Resultados obtidos analisando o parênquima segmentado por volume na proporção $D/8$ nas abordagens 1, 2, 3, 4 e 5.

Ao analisarmos os resultados do parênquima segmentado por volume (caso $D/8$), a AUC de maior valor foi obtida pela arquitetura 2 com 0.93 na abordagem quatro; a maior taxa de sensibilidade foi conseguida pela arquitetura 3 com 91% na abordagem cinco; a maior taxa de especificidade foi conseguida pela arquitetura 2 com 100% na abordagem três; a maior taxa de f -measure foi alcançada pela arquitetura 3 na abordagem quatro com 83%; e a maior acurácia foi conseguida pela arquitetura 3 na abordagem quatro com 82%, e consequentemente a menor taxa de erro com 18%. Comparando com os resultados aqui com a avaliação do parênquima segmentado (caso $D/8$) por corte, a AUC por volume também foi maior, diferença de 0.07, 0.93 contra 0.86 (Figura 4.1).



(a) AUC obtida na segmentação do parênquima na proporção $D/8$ por corte.

(b) AUC obtida na segmentação do parênquima na proporção $D/8$ por volume.

Figura 4.1: AUC obtidas na segmentação do parênquima na proporção $D/8$ por cortes (a) e por volume (b). Fonte: elaborado pelo autor.

A seguir na tabela 4.23, está resumido os melhores resultados alcançados na avaliação dos nódulos por volume em cada métrica, a arquitetura e a abordagem (Seção 3.5) utilizadas:

| Métrica | Resultado | Arquitetura - abordagem |
|----------------|-----------|---|
| Err.(%) | 16.30 | 3 - nódulos segmentados, abordagem 5; 5 - nódulos com parênquima (caso $D/2$), abordagens 1 e 5. |
| SEN(%) | 91.30 | 3 - parênquima segmentado (caso $D/8$), abordagem 5. |
| ESP(%) | 100.00 | 6 - parênquima segmentado (caso $D/2$), abordagem 3; 2 - parênquima segmentado (caso $D/8$), abordagem 3. |
| FM(%) | 84.21 | 5 - nódulos com parênquima (caso $D/2$), abordagem 5. |
| ACC(%) | 83.70 | 3 - nódulos segmentados, abordagem 5; 1 - nódulos com parênquima (caso $D/2$); 5 - nódulos com parênquima (caso $D/2$). |
| AUC | 0.93 | 2 - parênquima segmentado (caso $D/8$), abordagem 4. |

Tabela 4.23: Melhores resultados na avaliação por corte.

Pela tabela 4.23 vemos que, com exceção da sensibilidade e da AUC, os outros resultados foram conseguidos em mais de uma abordagem. Diferente dos resultados dos nódulos na avaliação por corte (Tabela 4.15) que obteve os melhores resultados através apenas da avaliação do parênquima segmentado na proporção $D/8$, aqui houve variação. Comparando os resultados entre as avaliações por corte e por volume, a avaliação por volume obteve apenas especificidade e AUC melhores, nas outras métricas os resultados da avaliação por corte foram melhores.

Em experimentos primeiramente feitos com os nódulos segmentados, foi delimitado o limite para os hiperparâmetros “Filtro” e “Unidades” de até 1024, porém os valores selecionados

estavam sendo abaixo de 512, então foi fixado a exploração desses hiperparâmetros em até 512 nas seis arquiteturas exploradas nas sete bases. Também foi realizado testes variando as funções de ativações com o algoritmo TPE, mas os resultados não estavam sendo satisfatórios, então foi fixado as funções de ativação como *ReLU* nas camadas de convolução e totalmente conectadas, com exceção da última camada totalmente conectada onde foi estabelecido a função de ativação *Softmax* e os resultados começaram a melhorar, tomando como base o trabalho de Silva et al. (SILVA; PAIVA; SILVA, 2017), que conseguiu bons resultados.

