



Universidade do Estado do Rio de Janeiro  
Centro de Tecnologia e Ciências  
Faculdade de Engenharia

Cesar Affonso de Pinho Pinheiro

**Detecção e Classificação de Nódulos Pulmonares  
Utilizando Transferência de Aprendizagem Profunda  
e Inteligência de Enxame**

Rio de Janeiro  
2019

Cesar Affonso de Pinho Pinheiro

**Detecção e Classificação de Nódulos Pulmonares  
Utilizando Transferência de Aprendizagem Profunda  
e Inteligência de Enxame**



Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Engenharia Eletrônica, da Universidade do Estado do Rio de Janeiro. Área de concentração: Sistemas Inteligentes e Automação.

Orientadora: Prof.<sup>a</sup> Dr.<sup>a</sup> Nadia Nedjah

Orientadora: Prof.<sup>a</sup> Dr.<sup>a</sup> Luiza de Macedo Mourelle

Rio de Janeiro  
2019

Cesar Affonso de Pinho Pinheiro

# Detecção e Classificação de Nódulos Pulmonares Utilizando Transferência de Aprendizagem Profunda e Inteligência de Enxame

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Engenharia Eletrônica, da Universidade do Estado do Rio de Janeiro. Área de concentração: Sistemas Inteligentes e Automação.

Aprovado em: 06 de junho de 2019.

Banca Examinadora:

---

Prof.<sup>a</sup> Dr.<sup>a</sup> Nadia Nedjah (Orientadora)  
Faculdade de Engenharia, UERJ

---

Prof.<sup>a</sup> Dr.<sup>a</sup> Luiza de Macedo Mourelle (Orientadora)  
Faculdade de Engenharia, UERJ

---

Prof. Dr. José Manoel de Seixas  
Programa de Engenharia Elétrica, COPPE/UFRJ

---

Prof.<sup>a</sup> Dr.<sup>a</sup> Vera Maria Benjamim Werneck  
Instituto de Matemática e Estatística, UERJ

Rio de Janeiro  
2019

## DEDICATÓRIA

Dedico este trabalho a meus pais Luciana e Ricardo, que me criaram com todo o amor possível e não mediram esforços para investir em minha educação. Dedico também à minha futura esposa Letícia, por todo o suporte, carinho e companheirismo. Vocês são meus exemplos e minha motivação!

## AGRADECIMENTOS

Agradeço às professoras Nadia Nedjah e Luiza de Macedo Mourelle pela orientação, pelos ensinamentos e pelos incentivos dados ao longo do desenvolvimento deste trabalho.

Agradeço aos professores Jorge Luís Machado do Amaral e José Franco Machado do Amaral pelos ensinamentos que recebi durante o curso.

Agradeço aos meus colegas de mestrado Victor, Ramon e Alexandre pela troca de conhecimentos, pelo companheirismo e pelos momentos de descontração.

Agradeço à FAPERJ e ao PEL-UERJ pelos recursos investidos nesta pesquisa.

Agradeço à minha mãe, minha futura esposa Letícia e à minha sogra Daniela por todo o carinho e suporte dado a mim durante esta jornada.

Não existe um caminho para a felicidade. A felicidade é o caminho.

Mahatma Gandhi

## RESUMO

PINHEIRO, Cesar Affonso de Pinho. *Detecção e Classificação de Nódulos Pulmonares Utilizando Aprendizagem Profunda e Inteligência de Enxame*. 2019. 98f. Dissertação (Mestrado em Engenharia Eletrônica) – Faculdade de Engenharia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2019.

O diagnóstico do câncer é geralmente uma tarefa árdua para a medicina, especialmente quando se trata do câncer pulmonar, um dos tipos mais letais e difíceis de tratar dessa terrível doença. A detecção tradicional do câncer de pulmão começa com uma inspeção visual em exames de tomografia computadorizada, feita por médicos especialistas. Por esse motivo, a detecção de nódulos pulmonares cancerígenos em estágios iniciais é um problema muito mais difícil. No entanto, deve-se investir nesta detecção precoce, pois ela aumenta drasticamente as chances de cura. Para ajudar a melhorar a detecção do câncer e as taxas de sobrevivência, engenheiros e cientistas têm desenvolvido sistemas de diagnóstico assistido por computador, os chamados “Sistemas CAD”. Este trabalho utiliza técnicas de inteligência computacional para sugerir uma nova abordagem à solução do problema de detecção de nódulos cancerígenos pulmonares em exames de tomografia computadorizada. A tecnologia aplicada consiste em usar Aprendizagem Profunda e Inteligência de Enxame para desenvolver diferentes modelos de classificação de nódulos. No total, 14 algoritmos de Inteligência de Enxame e uma Rede Neural Convolutiva usada para segmentação de imagens biomédicas foram usados para classificar nódulos pulmonares na base de dados *Lung Image Database Consortium and Image Database Resource Initiative* (LIDC-IDRI). O principal objetivo deste trabalho é utilizar os algoritmos de enxame para treinar modelos de redes neurais convolucionais por transferência de aprendizado e, com isso, verificar se este método é mais eficiente que os algoritmos de treinamento tradicionais, como o *Backpropagation* e Descida do gradiente. Os modelos desenvolvidos neste trabalho conseguiram alcançar desempenhos elevados, produzindo resultados compatíveis com os atuais modelos estado-da-arte em aplicações de detecção de câncer de pulmão. Com os experimentos realizados neste trabalho, foi possível verificar a real eficácia do uso de algoritmos de inteligência de enxame para treinar modelos de transferência de aprendizagem profunda e afirmar a superioridade destas técnicas sobre os modelos treinados com *backpropagation* para este tipo de aplicação. Dos 14 algoritmos de enxame testados, 7 obtiveram melhor desempenho que o *backpropagation*, alcançando até 93,78% em acurácia, 93,59% em precisão, 93,04% em sensibilidade, 98,56% em especificidade e operando até 25% mais rápido.

Palavras-chave: Câncer de Pulmão; Aprendizagem Profunda; Inteligência de Enxame; Sistemas de Diagnóstico Auxiliado por Computador.

## ABSTRACT

PINHEIRO, Cesar Affonso de Pinho. *Detection and Classification of Pulmonary Nodules using Deep Learning and Swarm Intelligence*. 2019. 98f. Dissertação (Mestrado em Engenharia Eletrônica) – Faculdade de Engenharia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2019.

Cancer diagnosis is usually an arduous task for medicine, specially when it comes to pulmonary cancer, one of the most deadly and hard to treat types of that terrible disease. Traditional lung cancer detection starts with a visual inspection, done by medical specialists on tomography scans. For this reason, detecting pulmonary cancerous nodules in early stages is a much harder problem. However, it is definitely worth trying, as it drastically increases surviving chances. To help improving cancer detection and surviving rates, engineers and scientist have been developing computer-aided diagnosis systems, also called CAD systems. This work uses computational intelligence techniques to suggest a new approach towards solving the problem of detecting pulmonary carcinogenic nodules in computed tomography scans. The technology applied consists in using Deep Learning and Swarm Intelligence to develop different nodule classification models. In total, 14 different Swarm Intelligence algorithms and a Convolutional Neural Networks for biomedical image segmentation called “U-Net” were used to find cancerous pulmonary nodules in the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) databases. The main goal of this work is to utilize Swarm algorithms to train Convolutional Neural Networks models via Transfer Learning, and check whether this method is more efficient than the regular training algorithms, such as Backpropagation and Gradient Descent. The models developed in this work managed to reach significantly high performances, producing results compatible with the current state-of-art models for lung cancer detection applications. With the experiments conducted in this work, it was possible to verify the real effectiveness of using Swarm Intelligence algorithms to train Deep Learning models and to state their superiority over the Backpropagation models for this application. From the 14 tested swarm algorithms, 7 performed better than backpropagation, reaching up to 93.78% in accuracy, 93.59% in precision, 93.04% in sensitivity, 98.56% specificity and operating up to 25% faster.

Keywords: Lung Cancer; Deep Learning; Swarm Intelligence; Computer Aided Diagnosis Systems.

## LISTA DE FIGURAS

1	Principais causas de morte no século XX . . . . .	17
2	Mortes por câncer no estado de São Paulo em centenas de milhares . . . . .	18
3	Mortes por câncer em todo o mundo nas últimas 2 décadas . . . . .	18
4	Tomografia pulmonar. . . . .	19
5	Imagem DICOM de uma tomografia de pulmão . . . . .	22
6	Partículas buscando um resultado ótimo . . . . .	37
7	Exemplos de topologias . . . . .	37
8	Uso e complexidade dos principais algoritmos de aprendizagem de máquinas	50
9	Experimento de Hubel e Wiesel . . . . .	52
10	Ilustração do aprendizado por transferência . . . . .	54
11	Arquitetura da U-Net . . . . .	55
12	Resumo do fluxo de processos desenvolvidos na aplicação. . . . .	58
13	Imagem DICOM de exame de tomografia computadorizada de pulmão. . .	59
14	Corte feito ao redor das coordenadas apontadas pelos especialistas como promissoras. . . . .	61
15	Transformações de <i>data augmentation</i> . . . . .	62
16	Operação de <i>Dropout</i> . . . . .	63
17	Treinamento de redes neurais com algoritmos de enxame . . . . .	66
18	Desempenho dos pares de hiper-parâmetros analisado . . . . .	70
19	Quantidade de iterações de treinamento. . . . .	72
20	Acurácia média para todos os algoritmos . . . . .	76
21	Precisão média para todos os algoritmos . . . . .	76
22	Sensibilidade média para todos os algoritmos . . . . .	77
23	Especificidade média para todos os algoritmos . . . . .	78
24	Tempo médio de treinamento. . . . .	78
25	Acurácia média . . . . .	79
26	Precisão média. . . . .	80
27	Sensibilidade média . . . . .	80
28	Especificidade média . . . . .	81
29	Tempo médio de treinamento. . . . .	81
30	Comparação de desempenho entre algoritmos e especialistas . . . . .	82

## LISTA DE TABELAS

1	Valores <i>Hounsfield</i> de alguns tecidos encontrados em tomografias. . . . .	60
2	Teste-T de <i>Student</i> dos casos $C_9$ , $C_{10}$ e $C_{11}$ para 95% de confiança. . . . .	71
3	Parâmetros do BFO. . . . .	72
4	Parâmetros do BOA . . . . .	73
5	Parâmetros do FOA. . . . .	73
6	Parâmetros do FA . . . . .	73
7	Parâmetros do WSA . . . . .	74
8	Parâmetros do CSA. . . . .	74
9	Parâmetros do HS . . . . .	74
10	Parâmetros do PSO. . . . .	74
11	Parâmetros do SSA . . . . .	74
12	Parâmetros do CSO. . . . .	75
13	Teste-T de <i>Student</i> para a acurácia. . . . .	83
14	Teste-T de <i>Student</i> para a precisão. . . . .	83
15	Teste-T de <i>Student</i> para a sensibilidade. . . . .	83
16	Teste-T de <i>Student</i> para a especificidade. . . . .	84
17	Impacto dos hiper-parâmetros da camada totalmente conectada da rede nas métricas de desempenho . . . . .	96
18	Acurácia média para as funções de ativação e matrizes de <i>max-pooling</i> . . .	97
19	Precisão média para as funções de ativação e matrizes de <i>max-pooling</i> . . .	97
20	Sensibilidade média para as funções de ativação e matrizes de <i>max-pooling</i>	98
21	Especificidade média para as funções de ativação e matrizes de <i>max-pooling</i>	98

## LISTA DE SIGLAS

ABC	<i>Artificial Bee Colony</i>
BFO	<i>Bacterial Foraging Optimization</i>
BKP	<i>Backpropagation</i>
BOA	<i>Bat Optimization Algorithm</i>
CAD	<i>Computer-aided Diagnosis System</i>
ChSO	<i>Chicken Swarm Optimization</i>
CSA	<i>Cuckoo Search Algorithm</i>
CSO	<i>Cat Swarm Optimization</i>
DICOM	<i>Digital Imaging and Communications in Medicine</i>
FA	<i>Firefly Algorithm</i>
FN	<i>False negative</i>
FOA	<i>Firework Optimization Algorithm</i>
FP	<i>False positive</i>
GSA	<i>Gravitational Search Algorithm</i>
GWO	<i>Gray Wolf Optimization</i>
HS	<i>Harmony Search</i>
HU	<i>Hounsfield Units</i>
KDSB17	<i>Kaggle Data Science Bowl 2017</i>
LIDC-IDRI	<i>Lung Image Database Consortium and Image Database Resource Initiative</i>
LUNA16	<i>Lung Nodule Analysis 2016</i>
PSO	<i>Particle Swarm Optimization</i>
ReLU	<i>Rectified Linear Units</i>
SSA	<i>Social Spider Algorithm</i>
SVM	<i>Support Vector Machines</i>
TN	<i>True negative</i>
TP	<i>True positive</i>