A partir dos resultados por corte e comparando as proporções $D/2$, $D/4$ e $D/8$ do nódulo com o parênquima, a maior AUC foi na proporção $D/8$ de 0.84, na menor proporção, o mesmo valor conseguido dos nódulos segmentados, seguido da proporção $D/4$ com AUC de 0.83, e da proporção $D/8$ com AUC de 0.82. Esses resultados da AUC nessas três proporções mostraram que nessa análise a menor proporção contribuiu mais do que as outras duas, porém com a mesma eficiência que na análise dos nódulos segmentados. Já, comparando as proporções $D/2$, $D/4$ e $D/8$ na análise do parênquima segmentado, a maior AUC foi de 0.86 na proporção $D/8$, a menor proporção, e maior do que conseguido dos nódulos segmentados, com AUC de 0.84, seguido da proporção $D/2$ com AUC de 0.84 e *f-measure* de 90% e da proporção $D/4$ com AUC de 0.84 e *f-measure* de 89%. Esses resultados nessas três proporções mostraram que nessa análise a menor proporção também contribuiu mais do que as outras duas, e já passou a ter uma AUC melhor do que a análise dos nódulos segmentados. Esses resultados mostraram que, analisar o parênquima segmentado por corte na proporção $D/8$ contribui mais para a performance de classificação dos nódulos seguido da análise do nódulo com parênquima, e da análise de apenas o nódulo.

A partir dos resultados por volume e comparando as proporções $D/2$, $D/4$ e $D/8$ do nódulo com o parênquima, a maior AUC foi na proporção $D/4$ de 0.92, o mesmo valor conseguido dos nódulos segmentados, seguido da proporção $D/8$ com AUC de 0.91, e da proporção $D/2$ com AUC de 0.89. Esses resultados da AUC nessas três proporções mostraram que a proporção $D/4$ contribuiu mais do que as outras duas, porém com a mesma eficiência que na análise dos nódulos segmentados. Já, Comparando as proporções $D/2$, $D/4$ e $D/8$ na análise do parênquima segmentado, a maior AUC foi de 0.93 na proporção $D/8$, maior AUC conseguida neste trabalho, e na menor proporção, seguido da proporção $D/2$ com AUC de 0.91 e especificidade de 100% e da proporção $D/4$ com AUC de 0.91 e especificidade de 97%. Esses resultados nessas três proporções mostraram que ao analisar o parênquima segmentado na menor proporção foram melhores e isso pode ser entendido pelas seguintes razões: o parênquima na menor proporção apresenta mais detalhes para diferenciar os nódulos benignos dos nódulos malignos em imagens de TC para uma RNC, uma ROI com o nódulo e o parênquima pode conter ruídos como a pleura e os vasos, e o parênquima segmentado contém no seu interior a forma do nódulo alinhado a textura do próprio parênquima que deve ser mais significativo do que a textura apenas dos nódulos segmentados, que devem contribuir mais com os seus contornos do que a textura para uma RNC.

Logo, entre analisar nódulo segmentado ou nódulo com parênquima ou parênquima segmentado utilizando uma RNC para a classificação dos nódulos pulmonares, é melhor utilizar uma base do parênquima segmentado e na proporção $D/8$, a mais próxima dos nódulos, para alcançar resultados da AUC maiores, visto que, a AUC é a métrica mais importante para avaliar um classificação de nódulos pulmonares.

A tabela 4.24 apresenta uma comparação entre os resultados encontrados neste trabalho e os trabalhos relacionados apresentados na seção 2.7. É importante salientar que para realizar uma comparação justa, seria necessário que estes trabalhos relacionados tivessem utilizado os mesmos dados nos conjuntos de treinamento, valiação e teste da base LIDC-IDRI, as mesmas métricas de avaliação dos resultados, entre outros parâmetros. Mesmo assim, a partir dos resultados alcançados neste trabalho, foi possível fazer uma comparação com os trabalhos que também focaram na classificação de nódulos pulmonares.

| Trabalho | SEN (%) | ESP (%) | ACC (%) | AUC |
|---|--------------|--------------|--------------|-------------|
| Obayya et al. (2015) | 96 | 96 | 96 | - |
| Silva et al. (2016) | 84.22 | 90.06 | 88.44 | 0.871 |
| Wang et al. (2016) | 74.60 | 78.90 | 76.10 | - |
| Dhara et al. (2016) | 89.73 | 86.36 | - | 0.950 |
| Ferreira Junior et al. (2018) | 70.2 | 85.6 | 80.0 | 0.858 |
| Reeves et al. (2015) | - | - | - | 0.708 |
| Hua et al. (2015b) | 73.4 | 82.2 | - | - |
| Shen et al. (2017) | 77 | 93 | 87.14 | 0.93 |
| Silva et al. (2017) | 94.66 | 95.14 | 94.78 | 0.949 |
| Song et al. (2017) | 83.96 | 84.32 | 84.15 | - |
| Dilger et al. (2015) | 90.91 | 92.86 | 92.0 | 0.938 |
| Felix et al. (2016) | - | - | - | 0.875 |
| Metodologia proposta para o diagnóstico dos nódulos por corte | 95.72 | 92.33 | 85.78 | 0.86 |
| Metodologia proposta para o diagnóstico dos nódulos por volume | 91.30 | 100 | 83.71 | 0.93 |

Tabela 4.24: Comparação com outros trabalhos publicados com foco em classificar nódulos pulmonares por meio de imagens de TC.

Como pode ser observado na tabela 4.24, conseguimos uma taxa de sensibilidade de 95%, superando as taxas de sensibilidade conseguida por Silva et al. (SILVA; PAIVA; SILVA, 2017), que conseguiu uma taxa de sensibilidade de 94.66%, e de Dilger (DILGER et al., 2015), que conseguiu uma taxa de sensibilidade de 90.91%, mas não de Obayya et al. (OBAYYA; GHAN-DOUR, 2015), que conseguiu uma sensibilidade de 96%. Contudo, Obayya não mencionou o tamanho dos nódulos utilizados, podendo ter utilizado nódulos maiores que 30mm. Na taxa de

especificidade, conseguimos o melhor resultado de 100%, onde as taxas de sensibilidade acabaram sendo baixas na consequência. A melhor acurácia conseguida neste trabalho foi de 85.78%, porém Obayya et al. (OBAYYA; GHANDOUR, 2015), Silva et al. (SILVA; PAIVA; SILVA, 2017), Shen et al. (SHEN et al., 2017) e Dilger et al. (DILGER et al., 2015) conseguiram acurácia melhores. Todavia, Dilger et al. utilizou uma base de nódulos diferente da utilizada neste trabalho. Já analisando a AUC, conseguimos a melhor AUC de 0.93, porém não superior a de Dhara et al. (DHARA et al., 2016), que conseguiu uma AUC de 0.950, de Silva et al. (SILVA; PAIVA; SILVA, 2017), que conseguiu uma AUC de 0.949 e de Dilger et al. (DILGER et al., 2015), que conseguiu uma AUC de 0.938, entretanto ficou próximo.

5 Conclusão

O câncer de pulmão já é o câncer que mais causa óbito no mundo e a tendência é de causar mais óbitos nos próximos anos. As altas taxas de mortalidade demonstram a importância da necessidade de uma ferramenta computacional que possa auxiliar os especialistas, e assim contribuir para facilitar o diagnóstico.

Este trabalho apresentou um estudo e um modelo computacional cumprindo o objetivo proposto no resumo para a classificação de nódulos pulmonares utilizando RNC aliado ao algoritmo TPE para otimizar os hiperparâmetros em três regiões: nódulos segmentados, nódulo com parênquima e parênquima segmentado.

As RNCs têm a capacidade de extrair as características dos detalhes mais simples aos mais complexos, e no fim ainda classificá-las selecionando as características mais importantes para diferenciar as classes (benigna e maligna). A otimização hiperparamétrica é outro passo necessário para o sucesso das RNCs, e um algoritmo que consiga isso de forma rápida, fazendo toda a exploração possível do tamanho da arquitetura e de todos os hiperparâmetros de uma RNC e ainda de forma eficiente é um desafio, porém o algoritmo TPE nos proporcionou ótimos resultados de forma eficiente e explorando praticamente todos os hiperparâmetros possíveis.

Na literatura encontramos muitos trabalhos que tratam do problema da classificação de nódulos pulmonares em imagens de TC, porém pouquíssimos analisaram o parênquima, e estes mostraram que o parênquima elevou a performance de classificação, e neste trabalho comprovamos isso mais vez. Contudo, nossos resultados evidenciaram que, a área do parênquima a ser considerada no processo de classificação é importante. A partir dos resultados obtidos, podemos afirmar que, a região que deve ser utilizada para obter a melhor classificação dos nódulos pulmonares é a do parênquima segmentado na proporção $D/8$, que gerou a maior AUC de 0.93.