# SUMÁRIO

<b>INTRODUÇÃO</b> .....	<b>13</b>
<b>1 CÂNCER DE PULMÃO: ESTATÍSTICAS E DIAGNÓSTICO</b> .	<b>16</b>
1.1 Estatísticas a redor da doença. ....	16
1.2 Métodos tradicionais de diagnóstico .....	19
1.3 Imagens médicas. ....	21
1.4 Sistemas CAD .....	22
1.5 Motivação do trabalho .....	23
1.6 Considerações finais .....	24
<b>2 TRABALHOS RELACIONADOS</b> .....	<b>26</b>
2.1 Aprendizagem profunda para classificação. ....	26
2.2 Classificação com rede U-Net .....	27
2.3 Diagnósticos por redes neurais e PSO .....	28
2.4 Detecção de câncer com redes neurais. ....	29
2.5 Utilização de outros métodos .....	30
2.5.1 Detecção por crescimento de regiões. ....	31
2.5.2 Detecção por agrupamento e classificação .....	31
2.5.3 Detecção por classificação Bayesiana .....	32
2.6 Técnicas para tratamento de imagens em sistemas CAD .....	32
2.6.1 Fusão de características .....	32
2.6.2 Remoção de ruídos e aguçamento de imagens. ....	32
2.7 Considerações finais .....	33
<b>3 INTELIGÊNCIA DE ENXAME</b> .....	<b>34</b>
3.1 Visão geral .....	34
3.2 Algoritmos de inteligência de enxame .....	35
3.2.1 Otimização por enxame de partículas .....	38
3.2.2 Colônia artificial de abelhas .....	38
3.2.3 Busca otimizada com morcegos .....	40
3.2.4 Otimização bacteriana .....	40
3.2.5 Busca por alcateia .....	41
3.2.6 Algoritmo de vaga-lumes .....	41
3.2.7 Busca otimizada dos cucos .....	42
3.2.8 Enxame de baleias .....	42
3.2.9 Busca por fogos de artifício .....	43
3.2.10 Otimização por enxame de galinhas .....	43
3.2.11 Colônia de aranhas .....	44
3.2.12 Otimização com gatos .....	45

# SUMÁRIO

---

3.2.13	<u>Busca pela harmonia</u> . . . . .	45
3.2.14	<u>Algoritmo de busca gravitacional</u> . . . . .	46
<b>3.3</b>	<b>Considerações finais</b> . . . . .	<b>46</b>
<b>4</b>	<b>APRENDIZAGEM PROFUNDA</b> . . . . .	<b>48</b>
4.1	Aprendizagem de máquinas . . . . .	48
4.2	Redes neurais de alta complexidade . . . . .	50
4.2.1	<u>Transferência de conhecimento</u> . . . . .	53
4.2.2	<u>Rede neural <i>U-Net</i></u> . . . . .	54
4.3	Considerações finais . . . . .	55
<b>5</b>	<b>ASPECTOS DE IMPLEMENTAÇÃO</b> . . . . .	<b>57</b>
5.1	A metodologia . . . . .	57
5.2	Obtenção das imagens de tomografia . . . . .	58
5.3	Pré-processamento dos dados . . . . .	59
5.4	Construção do modelo . . . . .	61
5.5	Treinamento dos modelos . . . . .	64
5.5.1	<u>Treinamento com <i>backpropagation</i></u> . . . . .	64
5.5.2	<u>Treinamento utilizando algoritmos de enxame</u> . . . . .	64
5.6	Considerações finais . . . . .	66
<b>6</b>	<b>RESULTADOS DE DESEMPENHO</b> . . . . .	<b>68</b>
6.1	Métricas de avaliação . . . . .	68
6.2	Obtenção de hiper-parâmetros da rede . . . . .	69
6.3	Resultados utilizando técnicas de enxame . . . . .	71
6.3.1	<u>Parâmetros das técnicas de enxame</u> . . . . .	71
6.3.2	<u>Testes e resultados</u> . . . . .	75
6.4	Análises de desempenho . . . . .	78
6.5	Validação estatística . . . . .	82
6.6	Considerações finais . . . . .	84
<b>7</b>	<b>CONCLUSÕES E TRABALHOS FUTUROS</b> . . . . .	<b>86</b>
7.1	Conclusões . . . . .	86
7.2	Trabalhos Futuros . . . . .	89
	<b>REFERÊNCIAS</b> . . . . .	<b>91</b>
	<b>APÊNDICE A – Resultados dos experimentos</b> . . . . .	<b>96</b>

# INTRODUÇÃO

A luta contra o câncer é um desafio constante para a medicina moderna. As dificuldades estão presentes em todos os aspectos desta doença, pois, em muitos dos casos não se sabe o motivo de sua origem, não existem curas efetivas, o estágio da doença avança rapidamente, o diagnóstico é dificultado e o tratamento caro e agressivo. Ainda sim, o problema é mais grave quando analisado em maior escala, e estudos como o apresentado em (HANNAH et al., 2018) mostram um panorama assustador desta doença.

Com o avanço da medicina e das tecnologias de segurança, as principais causas de morte ao redor do mundo vêm perdendo força ao longo dos anos. Alarmantemente, este comportamento não vem sendo observado no câncer, pois sua ocorrência vem aumentando e as taxas de sobrevivência ainda são baixas.

Tendência genética, exposição a substâncias tóxicas e até o consumo de alguns alimentos do dia-a-dia são apontados como possíveis causas dos diversos tipos de câncer. Desta forma torna-se cada vez mais difícil conter a aparição de novos casos desta doença.

O estágio do câncer é medido principalmente pelo grau de espalhamento dos tumores pelo corpo. Uma vez que um tumor atinja a corrente sanguínea e consiga se espalhar para outros órgãos, as chances de sobrevivência do paciente são reduzidas drasticamente, este processo é chamado de metástase.

Esses fatos reforçam a necessidade de se diagnosticar e combater a doença em seus estágios iniciais, onde as chances de cura são significativamente altas. No entanto, em estagio inicial, o câncer não costuma apresentar sintomas e, com isso, dificilmente são levantadas as hipóteses de análise para diagnóstico. Mesmo que apresente a doença manifeste sintomas no início, o diagnóstico de pequenos nódulos é bastante dificultado, pois geralmente os exames dependem de inspeções visuais ou análises laboratoriais de tecido humano. Devido à essa necessidade de detecção dos nódulos por inspeção visual, aliar exames médicos às tecnologias de visão computacional e reconhecimento de padrões são alternativas ótimas para a solução do problema.

Com as atuais tecnologias de inteligência artificial focadas em visão computacional, novos modelos são capazes de apresentar desempenhos sobre-humanos em reconhecimento de padrões. Sendo assim, utilizar sistemas como esses para a detecção e classificação de nódulos cancerígenos é uma estratégia bastante promissora.

O uso desta técnica é um aliado ainda mais forte quando se trata do câncer de pulmão, pois além de ser um dos tipos mais mortíferos da doença, seu diagnóstico é um dos mais difíceis, devido à dificuldade de acesso ao órgão e a sua estrutura vascular altamente ramificada. Dentre as principais técnicas utilizadas para aplicações de visão computacional, destaca-se o desenvolvimento das redes neurais convolucionais, que são modelos usados para simular o procedimento aplicado pelo cérebro humano para reconhecer e aprender novos padrões visuais.

Estas redes, são parte de uma ramificação da área de inteligência artificial, chamada de aprendizagem profunda. Neste ramo, são desenvolvidos modelos de alta complexidade, que quando treinados com grandes quantidades de dados, conseguem adquirir uma inteligência avançada suficiente para reconhecer quase qualquer tipo de padrão.

Para o treinamento de uma rede neural como esta, são necessários não só muitas imagens mas um bom algoritmo de otimização, para que exista uma boa busca pelos melhores parâmetros do modelo. Como em redes neurais de menor complexidade, geralmente são utilizados algoritmos como o *backpropagation* e a descida do gradiente para o treinamento de redes neurais convolucionais. No entanto, outros ramos da inteligência artificiais fornecem algoritmos de otimização ainda mais eficientes, que podem ser combinados com estes modelos de reconhecimento de padrões para produzir melhores resultados.

A inteligência de enxame é um dos ramos da inteligência artificial focado em aplicações de otimização de processos. Estas técnicas simulam o comportamento de colônias de animais ou de fenômenos da natureza para desenvolver estratégias inteligentes de buscas por configurações ótimas para a solução de um problema.

Os principais usos da inteligência de enxame são aplicados à robótica, com robôs de sensoriamento ou *drones*, e ao encontro de valores ótimos de funções de engenharia, como em estudos de aerodinâmica e projetos de peças mecânicas.

Este trabalho traz uma nova abordagem para a resolução do problema de detecção e classificação de nódulos cancerígenos pulmonares, utilizando aprendizagem profunda e inteligência de enxame.

Como objetivo, este estudo visa treinar modelos de rede neural convolucional por transferência de aprendizado, utilizando 14 algoritmos de inteligência de enxame e, com isso, comprovar a superioridade do uso destes algoritmos em relação aos métodos tradicionais de treinamento, que utilizam o *backpropagation*.

Como resultado, esta pesquisa conseguiu atingir o objetivo proposto, onde 7 dos 14 algoritmos de enxame testados obtiveram desempenhos superiores ao modelo treinado com *backpropagation*. Além disso, pode-se atingir performances compatíveis com os modelos estado-da-arte, com acurácias de até 93,71%, precisões de até 93,53%, sensibilidades de até 92,96% e especificidades de até 98,52%. Além da alta performance, os modelos treinados com algoritmos de enxame operaram até 25% mais rápido quando comparados com o modelo treinado com *backpropagation*.

Este trabalho deu origem a 2 artigos, um publicado na conferência *International Conference on Computer Science 2019* (PINHEIRO; MOURELLE; NEDJAH, 2019) e outro no periódico *Multimedia Tools and Applications* (PINHEIRO; NEDJAH; MOURELLE, 2019).

Nesta dissertação, as principais características e estatísticas ao redor do problema são apresentados no Capítulo 1. No Capítulo 2 são descritos alguns dos trabalhos relacionados à esta pesquisa, onde outros métodos e resultados são apresentados para aprofundar o contexto de solução do problema em análise. Em seguida, no Capítulo 3, a inteligência de enxame é descrita e cada algoritmo utilizado no trabalho é apresentado. Ainda na fundamentação teórica deste trabalho, o Capítulo 4 apresenta os principais aspectos em torno da aprendizagem profunda, descrevendo os métodos mais utilizados em aplicações como esta.

O Capítulo 5, é dedicado à descrição detalhada de todos os procedimentos envolvidos no desenvolvimento e conclusão deste trabalho. Nele são apresentados os passos percorridos desde a pesquisa dos dados de treinamento à obtenção dos resultados finais. Os resultados obtidos após os experimentos são apresentados no Capítulo 6, onde cada configuração testada é explicitada e suas consequências analisadas. Por fim, o Capítulo 7 apresenta as conclusões obtidas com o trabalho desenvolvido. Este capítulo também sugere possíveis modificações e melhorias para trabalhos futuros em aplicações como esta.

# Capítulo 1

## CÂNCER DE PULMÃO: ESTATÍSTICAS E DIAGNÓSTICO

**N**ESTE capítulo, são apresentadas as principais características e estatísticas ao redor do câncer de pulmão, nele são expostos dados de organizações e pesquisadores de diferentes regiões do mundo, que atestam a gravidade desta doença e evidenciam a necessidade constante de melhorias nas técnicas de diagnóstico. Além de apresentar estes dados, este capítulo traz também explicações sobre as principais técnicas de diagnóstico da doença e uma descrição das motivações que levaram à decisão do desenvolvimento deste trabalho, explicando as relações entre a medicina e a engenharia da computação no combate ao câncer.

### 1.1 Estatísticas a redor da doença

De acordo com o estudo apresentado em (HANNAH et al., 2018), o câncer é a segunda maior causa de morte em todo o mundo desde a década de 1940, estando atrás somente das mortes causadas por problemas cardiovasculares. O estudo (HANNAH et al., 2018) contém dados estatísticos que comprovam o poder devastador do câncer, que causou aproximadamente 9 milhões de óbitos em 2016. A Figura 1, retirada do mesmo estudo apresenta o comportamento alarmante do câncer, pois enquanto as doenças mais comuns tiveram sua incidência reduzida com o passar do tempo e o avanço da medicina, o câncer não parou de crescer.

Dos tipos conhecidos de câncer, o câncer de pulmão ocupa o topo da lista de vítimas fatais e seus números crescem alarmantemente a cada ano, devido à alta incidência do tabagismo e à exposição à poluição do ar em grandes metrópoles.

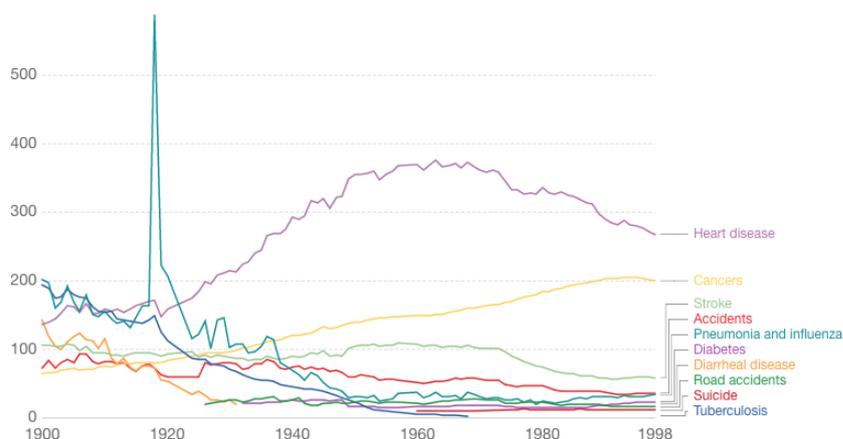


Figura 1: Principais causas de morte no século XX [Fonte: (HANNAH et al., 2018)]

O *National Cancer Institute*, (NCI, 2018), estima que 1,73 milhões de casos de câncer foram diagnosticados em 2018 nos Estados Unidos. No ano de 2017, o instituto publicou outra lista de casos e mortes causadas por câncer. Nesta lista, o câncer de pulmão foi o segundo mais comum entre a população, com 222.500 casos detectados, estando pouco atrás do câncer de mama com 252.710 casos. Nos casos de morte, o câncer de pulmão atingiu o primeiro lugar com 155.870 óbitos, cerca de três vezes mais do que o segundo lugar, o câncer de pâncreas. Além da alta incidência, o câncer de pulmão é também o segundo mais mortífero dentre os listados, com uma taxa de mortalidade de 71% nos Estados Unidos e potencialmente ainda maior em países menos desenvolvidos. Estimou-se também que um total de 12 bilhões de dólares são gastos com tratamentos relacionados ao câncer de pulmão a cada ano, e que são esperados 23,6 milhões de casos de câncer no mundo em 2030. As Figuras 2 e 3 apresentam respectivamente, panoramas das mortes por câncer no estado de São Paulo e no mundo, comprovando a importância do combate ao câncer de pulmão, que se mostra como o mais mortífero tipo da doença.

Estas altas taxas de mortalidade devem-se à própria natureza dos pulmões, que devido à presença dos brônquios e da alta carga vascular, formam uma região altamente complexa, que facilita o espalhamento de tumores, dificultam intervenções médicas e atrapalham os métodos de diagnóstico.

O estágio de um câncer de pulmão é medido por seu grau de espalhamento no organismo humano, processo chamado de metástase. A grande busca por técnicas de diagnóstico de câncer de pulmão se concentra em achar métodos que consigam detectar os primeiros estágios da doença, onde o tratamento oferece chances significativamente maiores

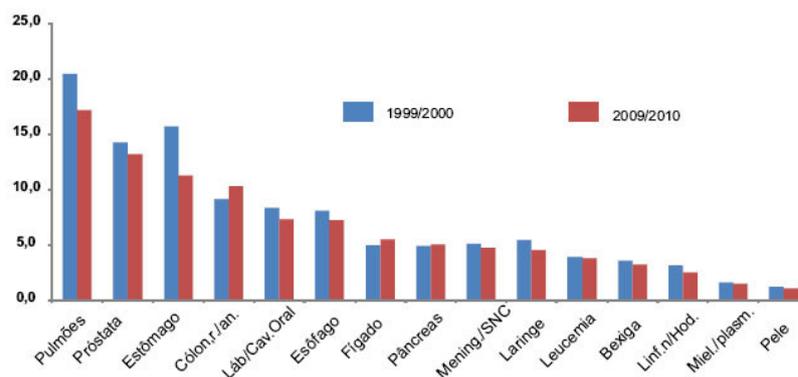


Figura 2: Mortes por câncer no estado de São Paulo em centenas de milhares [Fonte: FSEADE Censo IBGE 00/10, (MENDES; CECILIO, 2012)]

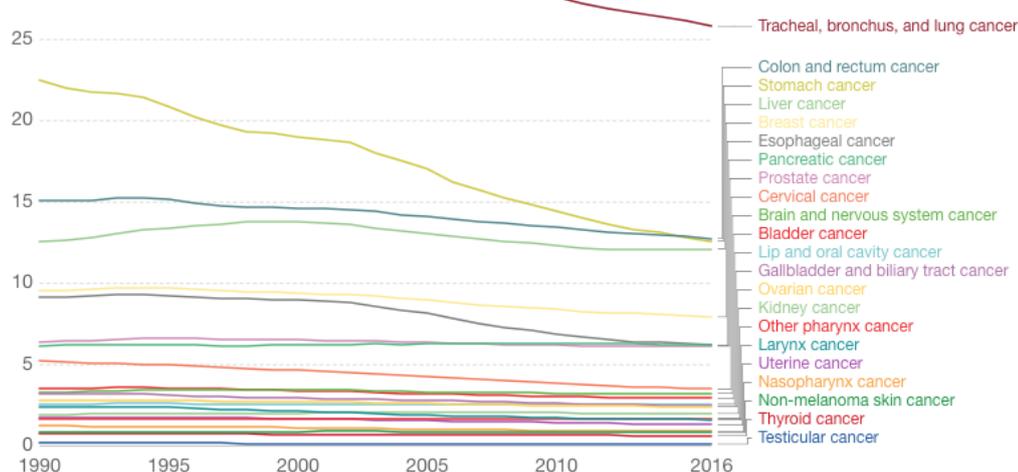


Figura 3: Mortes por câncer em todo o mundo nas últimas 2 décadas [Fonte: (HANNAH et al., 2018)]

de cura, no entanto, o diagnóstico precoce é bastante dificultado, devido à ausência de sintomas. Atualmente, os métodos mais utilizados para o diagnóstico do câncer de pulmão são a biópsia, exame que retira parte de um tecido humano para análise em laboratório, e as análises de tomografias computadorizadas, como mostra a Figura 4, realizadas por profissionais da área de radiologia ou por sistemas automatizados de visão computacional.

Estes métodos apresentam características bastante distintas. A biópsia requer uma logística complexa, que envolve o procedimento de intervenção médica para a retirada do tecido e o envio do material para laboratório, para ser analisado por especialistas. Por outro lado, o diagnóstico auxiliado por computador mostra-se mais eficiente e simplificado. Com este, basta que o paciente seja submetido a um exame de tomografia

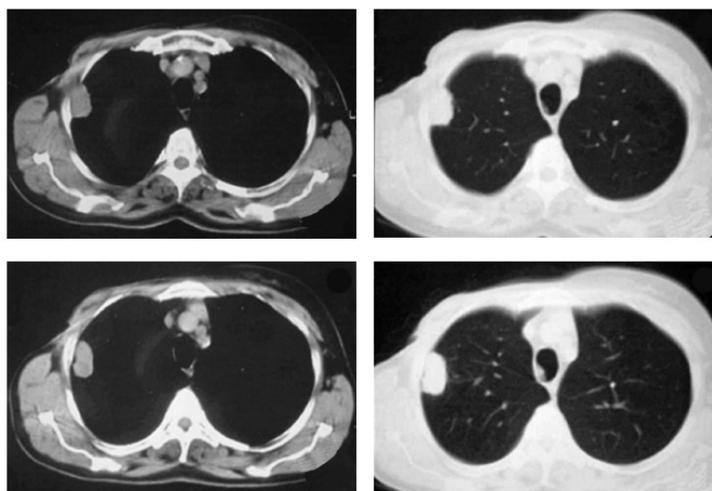


Figura 4: Tomografia pulmonar [Fonte: (SILVA et al., 2007)]

computadorizada, similar a um exame de raio x, sem qualquer incisão ou cirurgia. Após o exame as imagens geradas são analisadas por um sistema de inteligência computacional treinado para reconhecer certos tipos de anomalias e assim, classificar o paciente como enfermo ou não.

Tendo em vista esta facilidade traga pelos sistemas de diagnóstico computadorizado, hoje se observa um grande movimento de desenvolvimentos de sistemas que unificam o conhecimento médico com a tecnologia da ciência da computação. Desta forma, é possível automatizar e otimizar a obtenção de diagnósticos para que se possa detectar mais enfermidades em menos tempo e com altas taxas de precisão, garantindo aumentos expressivos nas taxas de sobrevivência de doenças como o câncer.

## 1.2 Métodos tradicionais de diagnóstico

O câncer é uma patologia que pode atingir diversas partes do organismo humano e em boa parte dos casos leva o indivíduo ao óbito. O grande problema na detecção deste mal é o fato da ausência de sintomas e sinais que indiquem uma suspeita da doença em estágios iniciais. Apesar de grandes avanços na medicina diagnóstica, ainda hoje não se sabe a real origem da doença, tão pouco se sabe uma maneira de curá-la com alta efetividade. O Instituto do Câncer do Brasil, INCA, estima que cerca de setenta por cento dos diagnósticos de câncer são detectados por médicos cuja especialidade não é a oncologia. Em geral, o paciente com câncer procura o médico sem saber a natureza de sua doença e por isso existe a necessidade de métodos de diagnóstico de câncer cada vez

mais eficazes em detectar a doença em estágios iniciais, pois uma vez que os nódulos cancerígenos consigam se espalhar pela corrente sanguínea do paciente, suas chances de sobrevivência são reduzidas drasticamente.

A detecção do câncer é sempre um assunto de alta complexidade, pois em grande maioria dos casos, envolve a aplicação de vastas rotinas médicas e exames custosos e delicados. Para canceres como o câncer de mama, onde os tumores em geral estão localizados em uma região de mais fácil detecção, é possível detectar nódulos apenas com exames de rotina sem nenhum tipo de intervenção cirúrgica ou radiográfica. No entanto, quando se trata de órgãos como os pulmões, não existe maneira de se realizar inspeções manuais para a detecção de nódulos. Sendo assim, os procedimentos para a detecção deste tipo de câncer tendem a envolver radiografias, incisões e procedimentos cirúrgicos.

A broncoscopia, ou endoscopia da árvore brônquica é um exame muito utilizado para a detecção de tumores nos brônquios, estruturas ramificadas do trato respiratório. Neste exame o paciente é submetido à uma anestesia e um duto equipado com uma micro câmera é inserido por via oral até o pulmão. Neste exame, é realizada uma inspeção visual em tempo real das condições pulmonares do paciente. A cintilografia é outro método que pode ser utilizado para a detecção de tumores na região pulmonar. Trata-se de um exame de imagem radiológica onde um radiofármaco radioativo é inserido de forma venosa no paciente e a análise de anomalias é feita por um especialista que busca por regiões suspeitas. A análise citológica de fluidos pleurais é outra técnica utilizada para a detecção de câncer de pulmão. Neste exame, o médico deve coletar uma amostra de fluido produzido na pleura do paciente. Este procedimento é similar a uma biópsia, pois é um método intrusivo que requer uma rotina cuidadosa e detalhada. Após a coleta do líquido, este é levado para um laboratório, para uma análise de células. Outro método intrusivo bastante utilizado para a detecção de câncer de pulmão é a biópsia de tecido pulmonar. Neste exame, o paciente é submetido a um procedimento cirúrgico de baixa complexidade onde uma pequena amostra de seu tecido pulmonar é retirado e enviado para uma análise laboratorial. Nesta análise as células serão vasculhadas para a detecção de algum padrão que possa indicar a presença de um câncer de pulmão.

Não obstante aos exames citados, a tomografia computadorizada do tórax é o exame mais utilizado em casos de suspeita de câncer de pulmão, pois trata-se de

um procedimento não intrusivo, que não envolve rotinas médicas complexas e de fácil execução. Neste exame o paciente é submetido a uma série de impulsos radiológicos que, como resultado, compõem uma série de imagens em seções do órgão analisado. No entanto, este método carrega uma dificuldade quando se trata de diagnóstico de tumores. Na tomografia computadorizada o especialista deve analisar todas as imagens e buscar por formas geométricas que despertem suspeitas sobre serem nódulos cancerígenos.

Esta inspeção visual nem sempre é eficaz, pois depende de variáveis como a qualidade da imagem, a experiência do especialista, o estado psicológico do especialista na hora da análise, entre outras. A análise deste tipo de imagem é complexa pois componentes como os vasos sanguíneos ou os próprios brônquios podem se parecer com tumores em imagens seccionadas. Contudo, deve-se reconhecer que, apesar da limitação humana, este exame é bastante otimizado para a aplicação pois não necessita de procedimentos médicos complexos e produz imagens de alta fidelidade ao que se encontra no organismo do paciente. Desta maneira, aliar as vantagens deste exame com a eficiência de uma máquina treinada para detectar nódulos é uma estratégia altamente eficaz. Por isso existem grandes estudos empenhados no desenvolvimento de sistemas capazes de reconhecer padrões em imagens médicas.

### 1.3 Imagens médicas

A engenharia da computacional é uma aliada cada vez mais forte da medicina moderna, pois as ferramentas tecnológicas desenvolvidas permitem exames e estudos cada vez mais precisos e aprofundados. Devido à alta demanda de exames como tomografias, ultrassonografias e raios-x, foi criado pelo Comitê Norte Americano de Radiologia em 1983 um padrão de normas para tratamento, armazenamento e transmissão de imagens médicas num formato eletrônico, instalando assim um novo protocolo de comunicação de imagens digitais em medicina, (*Digital Imaging and Communications in Medicine - DICOM*).

Com o protocolo, foi criada a extensão de arquivos “.dicom”, utilizada em todos os aparelhos de radiologia e tomografia. Esta extensão contém não somente as imagens do exame, como também informações em texto sobre quaisquer anotações relevantes sobre o paciente, exame, horário, equipe médica, etc. Estas informações são chamadas de *metadata* e costumam estar organizadas em arquivos de extensão XML, (*Extensible*

*Markup Language*), um formato de marcação utilizado para estruturar dados de maneira simplificada. A Figura 5 apresenta um exemplo de imagem DICOM de uma tomografia de pulmão.