As principais contribuições deste trabalho foram:

- Mostrar o quanto de parênquima deve ser utilizado ao usar uma RNC;
- Mostrar/testar/a utilização do algoritmo TPE através da ferramenta Hyperas para otimização hiperparamétrica de uma RNC;
- Evidenciar/confirmar que usar a CNN no parênquima segmentado contribui mais do que nos nódulos segmentados e do que nos nódulos com o parênquima.

5.1 Trabalhos Futuros

Com o objetivo de melhorar o modelo de classificação para nódulos pulmonares, e ainda, incluir nódulos semi-sólidos e não sólidos, os planejamentos futuros são os seguintes:

- Utilizar redes neurais convolucionais 3D;
- Extrair atributos das imagens por meio de técnicas de aprendizagem de máquina, como: atributos de textura, atributos de forma, etc;
- Juntar estes atributos extraídos a redes neurais convolucionais 3D;
- Explorar outras técnicas de otimização hiperparamétrica.

5.2 Contribuições Científicas

Os trabalhos científicos aceitos para apresentação com Qualis CAPES foram:

- Resultados parciais foram aceitos no 18º Simpósio Brasileiro de Computação Aplicado à Saúde (SBCAS), Qualis B4, com o título “Using Deep Learning for Classification of Early Lung Nodules on Computed Tomography Images”. No qual, foi utilizada uma CNN para classificação de nódulos pulmonares precoces com diâmetro de 5-10mm, técnica de aumento de base, e foi utilizado o corte central de cada nódulo. Os resultados alcançados foram uma AUC de 0,992 (desvio padrão 0,001) com uma validação cruzada 10-*fold*;
- Recentemente, enviamos um artigo para o Simpósio Brasileiro de Sistemas de Informação (SBSI) 2019, Qualis B2, e estamos no aguardo do aceite da segunda etapa e última previsto para o dia 01 de Março. Neste artigo, abordamos a mesma metodologia apresentada nesta dissertação, porém utilizamos apenas os nódulos segmentados precoces com diâmetro entre 5-10mm, sem uso do parênquima, e alcançamos os seguintes resultados: taxa de erro de 12%, sensibilidade de 94%, especificidade de 83%, acurácia de 88% e *F-measure* de 89%.

Referências Bibliográficas

- ABADI, M. et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. 2015. Software available from tensorflow.org. Disponível em: <<https://www.tensorflow.org/>>.
- ALBERTS, B. et al. *Biologia molecular da célula*. [S.l.: s.n.], 2010.
- ALILOU, M. et al. A comprehensive framework for automatic detection of pulmonary nodules in lung ct images. *Image Analysis & Stereology*, v. 33, n. 1, p. 13–27, 2014. ISSN 1854-5165.
- ARMATO, S. G. et al. The lung image database consortium (lidc) and image database resource initiative (idri): A completed reference database of lung nodules on ct scans. *Medical Physics*, v. 38, n. 2, p. 915–931, 2011.
- AZEVEDO-MARQUES, P. M. *Diagnóstico auxiliado por computador na radiologia*. [S.l.]: Radiologia Brasileira, 2001. v. 34. (285-293, 5).
- BARTHOLMAI, B. J. et al. Pulmonary nodule characterization, including computer analysis and quantitative features. *Journal of Thoracic Imaging*, v. 30, n. 2, p. 139–156, March 2015. ISSN 1076-6332.
- BBC. *Lung cancer cells spread like unanchored tents, study says*, <<https://www.bbc.com/news/health-38095297>>. 2016. Acessado em 11-12-2018.
- BERGSTRA, J. et al. Algorithms for Hyper-Parameter Optimization. In: SHAWE-TAYLOR, J. et al. (Ed.). *Advances in Neural Information Processing Systems 24*. Curran Associates, Inc., 2011. p. 2546–2554. Disponível em: <<http://papers.nips.cc/paper/4443-algorithms-for-hyper-parameter-optimization.pdf>>.
- BERGSTRA, J.; BENGIO, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.*, JMLR.org, v. 13, n. 1, p. 281–305, fev. 2012. ISSN 1532-4435. Disponível em: <<http://dl.acm.org/citation.cfm?id=2503308.2188395>>.
- BERGSTRA, J.; YAMINS, D.; COX, D. Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In: DASGUPTA, S.; MCALLESTER, D. (Ed.). *Proceedings of the 30th International Conference on Machine Learning*. Atlanta, Georgia, USA: PMLR, 2013. (Proceedings of Machine Learning Research, 1), p. 115–123. Disponível em: <<http://proceedings.mlr.press/v28/bergstra13.html>>.
- BERGSTRA, J.; YAMINS, D.; COX, D. D. *Hyperopt: A Python Library for Optimizing the Hyperparameters of Machine Learning Algorithms*. 2013.
- BISHOP, C. M. *Pattern Recognition and Machine Learning*. [S.l.]: Springer, 2007.