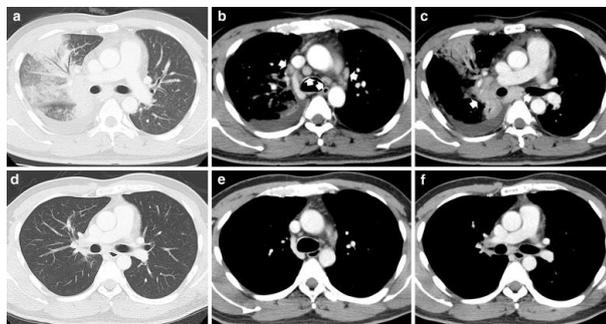


Figura 5: Imagem DICOM de uma tomografia de pulmão [Fonte: (NAITO et al., 2015)]

Durante anos, milhares de centros médicos por todo o mundo vêm coletando imagens médicas como estas, criando bancos de dados utilizados para fomentar a pesquisa e a criação de modelos capazes de produzir diagnósticos cada vez mais precisos. Este é o caso do *Cancer Imaging Archive*, uma organização focada em fornecer gratuitamente centenas de bancos de dados de imagens médicas sobre detecção de inúmeros tipos de cânceres.

Este trabalho utiliza o banco de dados LIDC-IDRI, *Lung Image Database Consortium* como base para o treinamento da rede neural convolucional. Este banco é fornecido pelo *Cancer Imaging Archive* e contém cerca de 250.000 imagens médicas de exames de tomografias computadorizadas de pulmões sadios e cancerígenos. As imagens deste banco de dados foram analisadas por quatro especialistas e o diagnósticos de cada um deles está contido nas *metadatas* das imagens. Foram também utilizados os bancos LUNA16, *Lung Nodule Analysis 2016* e KDSB17, *Kaggle Data Science Bowl 2017*, pois estes foram utilizados em competições de visão computacional e por autores de trabalhos relacionados a este.

## 1.4 Sistemas CAD

Sistemas de diagnóstico auxiliado por computador (*Computer Aided Diagnosis - CAD*) são métodos computacionais que auxiliam profissionais da área médica a realizar diagnósticos de doenças, pois oferecem uma segunda opinião sobre o estado do paciente. Estes sistemas são utilizados majoritariamente em diagnósticos de radiologia, onde um equipamento

fornece imagens de órgãos de um paciente e o diagnóstico é dado pelo especialista da área. É importante ressaltar que sistemas CAD não devem ser utilizados como uma ferramenta de diagnóstico automatizado, e sim como um método de auxílio para um profissional. Este cuidado deve-se pela natureza de aprendizagem dos métodos CAD, que é derivada do conhecimento humano sobre a doença em questão.

Estes sistemas são treinados a partir de grandes quantidades de dados categorizados por especialistas da área médica. Sistemas de diagnóstico automatizado são utilizados na detecção de doenças do sangue por contagem de células, por exemplo, onde o olho humano não possui funcionalidade e a classificação do conteúdo é feita por características físico-químicas pré-determinadas. Sistemas CAD utilizam algoritmos de visão computacional para extrair e reconhecer padrões em imagens médicas digitais. Este processo é conduzido através da aplicação de técnicas de processamento de imagens, onde a imagem em análise é vasculhada e mapeada para a detecção de bordas, linhas, formatos suspeitos, padrões de cores, etc.

Geralmente, antes da extração de atributos, são aplicados filtros de redução de ruído e de aguçamento de bordas às imagens para facilitar o reconhecimento das características citadas. Em sua maioria, algoritmos de sistemas CAD são equipados com redes neurais artificiais capazes de julgar e dar diagnósticos de uma doença baseando-se no mapeamento de características da fase de processamento de imagens. Estas redes são treinadas a partir de bancos de dados de diagnósticos realizados por um ou mais especialistas da área médica, portanto, é necessário um grande volume de imagens para criar um sistema de boa precisão. Para analisar o desempenho de um sistema CAD, geralmente são utilizadas métricas como a acurácia, a precisão e as curvas ROC de cada modelo.

## 1.5 Motivação do trabalho

Este trabalho foi idealizado para ser mais uma contribuição para a comunidade de pesquisadores que buscam o aprimoramento das técnicas de detecção deste mal que carrega números tão expressivos e faz milhões de vítimas todos os anos. Temas como estes incitam a necessidade da união de profissionais das áreas médicas e tecnológicas para o desenvolvimento de métodos capazes de salvar vidas, sendo este um dos principais motivos e inspirações para a inicialização deste trabalho. Técnicas como as desenvolvidas

neste projeto são de extrema ajuda para os especialistas da área médica, pois podem em relativamente pouco tempo, obter acurácias e eficiências superiores às do ser humano. Por estas questões, as ferramentas de diagnóstico computadorizado têm se tornado cada vez mais indispensáveis aos laboratórios médicos.

Além da motivação relacionada à detecção da enfermidade, este projeto tem como objetivo sugerir uma nova abordagem ao problema de detecção de nódulos cancerígenos de pulmão. A utilização de algoritmos de aprendizagem profunda para a detecção de nódulos cancerígenos já é uma realidade, pois graças a popularização de ferramentas de programação, como as bibliotecas de aprendizado de máquinas, muitos desenvolvedores independentes puderam desenvolver seus próprios modelos classificadores de nódulos. Estes eventos são possíveis pois estão disponíveis também grandes bancos de dados de imagens médicas para treino de sistemas de diagnóstico.

No entanto, este trabalho traz uma contribuição que visa sugerir e comparar a utilização de algoritmos de enxame para a otimização de parâmetros de modelos classificadores de nódulos cancerígenos pulmonares. O objetivo é treinar um modelo utilizando diversas técnicas de otimização de enxame, comparar os resultados com os de modelos que não utilizam estas otimizações e eleger o melhor modelo classificador para esta aplicação. Sendo assim, será possível saber a real contribuição deste tipo de otimização de parâmetros em aplicações de detecção de nódulos cancerígenos pulmonares.

## 1.6 Considerações finais

Com o conteúdo deste capítulo é possível ter noção das dimensões dos números atrelados ao câncer de pulmão e com isso perceber a real grandeza do problema, deixando clara a importância do desenvolvimento de ferramentas que possam oferecer qualquer tipo de auxílio no diagnóstico ou tratamento deste mal que atinge centenas de milhares de pessoas todos os anos. A partir destas informações foi explicitada também a necessidade da cooperação multidisciplinar entre profissionais da área médica e da computação para uma junção de esforços no combate ao câncer.

Foram apresentados também técnicas, procedimentos e protocolos utilizados em diferentes etapas da jornada de diagnóstico do câncer de pulmão. Foram abordadas as principais técnicas utilizadas atualmente na detecção do câncer bem como uma breve comparação entre estas e a técnica utilizada neste trabalho. Em seguida, o

protocolo DICOM de imagens médicas foi apresentado para explicar como funciona o mecanismo de criação, análise e armazenamento de dados médicos em imagem. Por fim, foram apresentadas as principais características de um sistema CAD, ferramenta computadorizada utilizada para auxiliar a tomada de decisão durante a análise de exames de tomografia.

No próximo capítulo serão apresentados os trabalhos relacionados a esta aplicação, que servirão como fundamentação técnica e teórica para possibilitar a conclusão deste trabalho.

## Capítulo 2

# TRABALHOS RELACIONADOS

**E**STE capítulo apresenta alguns trabalhos relacionados ao problema de detecção e classificação de nódulos pulmonares utilizando algoritmos de inteligência computacional em imagens de tomografia computadorizada. Os trabalhos introduzem ferramentas conhecidas na área médica como sistemas de diagnóstico auxiliado por computador (*Computer Aided Diagnosis - CAD*) e servem como uma grande base teórica para este estudo. Neste capítulo, os trabalhos estão segmentados de acordo com as técnicas utilizadas para tratar o problema. Estes são agrupados em trabalhos que utilizam aprendizagem profunda, redes neurais de menor complexidade, outros métodos de inteligência computacional e trabalhos que apresentam técnicas para tratamento de imagens para sistemas CAD.

São discutidas também neste capítulo as semelhanças e diferenças dos trabalhos perante a este estudo, bem como um pouco de cada metodologia utilizada nos artigos, apresentando seus desafios e resultados.

### 2.1 Aprendizagem profunda para classificação

Em (KUAN et al., 2017), uma ferramenta CAD que utiliza aprendizagem profunda através de uma rede neural convolucional é introduzida para detectar e classificar nódulos pulmonares. O trabalho foi desenvolvido para participar da competição *Kaggle Data Science Bowl 2017 Challenge*, que mobiliza pesquisadores e desenvolverem modelos capazes de detectar com precisão o câncer de pulmão. Os autores relatam os principais desafios deste tipo de aplicação, dentre eles a dificuldade da distinção entre nódulos malignos e benignos e a confusão causada por vasos sanguíneos que se parecem com nódulos nas imagens de tomografia.

Para o treinamento do modelo, foi utilizado o banco de imagens LUNA16 (*Lung Nodule Analysis 2016*), que conta com 888 imagens de tomografias de pulmões e fornece informações sobre as posições dos nódulos. Embora o LUNA16 fosse bastante completo, este não fornecia informações sobre o diagnóstico de câncer do paciente, portanto houve a necessidade do uso de outro banco de dados, o KDSB17 (*Kaggle Data Science Bowl 2017*), que conta com 2101 imagens com rótulos de diagnóstico de câncer.

O trabalho utiliza uma técnica bastante comum na área da aprendizagem profunda, o aprendizado por transferência, que consiste em utilizar um modelo de rede neural convolucional pré-treinado e ajustar somente alguns parâmetros de suas últimas camadas para adaptá-lo a sua aplicação.

A rede escolhida pelos autores foi a *ResNet*, criada pela *Microsoft* em (HE et al., 2016) e tida por muitos como o atual estado da arte para aplicações de visão computacional. Esta rede foi modificada para se adaptar às imagens de tomografia pulmonar e o modelo foi estruturado de acordo com a seguinte rotina: pré-processamento das imagens, detecção de nódulos, análise de malignidade, classificação dos nódulos e diagnóstico de câncer.

A rede foi treinada por 100.000 iterações, com taxa de aprendizagem 0,01 e regularização de pesos, classificando imagens em 3 classes: (nódulo maligno, nódulo benigno e sem nódulo) e diagnosticando pacientes conforme um valor percentual de saída da rede.

O desempenho do modelo foi medido pela métrica *Log-Loss*, estipulada pela competição *Kaggle Data Science Bowl 2017 Challenge* e após a submissão, o estudo atingiu a 41<sup>a</sup> colocação na competição entre 1971 competidores, atestando que os métodos utilizados são realmente eficientes para a aplicação.

## 2.2 Classificação com rede U-Net

Em (CHON; BALACHANDRA; LU, 2017), uma outra ferramenta CAD que utiliza aprendizagem profunda através de uma rede neural convolucional é introduzida para detectar e classificar nódulos pulmonares. Este trabalho foi também desenvolvido para participar da competição *Kaggle Data Science Bowl 2017 Challenge* e utilizou os mesmos bancos de imagens LUNA16 e KDSB17 da aplicação da seção anterior.

A principal diferença entre este trabalho e o desenvolvido por (KUAN et al., 2017), está nas redes escolhidas para a aplicação do aprendizado por transferência. Em (CHON; BALACHANDRA; LU, 2017), os autores utilizaram a rede convolucional U-Net, especializada em reconhecer padrões em imagens médicas de tomografia computadorizada, o que aumentou significativamente a performance da fase de segmentação de nódulos. Após esta fase de segmentação, os nódulos foram classificados por outra rede convolucional estado da arte modificada com o aprendizado por transferência, a *GoogLeNet* descrita em (SZEGEDY et al., 2015), capaz de atingir desempenhos sobre-humanos.

Este trabalho obteve resultados ainda mais expressivos do que o apresentado em (KUAN et al., 2017), atingindo maior acurácia, sensibilidade e especificidade. Em comparação com este estudo, ambos os trabalhos apresentados acima seguiram caminhos bem próximos, utilizando redes neurais convolucionais e aprendizado por transferência para desenvolver modelos de classificação de nódulos pulmonares.

A principal diferença é que estes não utilizaram algoritmos de inteligência de enxame para ajustar os parâmetros das redes. A seguir são apresentados trabalhos que utilizam redes neurais menos complexas, quando comparadas às redes convolucionais, para a detecção e classificação de nódulos cancerígenos em imagens médicas. Estes trabalhos fornecem técnicas alternativas que são úteis para reduzir o esforço computacional da aplicação ou para otimizar processos mais complexos.

## 2.3 Diagnósticos por redes neurais e PSO

Em (JAYANTHI; PREETHA, 2016), os autores apresentam um sistema CAD que utiliza uma rede neural artificial ajustada com o algoritmo de inteligência de enxame *Particle Swarm Optimization* para detectar e classificar nódulos cancerígenos em exames de mamografia.

O estudo utiliza o banco de imagens *Mammographic Image Analysis Society* (MIAS), que contém 322 imagens de mamografias classificadas em: presença de tumores malignos, presença de tumores benignos ou ausência de tumores, para treinar o modelo. O trabalho foi conduzido de acordo com a seguinte rotina: pré-processamento das imagens com filtros de mediana, segmentação utilizando *fuzzy C-Means*, extração de atributos como intensidade, textura e formato e classificação utilizando a rede neural.