- BUSHBERG, J. et al. *The Essential Physics of Medical Imaging*. 3. ed. [S.l.]: Wolters Kluwer Health, 2011. ISBN 9781451153941.
- CADENE, R. *Master's Thesis Deep Learning for Visual Recognition bibtex*. Dissertação (Mestrado), 2016.
- CHARTRAND, G. et al. Deep learning: A primer for radiologists. *Radiographics : a review publication of the Radiological Society of North America, Inc*, v. 37 7, p. 2113–2131, 2017.
- CHOLLET, F. et al. *Keras*. 2015. Disponível em: <<https://keras.io>>.
- CHUQUICUSMA, M. J. M. et al. How to fool radiologists with generative adversarial networks? A visual turing test for lung cancer diagnosis. *CoRR*, abs/1710.09762, 2017. Disponível em: <<http://arxiv.org/abs/1710.09762>>.
- CRUZ-ROA, A. et al. Automatic Detection of Invasive Ductal Carcinoma in Whole Slide Images with Convolutional Neural Networks. In: *SPIE Medical Imaging*. [s.n.], 2014. v. 9041, p. 904103–904103–15. Disponível em: <<http://dx.doi.org/10.1117/12.2043872>>.
- DENG, L. A tutorial survey of architectures algorithms and applications for deep learning. *APSIPA Transactions on Signal and Information Processing*, v. 3, n. e2, p. 1–29, 2014. Disponível em: <<https://doi.org/10.1017/atsip.2013.9>>.
- DEWANCKER, I.; MCCOURT, M.; CLARK, S. *Bayesian optimization primer*. 2015. Disponível em: <https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=2ahUKEwj8vnz_evfAhVhLLkGHZgdC10QFjAAegQICRAC&url=https%3A%2F%2Fsigopt.com%2Fstatic%2Fpdf%2FSigOpt_Bayesian_Optimization_Primer.pdf&usq=AOvVaw0iOwRc7FOem5rcS4qnvI0X>.
- DHARA, A. K. et al. A combination of shape and texture features for classification of pulmonary nodules in lung ct images. *Journal of Digital Imaging*, p. 1–10, 2016. ISSN 1618-727X.
- DICIOTTI, S.; AL. et. The LoG Characteristic Scale: A Consistent Measurement of Lung Nodule Size in CT Imaging. *IEEE Trans. Med. Imaging*, v. 29, n. 2, p. 397–409, 2010.
- DILGER, S. K. et al. Improved pulmonary nodule classification utilizing lung parenchyma texture features. *Proc. SPIE*, v. 9414, p. 94142T–10, 2015.
- DOI, K. Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. *Computerized Medical Imaging and Graphics*, v. 31, n. 4-5, p. 198–211, 2007.
- DOU., T. H. et al. Peritumoral radiomics features predict distant metastasis in locally advanced nsclc. *BPLoS One*, v. 13, n. 11, 2018.
- ERASMUS, J. J. et al. Solitary pulmonary nodules: Part i. morphologic evaluation for differentiation of benign and malignant lesions. *RadioGraphics*, v. 20, p. 43–58, 2010. PMID: 10682770. Disponível em: <<http://dx.doi.org/10.1148/radiographics.20.1.g00ja0259>>.
- ERHAN, D. et al. Why does Unsupervised Pre-training Help Deep Learning? *Journal of Machine Learning Research*, p. 625–660, 2010. Disponível em: <<http://jmlr.csail.mit.edu/papers/v11/erhan10a.html>>.