Devido à escassez de dados, os autores tiveram que desenvolver um modelo mais focado na extração de atributos através de técnicas de tratamento de imagens. Após

a aplicações destas técnicas, a classificação foi feita de duas maneiras diferentes, uma utilizando uma rede neural ajustada de maneira tradicional e outra utilizando uma rede neural ajustada através de um algoritmo de otimização por enxame de partículas (*Particle Swarm Optimization* - PSO). A abordagem que não utilizou o algoritmo PSO obteve 88,6% de acurácia, 72,7% de sensibilidade e 93,6% de especificidade, enquanto a que utilizou o PSO atingiu 95,6% de acurácia, 87,2% de sensibilidade e 97,3% de especificidade, sendo superior em todas as métricas registradas.

Os resultados obtidos confirmam que a utilização de algoritmos de inteligência de enxame como o PSO para o ajuste de parâmetros de redes neurais pode ser uma ótima ferramenta de aumento de performance para aplicações como esta. Por este motivo, este trabalho serviu como uma boa fonte de informações para este estudo, pois apresentou boas abordagens de tratamento de imagens e de ajuste de parâmetros.

## 2.4 Detecção de câncer com redes neurais

Em (MIAH; YOUSUF, 2015), os autores apresentam um sistema CAD que utiliza técnicas de processamento de imagens e uma rede neural artificial para detectar e classificar nódulos cancerígenos em exames de tomografia computadorizada de pulmões. Esse trabalho é uma boa fonte de comparação para os métodos aplicados neste estudo, pois o modelo obtido por (MIAH; YOUSUF, 2015) atinge uma acurácia de 0,966 utilizando somente técnicas de processamento de imagens e uma rede neural *Feed-Foward Multilayer Perceptron* para classificação.

Os autores do estudo utilizaram uma abordagem diferente dos apresentados até então, pois dividiram as imagens ao meio para aplicar as técnicas separadamente a cada pulmão. Dentre as técnicas de processamento de imagens utilizadas, destacam-se:

- Conversão para imagem em escala de cinza, seguida de normalização entre 0 e 1 e redução de ruídos.
- Binarização da imagem: Conversão de valores para 100% branco ou 100% preto.
- Remoção de partes não necessárias: Rotina para otimizar o reconhecimento de padrões.
- Segmentação por processos de morfologia: Rotinas de detecção de bordas.

Após estas etapas, as imagens resultantes são analisadas e fornecem 33 atributos de entrada para a rede classificadora. Estes atributos são calculados através de relações matemáticas entre partes do pulmão, densidades de pixels, etc. Como resultado, o modelo produz uma classificação que identifica nódulos cancerígenos e os localiza em cada pulmão. O estudo de (MIAH; YOUSUF, 2015) introduz uma abordagem simples, porém muito eficiente, que pode ser utilizada em redes mais complexas como as de aprendizagem profunda.

Em (MESSAY; HARDIE; TUINSTRA, 2015), os autores apresentam um sistema CAD que utiliza técnicas de processamento de imagens e uma rede neural de regressão para detectar nódulos cancerígenos em exames de tomografia computadorizada de pulmões. Este trabalho traz resultados interessantes pois, dentre os apresentados nesta lista, é o que trabalha com o banco de imagens mais extenso, ou seja, é o único a trabalhar com um volume de dados expressivos o suficiente para aplicações como esta. O banco utilizado nesta aplicação foi o *Lung Image Database Consortium and Image Database Resource Initiative* (LIDC-IDRI), que contém cerca de 245.000 imagens de tomografias computadorizadas de pulmões.

Devido à vasta quantidade de dados, os autores foram capazes de produzir um modelo altamente robusto, capaz de obter resultados expressivos diante de diversas situações. Após a finalização do estudo, foram produzidos dois modelos, um semiautomático treinado para classificar a malignidade e dar um diagnóstico uma vez que um humano aponte a localização de um nódulo e um automático, que detecta os nódulos, os classifica e infere um diagnóstico de câncer. Os modelos foram testados e validados em diversas partições diferentes, produzindo acurácias médias de 91,32% para o modelo semiautomático e de 88,65% para o modelo automático. Embora (MESSAY; HARDIE; TUINSTRA, 2015) tenham obtido resultados percentuais levemente inferiores aos de outros modelos apresentados aqui, é de se esperar que os índices obtidos por estes autores sejam mais confiáveis devido ao tamanho dos conjuntos de treinamento e teste.

## 2.5 Utilização de outros métodos

Nesta seção, são apresentados trabalhos que utilizam outros métodos que não redes neurais comuns ou redes convolucionais para a detecção e classificação de nódulos cancerígenos em imagens médicas. Estes trabalhos fornecem técnicas alternativas que são úteis para

comparação de modelos e também para extração de métodos que podem ser combinados para produzir resultados otimizados.

### 2.5.1 Detecção por crescimento de regiões

Em (PARVEEN; KAVITHA, 2013), os autores apresentam um sistema CAD que utiliza o banco de imagens LIDC-IDRI para aplicar procedimentos de tratamento de imagens e um classificador que opera com o *Region Growing Algorithm* para detectar nódulos pulmonares cancerígenos em imagens de tomografia computadorizada. Ao início do procedimento as imagens passam pelas seguintes etapas: tratamento com filtros de mediana, binarização da imagem, detecção de bordas, remoção de partes desnecessárias e aplicação do algoritmo *Flood Fill*.

Após o tratamento das imagens, estas são submetidas ao algoritmo de crescimento de regiões, que verifica regiões de *pixels* semelhantes e realiza uma seleção de regiões promissoras às características buscadas. Ao final de algumas iterações, o algoritmo retorna uma imagem binária que possui somente os nódulos. Neste estudo, os autores introduzem uma sequência de métodos muito úteis para aplicações como está.

### 2.5.2 Detecção por agrupamento e classificação

Em (SIVAKUMAR; CHANDRASEKAR, 2013), os autores apresentam um sistema CAD que utiliza procedimentos de agrupamento com diferentes algoritmos de lógica *fuzzy* e um classificador de máquinas de vetor de suporte (*Support Vector Machines - SVM*) para detectar e classificar nódulos cancerígenos em imagens de tomografia computadorizada de pulmões.

Os autores do estudo optaram por utilizar o banco de imagens LIDC-IDRI para treinar seus modelos, obtendo resultados de alta robustez. No trabalho, as imagens passam por uma rotina de pré-processamento semelhante à aplicada em (JAYANTHI; PREETHA, 2016), porém são aplicados 3 métodos de agrupamento diferentes, o *fuzzy C-Means*, o *fuzzy-Possibilistic C-Means* e o *Weighted fuzzy-Possibilistic C-Means*. Após as fases de tratamento de imagens e extração de características, as imagens foram submetidas a um classificador SVM. Para obter melhores resultados, o classificador SVM foi treinado com 3 funções de núcleo (*Kernels*) diferentes, uma linear, uma polinomial e uma de base radial. Dos trabalhos apresentados até então, este obteve o pior resultado, com uma acurácia de

80,36%, fato que pode estar relacionado com a dificuldade de se encontrar uma função núcleo ideal para um classificador SVM.

### 2.5.3 Detecção por classificação Bayesiana

Em (DHAWARE; PISE, 2016), os autores apresentam um sistema CAD que utiliza procedimentos de agrupamento de lógica *fuzzy* e um classificador *Bayesiano* para detectar e classificar nódulos cancerígenos em imagens de tomografia computadorizada de pulmões. Neste trabalho, os autores aplicam técnicas estatísticas durante o tratamento das imagens para extrair os atributos de entrada do classificador *Bayesiano*. Este modelo obteve desempenho pouco superior ao modelo anterior, que utilizou SVM.

## 2.6 Técnicas para tratamento de imagens em sistemas CAD

Nesta seção, são apresentados trabalhos que descrevem detalhadamente os principais métodos de tratamentos de imagem utilizados por sistemas CAD para detecção e classificação de nódulos cancerígenos pulmonares. Estes trabalhos servem como grandes fontes de informação para as fases de processamento de imagens desta aplicação.

### 2.6.1 Fusão de características

Em (FARAG et al., 2017), os autores apresentam uma série de técnicas de tratamento de imagens médicas de tomografia computadorizada de pulmões para identificação de nódulos. Os autores utilizam o banco de imagens LIDC-IDRI para aplicar métodos como filtragem *Gabor*, reconhecimento de padrões de textura local binária (*Local Binary Pattern* - LBP) e filtragem por distância combinada com LBP.

Após estes processos de tratamento, os autores demonstram que é possível extrair dezenas de características diferentes das imagens, como relações “altura x largura” dos nódulos, volume, densidade, textura, etc. Este trabalho providencia uma grande ferramenta para análise de imagens neste tipo de aplicação.

### 2.6.2 Remoção de ruídos e aguçamento de imagens

Em (YUVARANI, 2012), o autor apresenta técnicas de tratamento de imagens médicas de tomografia computadorizada de pulmões para identificação de nódulos. O objetivo

do estudo é aplicar procedimentos de filtragem por lógica *fuzzy* para eliminar ruídos e manter somente as informações relevantes de cada imagem. Após a descrição detalhada dos processos utilizados, o autor demonstra o desempenho do método *fuzzy* de filtragem em relação às filtragens *Gaussianas* e de mediana, comparando-as pelo erro médio quadrático em relação à imagem original.

Em todas as comparações, os filtros *fuzzy* obtiveram melhores resultados em remover ruídos, portanto este método se mostrou bastante relevante para esta aplicação, sendo uma importante base teórica para o presente estudo.

## 2.7 Considerações finais

Neste capítulo, foram apresentados trabalhos relacionados ao problema de detecção e classificação de nódulos pulmonares utilizando algoritmos de inteligência computacional em imagens de tomografia computadorizada. Com estes trabalhos foi possível identificar os principais desafios e características desta aplicação, obter informações sólidas de que os modelos baseados em redes neurais convolucionais são de fato muito eficientes, comprovar que os algoritmos de inteligência de enxame são realmente eficientes em aumentar a performance de alguns modelos, identificar técnicas que podem reduzir o esforço computacional e ainda sim produzir resultados expressivos, identificar os melhores softwares, *frameworks* e redes pré-treinadas para se utilizar nesta aplicação, obter métricas de comparação de modelos, identificar quais são as melhores técnicas para se utilizar, bem como as que não devem ser utilizadas, escolher os melhores bancos de imagens para a aplicação, aprender técnicas avançadas e muito eficientes de tratamento de imagens médicas, entre outros. Cada trabalho contribuiu com porções de conhecimento que, quando combinadas resultaram em uma base sólida para o desenvolvimento de um estudo como este. No próximo capítulo serão apresentadas as técnicas de inteligência de enxame utilizadas no desenvolvimento deste estudo.

## Capítulo 3

# INTELIGÊNCIA DE ENXAME

**I**NTELIGÊNCIA de enxame é o termo utilizado para denominar uma família de algoritmos de inteligência computacional que possuem a característica comum de simular uma população de indivíduos pouco inteligentes que, baseada em certas leis de convivência, consegue produzir comportamentos de alta complexidade. As principais aplicações destes algoritmos são destinadas à otimização de processos, como encontrar valores ótimos para uma função matemática, encontrar os melhores caminhos para um sistema de navegação ou organizar da melhor maneira uma colônia de robôs empenhados em realizar uma certa tarefa. A metodologia destes processos é baseada na observação do comportamento de animais ou micro-organismos, que quando vivem em uma comunidade ou ecossistema, conseguem com simples ações produzir comportamentos complexos de enxame e realizar tarefas de maneira organizada.

### 3.1 Visão geral

O termo enxame, ou população, é utilizado de forma genérica para se referir a qualquer coleção estruturada de agentes capazes de interagir entre si. O exemplo clássico de um comportamento de enxame é a colônia de abelhas, onde os indivíduos se dividem de acordo com suas funções sociais e trabalham de forma ordenada para que a colônia possa prosperar. Podemos considerar também outros sistemas, como por exemplo uma colônia de formigas, uma revoada de pássaros, uma colônia de bactérias, manadas de lobos, cardumes, entre outros.

A noção de enxame sugere um aspecto de movimento coletivo no espaço, todavia, como em um enxame de pássaros, estamos interessados em todos os tipos de comportamentos coletivos, não somente espacial. Logo, podemos dizer que as interações

coletivas de todos os agentes dentro do sistema muitas vezes leva a algum tipo de comportamento ou inteligência coletiva. Este tipo de inteligência artificial inclui qualquer tentativa de projetar algoritmos ou dispositivos distribuídos de solução de problemas sem ter um controle centralizado, inspirado no comportamento coletivo de agentes sociais e outras sociedades animais. A inteligência coletiva é uma propriedade de sistemas compostos por agentes pouco inteligentes com capacidade individual limitada, capazes de apresentar comportamentos coletivos inteligentes. Tais comportamentos, em geral, apresentam as seguintes propriedades:

- Comportamento social: Os agentes são capazes de interagir entre si.
- Auto avaliação: Os agentes devem ser capazes de avaliar seu próprio desempenho, para tomar decisões que visam a otimização.
- Estabilidade: Nem todas as variações ambientais devem afetar o comportamento de um agente.
- Adaptabilidade: Capacidade de se adequar a variações ambientais.