- FACELI, K. et al. *Inteligência Artificial: Uma Abordagem de Aprendizagem de Máquina*. [S.l.]: Livros Técnicos e Científicos Editora - LTC, 2011.
- FALCÃO, A. X. *Visualização de Volumes Aplicada à Área Médica*. [S.l.]: Tese de Mestrado - UNICAMP, 1993.
- FERNEDA, E. Redes neurais e sua aplicação em sistemas de recuperação de informação. *Ciência da Informação*, v. 35, n. 1, 2006. ISSN 1518-8353. Disponível em: <<http://dx.doi.org/10.18225/ci.inf.v35i1.1149>>.
- FERREIRA-JUNIOR, J. R.; OLIVEIRA, M. C. *Auxílio Computadorizado ao Diagnóstico do Câncer de Pulmão Otimizado por GPU*. Dissertação (Mestrado) — University of Alagoas - (UFAL), 2015.
- FERREIRA-JUNIOR, J. R.; OLIVEIRA, M. C.; AZEVEDO-MARQUES, P. M. de. Cloud-based nosql open database of pulmonary nodules for computer-aided lung cancer diagnosis and reproducible research. *Journal of Digital Imaging*, p. 1–14, 2016. ISSN 1618-727X.
- FERREIRA-JUNIOR, J. R.; OLIVEIRA, M. C.; AZEVEDO-MARQUES, P. M. de. Characterization of Pulmonary Nodules Based on Features of Margin Sharpness and Texture. *J. Digital Imaging*, v. 31, n. 4, p. 451–463, 2018.
- FILHO, A. L.; MACHADO, A. P.; OLIVEIRA, M. *Modelo para Classificação de Nódulos Pulmonares Pequenos usando Descritores Radiomics*. Dissertação (Mestrado) — University of Alagoas - (UFAL), 2016.
- FUJIMOTO, J.; WISTUBA, I. I. Current Concepts on the Molecular Pathology of Non-small Cell Lung Carcinoma. In: *ELSEVIER. Seminars in diagnostic pathology. United States*, v. 31, n. 4, p. 306–313, 2014.
- FUKUSHIMA, K. Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position. v. 36, p. 193–202, 02 1980.
- GILLIES, R. J.; KINAHAN, P. E.; HRICAK, H. Radiomics: Images are more than pictures, they are data. *Radiology*, v. 278, n. 2, p. 563–577, 2016. PMID: 26579733.
- GONZALEZ, R. C.; WOODS, R. E. *Processamento Digital De Imagens*. 3. ed. [S.l.]: Pearson Education, 2010.
- GUO, Y. et al. Deep learning for visual understanding. *Neurocomput.*, Elsevier Science Publishers B. V., Amsterdam, The Netherlands, The Netherlands, v. 187, n. C, p. 27–48, abr. 2016. ISSN 0925-2312. Disponível em: <<http://dx.doi.org/10.1016/j.neucom.2015.09.116>>.
- HENSCHKE, C. I. et al. Early lung cancer action project: overall design and findings from baseline screening. *The Lancet*, v. 354, n. 9173, p. 99 – 105, 1999. ISSN 0140-6736.
- HUA, K.-L. et al. Computer-aided classification of lung nodules on computed tomography images via deep learning technique. *OncoTargets and Therapy*, 8:2015–2022, 2015.
- HUA, K.-L. et al. Computer-aided classification of lung nodules on computed tomography images via deep learning technique. In: *OncoTargets and therapy*. [S.l.: s.n.], 2015.

HUANG, L.-K.; WANG, M.-J. J. Image thresholding by minimizing the measures of fuzziness. *Pattern Recognition*, v. 28, p. 41–51, 1995.

INCA. INCA, <<http://www1.inca.gov.br/>>. 2018. Acessado em 21-11-2018.

JUNIOR, W. J. de A.; LEMOS, A. P. *Métodos de Otimização Hiperparamétrica: Um Estudo Comparativo Utilizando Árvores de Decisão e Florestas Aleatórias na Classificação Binária*. Dissertação (Mestrado) — University of Minas Gerais - (UFMG), 2018.

KADOTA, K.; AL et. Tumor Spread through Air Spaces is an Important Pattern of Invasion and Impacts the Frequency and Location of Recurrences after Limited Resection for Small Stage I Lung Adenocarcinomas. *Journal of Thoracic Oncology*, v. 10, p. 806–814, 2015.

KNIGHT, S. B. et al. Progress and Prospects of Early Detection in Lung Cancer. *Royal Society Journals*, v. 7, n. 9, 2017.

KOOI, T.; AL et. Large Scale Deep Learning for Computer Aided Detection of Mammographic Lesions. *Medical image analysis*, v. 35, p. 303–312, 2017.

LANTZ, B. *Machine Learning with R*. 2. ed. [S.l.: s.n.], 2015.

LECUN, Y. et al. Gradient-based learning applied to document recognition. v. 86, p. 2278 – 2324, 12 1998.

LECUN, Y.; KAVUKCUOGLU, K.; FARABET, C. Convolutional Networks and Applications in Vision. p. 253–256, 05 2010.

LECUNN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *Nature*, v. 521, p. 436–444, 2015.

LITJENS, G. J. S. et al. A survey on deep learning in medical image analysis. *CoRR*, abs/1702.05747, 2017. Disponível em: <<http://arxiv.org/abs/1702.05747>>.

MAXUMPERIA. *Hyperas*. 2017. Disponível em: <<https://pypi.org/project/hyperas/>>.

MITCHELL, T. M. *Machine Learning*. 1. ed. [S.l.]: McGraw-Hill Education, 1997.

MOSMANN, M. P. et al. Solitary pulmonary nodule and 18F-FDG PET/CT. Part 1: epidemiology, morphological evaluation and cancer probability. *Radiologia Brasileira*, scielo, v. 49, p. 35 – 42, 02 2016. ISSN 0100-3984. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-39842016000100035&nrm=iso>.

OBAYYA, M.; GHANDOUR, M. Lung Cancer Classification using Curvelet Transform and Neural Network with Radial Basis Function. *International Journal of Computer Applications*, v. 120, n. 3, p. 33–37, 2015.

OLIVEIRA, M. C. *Visualização de volumes em morfometria de estruturas cerebrais a partir de imagens de ressonância magnética nuclear*. Tese (Doutorado) — Universidade de São Paulo (USP), 2002.

PEIXOTO; CÁMARA-CHÁVEZ; MENOTTI. *Brazilian License Plate Character Recognition using Deep Learning*. 2015.