## 3.2 Algoritmos de inteligência de enxame

Nesta seção, serão apresentados resumos dos algoritmos de enxame aplicados neste trabalho. Ao todo foram aplicados 14 algoritmos para encontrar os melhores parâmetros para a rede neural classificadora do modelo desta aplicação. São eles, *Particle Swarm Optimization* (KENNEDY, 2010), *Artificial Bee Algorithm* (KARABOGA; BASTURK, 2007), *Bat Optimization Algorithm* (ALIHODZIC; TUBA, 2013), *Bacterial Foraging Optimization* (PASSINO, 2010), *Gray Wolf Optimization* (MIRJALILI; MIRJALILI; LEWIS, 2014), *Firefly Optimization Algorithm* (RITTHIPAKDEE THAMMANO; JITKONGCHUEN, 2017), *Cuckoo Search Algorithm* (RAJABIOUN, 2011), *Whale Swarm Algorithm* (MIRJALILI; LEWIS, 2016), *Firework Algorithm* (TAN; ZHU, 2010), *Chicken Swarm Optimization* (MENG LIU; ZHANG, 2014), *Social Spider Algorithm* (CUEVAS; CIENFUEGOS, 2014), *Cat Swarm Optimization* (CHU; TSAI, 2007), *Harmony Search* (GEEM; KIM; LOGANATHAN, 2001), e o *Gravitational Search Algorithm* (KHAJEHZADEH; ESLAMI, 2012).

Em geral, algoritmos de otimização em enxame tendem a seguir rotinas similares, onde inicialmente um conjunto de agentes é gerado em posições aleatórias e, em cada

rodada, estes se movimentam levando em consideração fatores como, auto-avaliações de seu desempenho atual, sua memória sobre os melhores desempenhos que obteve, o instinto de seguir os líderes do enxame, entre outros. Estas rotinas são repetidas até uma ou mais partículas encontrem resultados ótimos ou suficientemente satisfatórios.

Nesta aplicação, algoritmos de enxame são utilizados para encontrar o melhor conjunto de parâmetros para uma rede neural de classificação de nódulos cancerígenos de pulmão. Neste caso, cada partícula do enxame é representada por possíveis valores para o conjunto de pesos da rede e a otimização acontece de acordo com o Algoritmo 1.

---

**Algoritmo 1** Treinamento com enxame

---

```
e = número total de épocas;
n = tamanho do enxame;
Criação do primeiro enxame;
para i = 1 até e faça
  para j = 1 até n faça
    Elege a j-ésima partícula para compor os pesos da rede neural de classificação;
    Avalia o fitness desta partícula através do erro médio quadrático na classificação do conjunto de
    treinamento;
  fim para
  Aplica a lógica do algoritmo de enxame para gerar um novo enxame;
fim para
Elege a melhor partícula para ser o conjunto final de pesos da rede;
```

---

Em geral, cada partícula de um algoritmo de enxame representa uma possível solução para o problema de otimização e pode ser representada por um vetor de coordenadas de posição da partícula no espaço de busca. Este vetor é atualizado constantemente de acordo com as regras de convivência entre partículas, estipuladas por cada algoritmo durante o processo de otimização.

O objetivo dos algoritmos de enxame é simular um ambiente onde indivíduos consigam vagar pelo espaço de busca em direções otimizadas, tendo suas velocidades e posições atualizadas dinamicamente de acordo com o histórico das experiências individuais e coletiva de todo o enxame. Logo, o desempenho do algoritmo está associada à trajetória percorrida pelo enxame e ao tempo gasto para encontrar a melhor solução para o problema. Basicamente, como mostrado na Figura 6, as partículas são espalhadas aleatoriamente pelo espaço de busca ao início do processo de otimização. Com a aplicação das lógicas de inteligência de enxame, espera-se observar uma convergência dos indivíduos para um ponto ou região mais promissora, encontrando um resultado ótimo para a aplicação desejada.

Outro componente importante, que influencia no desempenho de algoritmos de enxame é a estrutura de comunicação, ou topologia, das partículas. Ela rege como as

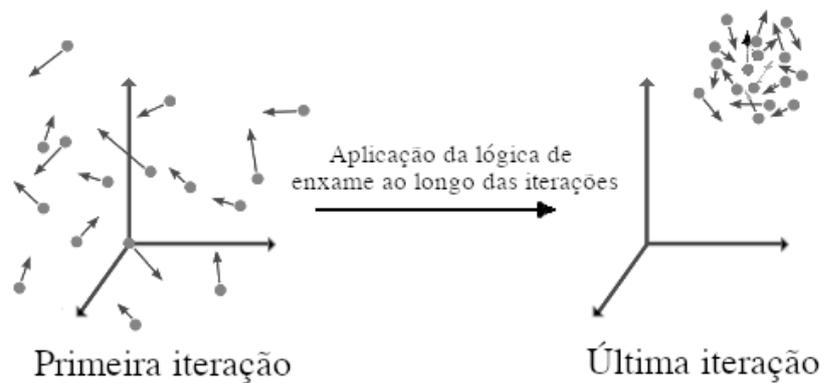


Figura 6: Partículas buscando um resultado ótimo

partículas do enxame trocam informações e se comunicam. Logo, a escolha da topologia influencia na avaliação cognitiva de cada partícula, fator responsável pela tomada de decisão sobre a intensidade de movimentação no espaço de busca. A depender de como as partículas se comunicam entre si e do problema a ser tratado, a busca pela solução ótima pode priorizar tanto a velocidade de convergência, a qualidade da solução ou ambas. As principais topologias utilizadas como mecanismos de comunicação entre as partículas são: a topologia estrela, onde todas as partículas se comunicam entre si e topologias locais, onde cada partícula se comunica apenas com um certo número de partículas vizinhas. A Figura 7 apresenta algumas destas topologias, onde temos à esquerda a topologia *lbest*, também chamada de anel, que conecta cada partícula a suas 2 partículas vizinhas; ao centro a topologia de roda, que conecta todas as partículas a uma partícula central, e à direita a topologia *gbest*, também chamada de estrela, onde cada partícula está ligada a todas as outras partículas do enxame.

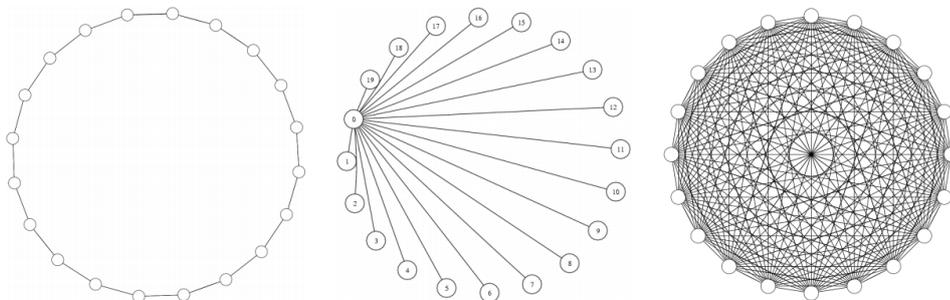


Figura 7: Exemplos de topologias [Fonte: (MENDES, 2004)]

### 3.2.1 Otimização por enxame de partículas

O *Particle Swarm Optimization* (PSO), desenvolvido por (KENNEDY, 2010), é um tipo de inteligência de enxame inspirado no comportamento de bandos de pássaros. A busca por alimentos e a interação entre aves ao longo do voo são modeladas como um mecanismo de otimização. Neste caso, a área sobrevoada é equivalente ao espaço de busca e encontrar o local com a maior quantidade ou qualidade de comida corresponde a encontrar a solução ótima.

O algoritmo é modelado por pássaros, ou partículas, que fazem uso de sua experiência e da experiência do próprio bando para encontrar a melhor região do espaço de busca. Para que o bando de pássaros sempre se aproxime do objetivo, ao invés de se perder ou nunca alcançar o alvo focado, utiliza-se o indicador denominado *fitness*, ou função de aptidão, que irá avaliar o desempenho das partículas em cada ponto do espaço de busca.

Durante a otimização, cada partícula consegue se lembrar do ponto no espaço em que ela obteve seu melhor resultado, este ponto é chamado de melhor pessoal, ou *personal best*, geralmente apresentado como *pbest*. Além de sua memória pessoal, uma partícula deve ser capaz de se comunicar com o enxame ou com seus vizinhos mais próximos para saber qual foi o melhor resultado obtido até agora pelo enxame como um todo. Este resultado é chamado de melhor global, ou global best, geralmente apresentado como *gbest*,

Neste algoritmo, os parâmetros necessários são: O coeficiente de inércia, que é responsável por manter parte da velocidade da partícula em seu curso atual, o coeficiente cognitivo, que é responsável por direcionar a partícula para o melhor ponto encontrado por ela até então, e o coeficiente social, que é responsável por direcionar a partícula para o melhor ponto encontrado pelo enxame até o momento.

### 3.2.2 Colônia artificial de abelhas

O *Artificial Bee Colony* (ABC) é uma técnica de otimização que simula o comportamento social das abelhas em suas colônias. Esta técnica tem sido aplicada com sucesso em vários problemas práticos. O ABC foi proposto por (KARABOGA; BASTURK, 2007).

No modelo ABC, a colônia consiste em três grupos de abelhas: abelhas empregadas, observadoras e escoteiras. Supõe-se que haja apenas uma abelha empregada para cada fonte de alimento. Em outras palavras, o número de abelhas empregadas na colônia é

igual ao número de fontes de alimento ao redor da colmeia. As abelhas empregadas vão à sua fonte de comida, voltam para a colmeia e dançam nesta área. A abelha empregada cuja fonte de alimento foi abandonada torna-se uma escoteira e começa a procurar por uma nova fonte de alimento. As observadoras assistem as danças de abelhas empregadas e escolhem as fontes de alimento dependendo das danças. Os principais passos do algoritmo são os seguintes:

- Fontes alimentares iniciais são produzidas para todas as abelhas empregadas.
- Cada abelha empregada vai para uma fonte de alimento em sua memória e determina uma fonte mais próxima, depois avalia sua quantidade de néctar e dança na colmeia.
- Cada observadora assiste à dança das abelhas empregadas, escolhe uma de suas fontes e vai à aquela fonte.
- Depois de escolher esta avalia o valor do néctar.
- Fontes alimentares abandonadas são substituídas pelas novas fontes de alimento descobertas pelas escoteiras.
- A melhor fonte de comida encontrada até agora é registrada.
- Os processos são repetidos até que o critério de parada seja atingido

No ABC, a posição de uma fonte de alimento representa uma possível solução para o problema de otimização e a quantidade de néctar de uma fonte de alimento corresponde à qualidade (aptidão) da solução associada. O número de abelhas empregadas é igual ao número de soluções na população. Na primeira etapa, é gerada uma população inicial distribuída aleatoriamente (posições da fonte de alimento). Após a inicialização, a população é submetida a repetir os ciclos dos processos descritos até o fim da otimização.

Uma abelha empregada guarda um registro de sua fonte alimentícia atual em sua memória. Caso descubra uma nova fonte, esta só trocará de posição se a quantidade de néctar da nova fonte seja maior. Neste caso a abelha memoriza a nova posição de origem e esquece a antiga. Caso contrário, ela mantém sua posição. Depois de todas as abelhas empregadas completarem o processo de busca, elas compartilham as informações de posição das fontes com as observadoras na área de dança. Cada observadora avalia as informações de néctar tomadas de todas as abelhas empregadas e, em seguida, escolhe

uma fonte de alimento dependendo das quantidades de néctar. Neste algoritmo, apenas os hiper-parâmetros de quantidade de partículas e número de épocas.

### 3.2.3 Busca otimizada com morcegos

O *Bat Optimization Algorithm* (BOA) é baseado na habilidade de eco-localização desenvolvida por estes animais. Utilizando a eco-localização, os morcegos são capazes de detectar comida, predadores e até diferenciar tipos de insetos no escuro. Morcegos emitem ondas sonoras e conseguem detectar e interpretar com alta definição o eco produzido pela reflexão das ondas nos objetos do entorno.

Este algoritmo foi introduzido por (ALIHODZIC; TUBA, 2013) e se baseia nos seguintes princípios; cada morcego voa aleatoriamente com uma certa velocidade, uma frequência variável e uma intensidade sonora. Ao pesquisar e encontrar sua presa, ela altera a frequência, a intensidade e a taxa de emissão de pulso. A seleção da melhor solução continua até que o critério de parada seja atendido.

Neste algoritmo, os parâmetros necessários são: Frequências máxima e mínima, responsáveis pela abrangência de sinal percebida, e os parâmetros que identificam o sinal emitido, que são o nível de emissão de pulso, o volume do som, a variação de volume e a alteração de pulso.

### 3.2.4 Otimização bacteriana

O *Bacterial Foraging Algorithm* (BFO), introduzido por (PASSINO, 2010), é inspirado no comportamento de forrageamento de bactérias como a *E.coli* e *M.xanthus*. Especificamente, o BFO é inspirado no comportamento de quimiotaxia de bactérias, que percebem gradientes químicos de nutrientes no ambiente e se movem em direção às regiões mais promissoras. A estratégia de processamento de informações do algoritmo é permitir que os indivíduos se movimentem de forma estocástica e coletiva em direção à otimização. Isto é possível através de uma combinação de três processos. São eles: Quimiotaxia, responsável pela movimentação das bactérias; Reprodução, responsável pela reprodução das bactérias que obtiverem bons desempenhos; Eliminação-dispersão, responsável por descartar bactérias de acordo com um fator estocástico.

Neste algoritmo, os parâmetros necessários são: Passos de quimiotaxia, responsável pela dosagem de passos dados por cada bactéria em uma iteração, a distância de

navegação, responsável por delimitar o quanto uma bactéria pode se mover, o tamanho do passo, responsável por ditar a distância percorrida em cada passo, e a probabilidade de eliminação, decisão estatística tomada para eliminar parte das bactérias

### 3.2.5 Busca por alcateia

O *Gray Wolf Optimization* (GWO) se inspira na hierarquia de liderança e nos mecanismos de caça desenvolvidos por estes mamíferos. Lobos vivem em alcateias compostas, em média, por 5 a 12 indivíduos. A hierarquia das alcateias é dividida em 4 níveis; Alfa, Beta, Delta e Omega. Lobos alfa são os líderes da alcateia e são responsáveis por tomar decisões, embora possam obedecer a outros lobos. Lobos beta ajudam os lobos alfa a tomarem decisões. Todos os betas são candidatos a se tornarem alfas caso o alfa morra ou fique velho demais. Os betas respeitam o alfa e passam os comandos para o resto da alcateia. Os lobos delta obedecem aos betas e dominam os ôegas, que são responsáveis por tomar conta dos filhotes e velhos. No modelo matemático deste algoritmo, a hierarquia é mapeada com o objetivo de se atingir um valor ótimo de uma função de aptidão. O lobo Alfa é considerado a melhor solução, beta e delta são a segunda e terceira melhor solução e todo o resto da população é considerada ômega. Os lobos alfa, beta e delta lideram a caça e os ôegas os seguem. Este algoritmo foi introduzido por (MIRJALILI; MIRJALILI; LEWIS, 2014). Como no ABC, este algoritmo só utiliza os hiper-parâmetros de tamanho do enxame e quantidade de épocas.

### 3.2.6 Algoritmo de vaga-lumes

O *Firefly Algorithm* (FA), introduzido por (RITTHIPAKDEE THAMMANO; JITKONGCHUEN, 2017) se baseia no comportamento dos vaga-lumes. Estes insetos podem produzir pequenos flashes de luz para atrair vaga-lumes do sexo oposto ou pequenas presas. O sinal luminoso serve também para sinalizar para predadores que estes vaga-lumes possuem um sabor amargo. O modelo matemático deste algoritmo se baseia em dois aspectos; a mudança na luz e atratividade. A atratividade de um vaga-lume é definida pelo brilho de sua luz, que está conectado à função objetivo.

O comportamento dos indivíduos do algoritmo segue as seguintes premissas: Todos os vaga-lumes podem se atrair independentemente de seu gênero, a atratividade de um vaga-lume é diretamente proporcional a seu brilho, vaga-lumes menos atraentes tendem a

voar em direção aos mais atraentes, o brilho percebido pelos vaga-lumes diminui conforme a distância entre eles, caso um vaga-lume não enxergue outro, este se move aleatoriamente.

Neste algoritmo, os parâmetros necessários são: índices de atração mútua e absorção de luz, que regem a interação entre insetos, e os parâmetros  $\alpha_1$ ,  $\alpha_2$ , primeiro parâmetro da *Gaussiana* e o segundo parâmetro da *Gaussiana*, que são responsáveis pela velocidade de convergência do algoritmo.

### 3.2.7 Busca otimizada dos cucos

O *Cuckoo Search Algorithm* (CSA), introduzido por (RAJABIOUN, 2011) se baseia no ciclo de vida das aves cucos. O modelo se baseia no comportamento de holoparasitismo de algumas espécies desta ave, que invadem ninhos de outras aves para botar seus ovos na esperança de que não sejam notados. Neste algoritmo, cada ovo em um ninho representa uma solução para o problema e o ovo recém posto pelo cuco uma nova solução. O objetivo é substituir soluções ruins por soluções boas em potencial. O algoritmo é baseado nas seguintes regras: Cada cuco bota um ovo por vez em um ninho aleatório, os ninhos com ovos de melhor qualidade irão gerar a nova geração, o número de ninhos anfitriões é fixado e o ovo posto por um cuco é descoberto pela ave dona do ninho de acordo com uma probabilidade, ovos descobertos são eliminados. Neste algoritmo, os parâmetros necessários são: a probabilidade de detecção do ovo e o número total de ovos.

### 3.2.8 Enxame de baleias

O *Whale Swarm Algorithm* (WSA), introduzido em (MIRJALILI; LEWIS, 2016) é um algoritmo de otimização baseado em baleias. Baleias são animais sociáveis e vivem em grupos, a comunicação entre indivíduos é feita através da produção de sons. Uma baleia é capaz de emitir uma grande variedade de sons, estes sons estão atrelados a padrões que indicam ações de migração, alimentação e amamentação. Na maioria dos casos, os sons emitidos são em ultrassom.

Quando uma baleia encontra um alimento ela emite um som para chamar outras baleias de seu bando e informar a quantidade e qualidade da comida. Sendo assim, cada baleia recebe constantemente vários sinais e se movimenta em direção à aquele que parece mais promissor. O algoritmo segue os seguintes passos: Todas as baleias se comunicam entre si por ultrassom dentro da área de busca, cada baleia possui a habilidade de calcular

sua distância a outras baleias, a qualidade e quantidade de comida encontrada é associada ao *fitness* desta baleia, o movimento de uma baleia é guiado pela baleia mais próxima que possuir um melhor *fitness* que esta.

Neste algoritmo, os parâmetros necessários são: Intensidade do ultrassom na origem, que norteia as baleias durante a movimentação, e a probabilidade de distorção da mensagem, que adiciona um caráter estocástico à solução.

### 3.2.9 Busca por fogos de artifício

O *Firework Optimization Algorithm* (FOA), introduzido por (TAN; ZHU, 2010) é um algoritmo baseado em fogos de artifício. Fogos de artifício explodem de maneira similar à forma como indivíduos procuram por soluções ótimas em algoritmos de enxame. O algoritmo de fogos de artifício consiste de 4 etapas, o operador de explosão, o operador de mutação, a regra de mapeamento e a seleção de estratégia. O efeito do operador de explosão é gerar faíscas em volta dos fogos de artifício. Depois disso, algumas faíscas são produzidas pelo operador de mutação. O operador de mutação utiliza um operador Gaussiano para produzir faíscas em uma distribuição Gaussiana. Diante desses 2 operadores, caso a faísca gerada não atinja uma região promissora, o mapeamento irá gerar novas faíscas em uma região promissora. Para selecionar faíscas para a próxima geração, a estratégia de seleção é utilizada.

Neste algoritmo, os parâmetros necessários são: o número de fogos normais e o número de fogos *Gaussianos*.

### 3.2.10 Otimização por enxame de galinhas

O *Chicken Swarm Optimization* (ChSO), introduzido por (MENG LIU; ZHANG, 2014), imita a ordem hierárquica pela qual galos, galinhas e pintinhos se comportam em enxame dentro de um galinheiro. O enxame dessas aves é dividido em vários grupos, onde cada grupo contém um galo, várias galinhas e pintinhos.

Aves diferentes seguem diferentes leis de movimento. Existem competições entre galinhas diferentes sob uma ordem hierárquica específica. Galinhas domésticas vivem em bandos. Existem mais de 30 sons distintos para a comunicação pelos quais eles podem comunicar muitas informações relacionadas a aninhamento, descoberta de alimentos, acasalamento e perigo. Além de aprender por tentativa e erro, as galinhas também

aprendem com sua experiência anterior e com as das outras para tomar decisões. A ordem hierárquica desempenha um papel significativo na vida social das galinhas. As galinhas preponderantes dominam as fracas. Existem as galinhas mais dominantes que permanecem perto dos galos, bem como as mais submissas que estão na periferia do grupo. Remover ou adicionar galinhas de um grupo existente causaria uma interrupção temporária na ordem social até que uma ordem específica seja estabelecida. O modelo segue os seguintes passos: No enxame de galinhas, existem vários grupos. Cada grupo compreende um galo dominante, galinhas e pintos. A divisão da identidade das aves (galos, galinhas e pintinhos) depende dos valores de aptidão de cada indivíduo. As aves com melhores aptidões são representadas como galos. As aves com piores valores de aptidão são designadas como pintinhos. As aves restantes são designadas como galinhas. As galinhas escolhem aleatoriamente em qual grupo viver. A relação mãe-filho entre as galinhas e os filhotes também é estabelecida aleatoriamente. A ordem hierárquica, o relacionamento de dominância e o relacionamento mãe-filho em um grupo permanecerão inalterados durante uma etapa do processo. Esses status são atualizados a cada nova etapa. As galinhas seguem seu galo companheiro de grupo para procurar comida, enquanto elas podem impedir que as pessoas comam sua própria comida. Os filhotes buscam comida em volta da mãe (galinha). Os indivíduos dominantes têm vantagem na competição por comida.

### 3.2.11 Colônia de aranhas

O *Social Spider Algorithm* (SSA), introduzido por (CUEVAS; CIENFUEGOS, 2014), imita o comportamento das colônias de aranhas. Algumas espécies de aranhas formam colônias que permitem que elas permaneçam juntas e seguras em uma rede comunitária. Uma colônia de aranhas sociais consiste em dois componentes principais: seus membros e sua teia comunitária. Todos os membros são divididos em dois grupos diferentes: homens e mulheres.

O SSA pressupõe que todo o espaço de busca é uma teia comunitária, onde todas as aranhas sociais interagem umas com as outras. Cada solução dentro do espaço de busca representa uma posição na teia. Cada aranha recebe um peso de acordo com o valor de adequação da solução simbolizada por ela. Dependendo de seu gênero, cada indivíduo é conduzido por um conjunto diferente de operadores evolutivos, que imitam

diferentes comportamentos cooperativos comumente assumidos dentro da colônia. Uma das características das aranhas sociais são as populações altamente influenciadas pelo sexo feminino. Neste algoritmo, o único parâmetro necessário é o fator estocástico.

### 3.2.12 Otimização com gatos

O *Cat Swarm Optimization* (CSO), introduzido por (CHU; TSAI, 2007), reproduz dois aspectos do comportamento dos gatos: modo de busca e modo de perseguição. Cada gato tem sua própria posição, composta de um certo número de dimensões, velocidades para cada dimensão, um valor de aptidão e uma identificação se o gato está em modo de busca ou modo de perseguição. A solução final será a melhor posição em um dos gatos, pois o CSA mantém a melhor solução até atingir o final das iterações. O modo de busca é usado para modelar a situação do gato que está descansando, olhando ao redor e buscando a próxima posição para se mover. Já o modo de perseguição modela o comportamento de ataque do gato que encontrou uma presa ou comida.

Neste algoritmo, os parâmetros necessários são: Número de gatos caçadores, coeficiente de memória, a contagem para a mudança de modo, o intervalo de dimensões, o limite de peso e o limite do custo.

### 3.2.13 Busca pela harmonia

O *Harmony Search* (HS), introduzido por (GEEM; KIM; LOGANATHAN, 2001) imita a técnica desenvolvida por músicos de jazz que possibilita que estes se adaptem rapidamente à harmonia da banda ou improvisem para manter uma boa melodia.

Neste modelo cada músico corresponde a um atributo em uma solução candidata, o tamanho e o alcance de cada instrumento correspondem aos limites e restrições da variável de decisão. A harmonia entre os músicos é tomada como uma solução candidata completa em um dado momento, e a apreciação da harmonia pelo público representa a função de custo específico do problema. Os músicos buscam a harmonia ao longo do tempo por meio de pequenas variações e improvisações, o que resulta em uma melhoria em relação à função custo. O objetivo é usar boas soluções já descobertas para influenciar a criação de novas soluções, visando encontrar as soluções ótimas.

Isso é obtido pela criação estocástica de soluções candidatas de maneira gradual, em que cada componente é extraído aleatoriamente de uma memória de soluções de alta

qualidade ou distribuídas aleatoriamente dentro dos limites do problema. A memória de soluções candidatas é inicialmente aleatória, e um critério de aceitação guloso é usado para admitir novas soluções candidatas somente se elas tiverem um valor objetivo aprimorado, substituindo um membro existente.

Neste algoritmo, os parâmetros necessários são: a taxa de ajuste do tom, a taxa de consideração da harmonia e a largura de banda do som.

### 3.2.14 Algoritmo de busca gravitacional

O *Gravitational Search Algorithm* (GSA), introduzido por (KHAJEZHADDEH; ESLAMI, 2012), se baseia nas leis de gravitação e atração de massas. Seu comportamento se assemelha ao PSO, pois este cria partículas que interagem por meio de leis de atração e relacionamento. O algoritmo opera com 2 leis:

- lei da gravitação: cada partícula atrai outras partículas e a força de gravidade entre duas partículas é diretamente proporcional ao produto de suas massas e inversamente proporcional à distância entre elas (deve-se atentar para o fato de que, diferentemente da lei da gravitação universal, não usamos o quadrado da distância, pois resulta em melhores resultados do algoritmo).
- lei do movimento: a velocidade de qualquer partícula é igual à soma de parte da velocidade no instante anterior com uma nova velocidade gerada pelo resultado de forças gravitacionais do sistema.

Como no ABC e no GWO, este algoritmo utiliza apenas os dois hiper-parâmetros comuns a todos os algoritmos desta aplicação.

## 3.3 Considerações finais

Este capítulo apresentou todas as técnicas de inteligência de enxame utilizadas neste estudo. Embora algumas técnicas sejam relativamente novas, estas são vastamente utilizadas em diversas aplicações de engenharia, garantindo resultados de otimização que seriam inviáveis de se atingir com métodos analíticos. Neste trabalho, estas técnicas foram utilizadas para otimizar parâmetros de modelos de aprendizagem profunda, mais especificamente para buscar os valores ideais para os pesos da rede neural totalmente

---

conectada de uma rede convolucional treinada para classificar imagens médicas e detectar nódulos cancerígenos de pulmão.