- PINAYA, W. H. L. et al. Using deep belief network modelling to characterize differences in brain morphometry in schizophrenia. *Scientific Reports*, 2016. Disponível em: <<http://dx.doi.org/10.1038/srep38897>>.
- RAVÌ, D. et al. Deep learning for health informatics. *IEEE Journal of Biomedical and Health Informatics*, v. 21, p. 4–21, 2017. Disponível em: <<http://dx.doi.org/10.1109/JBHI.2016.2636665>>.
- RAVI, D. et al. Deep learning for health informatics. *IEEE Journal of Biomedical and Health Informatics*, v. 21, n. 1, p. 4–21, 2017. Disponível em: <<http://dx.doi.org/10.1109/JBHI.2016.2636665>>.
- REEVES, A. P.; XIE, Y.; JIRAPATNAKUL, A. Automated pulmonary nodule ct image characterization in lung cancer screening. *International Journal of Computer Assisted Radiology and Surgery*, v. 11, n. 1, p. 73–88, 2015. ISSN 1861-6429.
- RUDER, S. An overview of gradient descent optimization algorithms. *CoRR*, abs/1609.04747, 2016.
- SCHMIDHUBER, J. Deep Learning in Neural Networks: An Overview. *Neural Networks*, v. 61, p. 85–117, 2015. Disponível em: <<http://dx.doi.org/10.1016/j.neunet.2014.09.003>>.
- SHEN, W. et al. Multi-crop Convolutional Neural Networks for Lung Nodule Malignancy Suspiciousness Classification. *Pattern Recognition*, v. 61, p. 663–673, 2017. Disponível em: <<http://dblp.uni-trier.de/db/journals/pr/pr61.html#ShenZYYDYZT17>>.
- SHIMADA, Y.; AL et. Extratumoral vascular invasion is a significant prognostic indicator and a predicting factor of distant metastasis in non-small cell lung cancer. *J Thorac Oncol*, v. 5, p. 970–975, 2010.
- SHIN, H. et al. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *CoRR*, abs/1602.03409, 2016. Disponível em: <<http://arxiv.org/abs/1602.03409>>.
- SILVA, A. C.; CARVALHO, P. C. P.; GATTASS, M. Diagnosis of lung nodule using semivariogram and geometric measures in computerized tomography images. *Computer Methods and Programs in Biomedicine*, v. 79, p. 31–38, 2005.
- SILVA, C. I. S. et al. Consenso brasileiro ilustrado sobre a terminologia dos descritores e padrões fundamentais da TC de tórax. *J. Bras. Pneumol*, v. 36, n. 1, p. 99–123, 2010. Disponível em: <<http://dx.doi.org/10.1590/S1806-37132010000100016>>.
- SILVA, G.; AL et. Taxonomic indexes for differentiating malignancy of lung nodules on ct images. *Research on Biomedical Engineering, SciELO Brasil*, v. 32, n. 3, p. 263–272, 2016.
- SILVA, G. L. F. da; PAIVA, A. C. de; SILVA, A. C. *Diagnóstico de nódulos pulmonares em imagens de tomografia computadorizada usando redes neurais convolucionais evolutivas*. Dissertação (Mestrado) — University of Maranhão - (UFMA), 2016. Disponível em: <<https://tedebc.ufma.br/jspui/handle/tede/1534>>.

- SILVA, G. L. F. da; PAIVA, A. C. de; SILVA, A. C. Lung Nodules Diagnosis Based on Evolutionary Convolutional Neural Network. *Multimedia Tools and Applications*, v. 76, n. 18, p. 19039–19055, 2017. ISSN 1573-7721. Disponível em: <<https://doi.org/10.1007/s11042-017-4480-9>>.
- SONG, Q. et al. Using Deep Learning for Classification of Lung Nodules on Computed Tomography Images. In: *Journal of healthcare engineering*. [s.n.], 2017. v. 2017, p. 1–7. Disponível em: <<http://dx.doi.org/10.1155/2017/8314740>>.
- SONODA, S.; MURATA, N. Neural network with unbounded activations is universal approximator. *CoRR*, abs/1505.03654, 2015. Disponível em: <<http://arxiv.org/abs/1505.03654>>.
- STRAUCH, C. *NoSQL databases*. [S.l.]: Stuttgart Media University, 2011.
- TAMMEMAGI, M.; LAM, S. Screening for Lung Cancer Using Low Dose Computed Tomography. *BMJ*, BMJ Publishing Group Ltd, v. 348, 2014. Disponível em: <<https://www.bmj.com/content/348/bmj.g2253>>.
- TIWARI, S. *Professional NoSQL*. [S.l.]: John Wiley and Sons, Inc, 2011.
- WANG, J. et al. Prediction of malignant and benign of lung tumor using a quantitative radiomic method. *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) - Orlando , United States*, IEEE, v. 34, p. 1272–1275, 2016. Disponível em: <<http://dx.doi.org/10.1109/EMBC.2016.7590938>>.
- WORMANNS, D.; HAMER, O. W. Glossary of Terms for Thoracic Imaging–German Version of the Fleischner Society Recommendations. *RoFo : Fortschritte auf dem Gebiete der Rontgenstrahlen und der Nuklearmedizin*, v. 187 8, p. 638–61, 2015.
- YU, D. et al. Feature Learning in Deep Neural Networks - A Study on Speech Recognition Tasks. *CoRR*, abs/1301.3605, 2013. Disponível em: <<https://dblp.org/rec/bib/journals/corr/abs-1301-3605>>.
- ZAPPA, C.; MOUSA, S. Non-small cell lung cancer: current treatment and future advances. *Translational Lung Cancer Research*, v. 5, n. 3, 2016. ISSN 2226-4477. Disponível em: <<http://tlcr.amegroups.com/article/view/8139>>.