No próximo capítulo serão apresentados alguns dos aspectos referentes à aprendizagem profunda, principal tópico abordado neste trabalho.

# Capítulo 4

## APRENDIZAGEM PROFUNDA

**N**ESTE capítulo, são apresentadas as principais técnicas envolvidas no assunto base deste trabalho, a aprendizagem profunda de máquinas. Serão apresentados também modelos de exemplo que serviram como inspiração para este projeto, bem como modelos estado da arte.

### 4.1 Aprendizagem de máquinas

A aprendizagem de máquinas é um ramo da engenharia da computação onde são desenvolvidas técnicas focadas em transmitir ao computador parte da inteligência desenvolvida por humanos, animais ou até ecossistemas e assim, possibilitar soluções diversificadas para problemas de alta complexidade. O campo de pesquisa, em inteligência artificial é altamente ramificado, e faz com que a engenharia da computação possa atuar em grande sintonia com praticamente qualquer outra área do conhecimento. Suas aplicações variam de simples soluções para problemas matemáticos de baixa complexidade ao desenvolvimento de máquinas capazes de controlar o tráfego aéreo de aeroportos ou detectar doenças com precisões sobre-humanas. Dentre as principais aplicações da aprendizagem de máquinas, destacam-se: Visão computacional, onde modelos de alta complexidade são construídos para detectar objetos em imagens ou vídeos; Análises numéricas e previsões, utilizadas principalmente em aplicações financeiras como previsões de investimentos ou valores de ações; Análises linguísticas, aplicações focadas em reconhecer linguagem humana para fins de traduções instantâneas ou para a criação de assistentes virtuais capazes de interagir através de diálogo linguístico; Otimização de marketing digital, desenvolvimento de sistemas capazes de traçar perfis de compra de indivíduos baseado em sua atividade na internet, dessa forma, são criados sistemas

de recomendação de produtos para aumentar as chances de conversão de vendas; Sistemas especialistas: Compreendem todas as aplicações onde um modelo recebe um treinamento com conhecimento humano ou proveniente de grandes bancos de dados sobre um determinado assunto. O modelo então absorve esse conhecimento e passa a atuar como um especialista, realizando tarefas com eficiências que podem facilmente ultrapassar a humana.

Além de uma vasta gama de aplicações, a aprendizagem de máquinas conta com um número crescente de algoritmos desenvolvidos para servirem como base técnica para estes modelos inteligentes. Estes algoritmos são geralmente agrupados em famílias, que levam em consideração fatores como, tipos de aplicações, métodos utilizados e formatos de resposta para classificá-los. Nestas famílias, temos algoritmos mais simples, que utilizam apenas algumas inferências estatísticas para tomadas de decisão. Temos algoritmos utilizados para aprender funções matemáticas chamados de modelos regressores, que podem variar de simples técnicas como a regressão por mínimos quadrados (GOLUB; LOAN, 1980), até modelos mais complexos como os *support vector machines* (SUYKENS; VANDEWALLE, 1999), considerados por muitos pesquisadores como os melhores separadores lineares. Existe também uma grande família de algoritmos agrupadores, também chamados de clusterizadores, que utilizam grandes massas de dados para encontrar padrões e formar grupos onde os pontos possuem características similares. Na área de desenvolvimento de jogos, uma família de algoritmos é altamente utilizada, são os algoritmos de aprendizado por reforço, capazes de treinar agentes virtuais a realizarem tarefas complexas e reagir a ações dos usuários. Dentre as famílias de algoritmos utilizados para otimização de processos, destacam-se os algoritmos de inteligência de enxame, que se inspiram em colônias de animais ou microrganismos para simular indivíduos artificiais pouco inteligentes, que quando em grupo conseguem realizar tarefas de alta complexidade. A Figura 8 apresenta uma relação entre a complexidade técnica e a probabilidade de uso das principais famílias de algoritmos de aprendizagem de máquinas. Nela é possível perceber que a aprendizagem profunda e o aprendizado por transferência, principais técnicas aplicadas neste estudo, aparecem como umas das mais cotadas para a maioria das aplicações de inteligência artificial na atualidade mesmo sendo técnicas de relativamente alta dificuldade técnica.

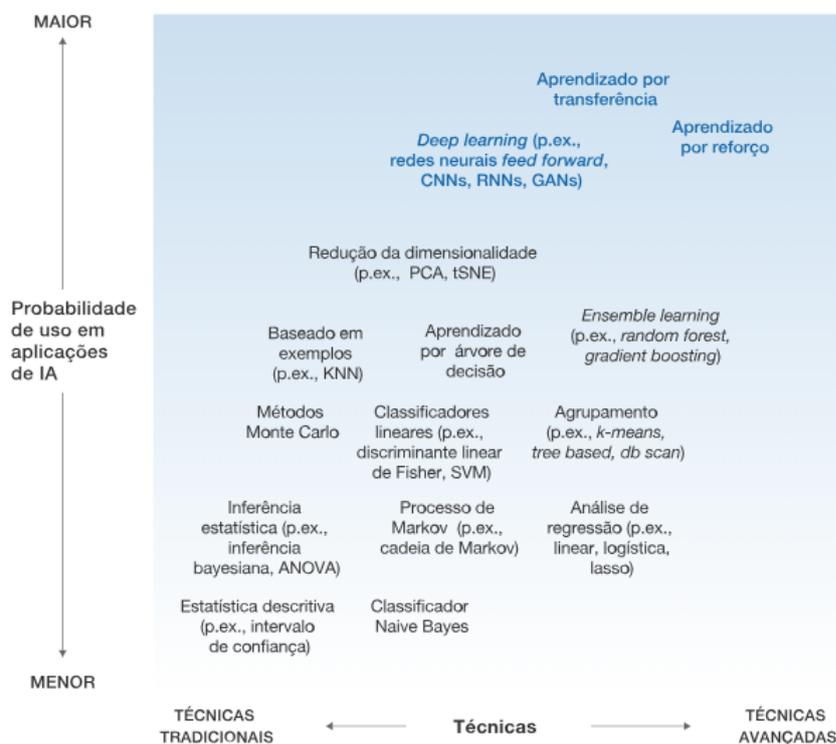


Figura 8: Uso e complexidade dos principais algoritmos de aprendizagem de máquinas [Fonte: (KELNAR, 2016)]

Redes neurais são um subconjunto de técnicas de aprendizagem de máquina. São, essencialmente, sistemas de inteligência artificial baseados em simular neurônios conectados e focados em aprender uma certa tarefa. Modelos computacionais inspirados por conexões neurais começaram a ser estudados na década de 1940 e, com o aumento do poder de processamento dos computadores e o uso de grandes conjuntos de dados de treinamento, voltaram aos holofotes e ganharam papel de destaque na análise bem-sucedida de dados como imagens, vídeo e fala.

## 4.2 Redes neurais de alta complexidade

Nos últimos anos, muito vem se falando sobre aplicações de redes neurais onde são necessárias estruturas altamente complexas para lidar com o aprendizado de conjuntos massivos de dados. De acordo com estas demandas, surgiu o termo aprendizagem profunda. A aprendizagem profunda é um subgênero do aprendizado de máquinas onde modelos de redes neurais de alta complexidade são criados para simular a grande capacidade do cérebro humano de reconhecer padrões mais abstratos ou sutis.

Assim como as redes neurais mais comuns, as redes de aprendizagem profunda são estruturadas em camadas, que realizam operações matemáticas de soma e multiplicação para ajustar seus parâmetros e aprender a realizar alguma tarefa mediante um treinamento. A principal diferença entre as redes neurais comuns e as de aprendizagem profunda, é a quantidade de camadas presentes em cada modelo. Enquanto as redes mais comuns costumam operar com não mais do que 5 ou 6 camadas, as redes de aprendizagem profunda são estruturadas em arquiteturas que podem ultrapassar a marca de 150 camadas, como é o caso de algumas das redes utilizadas por empresas como a *Google*<sup>™</sup>, em (SZEGEDY et al., 2015) e a *Microsoft*<sup>™</sup>, em (HE et al., 2016) para seus mecanismos de busca por imagem.

O reconhecimento de padrões em imagens e vídeos pode ser considerado uma das aplicações da aprendizagem profunda mais utilizada nos últimos anos. Este fato se deve à grande evolução do ramo da visão computacional a partir do desenvolvimento das redes neurais chamadas Redes Neurais Convolucionais. Redes neurais convolucionais são modelos de aprendizagem profunda geralmente utilizados para aplicações de reconhecimento e classificação de padrões em imagens e vídeos. Estas redes visam simular o processo de reconhecimento de padrões em nosso cérebro a partir da criação de regiões especializadas em reconhecer certos tipos de características morfológicas das imagens, como linhas, círculos, etc.

A fundamentação teórica que inspirou a construção dessas redes neurais convolucionais veio a partir de um estudo de dois pesquisadores do laboratório de neurofisiologia da *Harvard Medical School*, (HUBEL; WIESEL, 1962), que utilizaram um gato para mapear diferentes regiões de seu *cortex* visual e comprovaram que o cérebro se divide em parte especialistas em reconhecer certos tipos de padrões geométricos.

Com base em seus experimentos, eles propuseram um modelo hierárquico composto de dois tipos de células neuronais, células S (simples) e células C (complexas). As células S são ativadas quando detectam formas básicas, como linhas e pontos em posições específicas, enquanto as células C combinam a ativação das células simples, reconhecendo padrões mais abstratos. A Figura 9 ilustra o experimento de (HUBEL; WIESEL, 1962), onde os eletrodos conectados ao cérebro do gato respondem com sinais diferentes a cada tipo de padrão geométrico apresentado na tela de projeção.

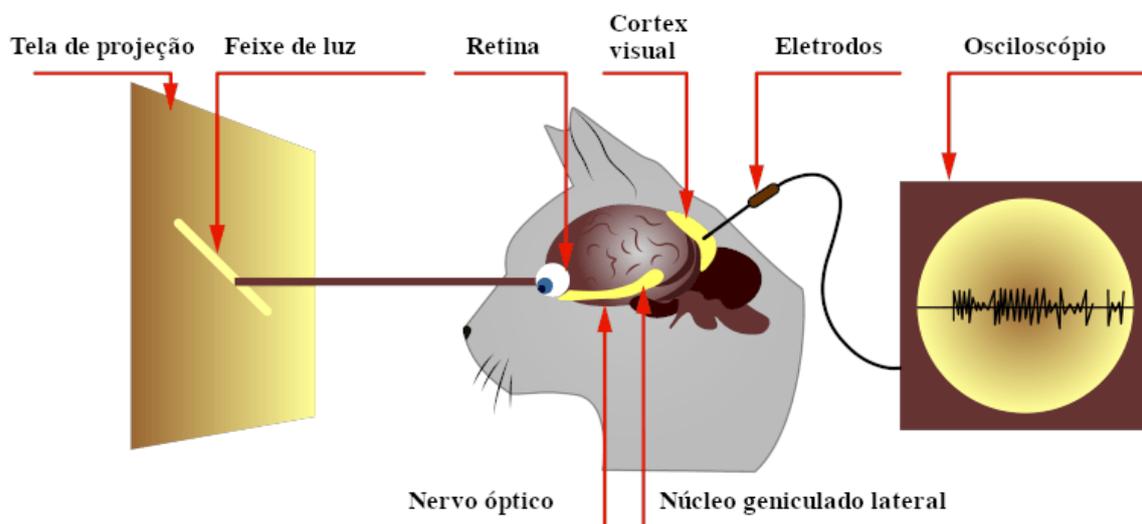


Figura 9: Experimento de Hubel e Wiesel [Fonte: (VEREMYOV, 2018)]

Uma rede neural convolucional é geralmente formada por: Uma camada de entrada que recebe as imagens, uma série de camadas que realizam operações de tratamento de imagens com mapeamento de características e uma última fase que contém uma rede neural de classificação, que recebe o mapeamento de características e fornece como saída o resultado da classificação. Dentre essas camadas de mapeamento de características, a três principais operações aplicadas são:

- **Convolução:** Operação baseada em um escaneamento de operações lineares, feita por uma matriz de características ao longo da imagem. Esta matriz de características carrega o formato do padrão a ser identificado de acordo com uma distribuição numérica de seus valores em sua extensão. Durante a operação, a imagem é varrida pela matriz de características, que vai “posição por posição” avaliando o grau de proximidade da região analisada com o padrão presente na matriz. As operações lineares resultam em uma nova matriz, menor do que a imagem, que contém o mapa de similaridade das regiões da imagem com o padrão analisado.
- **Max-Pooling:** Nesta operação, o foco está em reduzir as dimensões dos mapas de características, aproveitando somente as regiões que obtiveram maior afinidade com o padrão analisado pela matriz de convolução. Durante a operação, uma pequena matriz de *max-pooling*, geralmente 2x2 e sem valores, é passada pelo mapa gerado

na convolução. A cada iteração, o maior valor dentro do espaço 2x2 é extraído para compor o próximo mapa de características. Desta forma, pode-se assegurar que somente as maiores afinidades estão sendo consideradas.

- *ReLU*: As camadas chamadas de ReLU, ou *Rectified Linear Units* aplicam uma operação que visa reduzir a linearidade das operações da rede neural convolucional, pois sabe-se que modelos baseados somente em operações lineares podem apresentar problemas de aprendizagem. A função de ativação ReLU se assemelha a uma operação matemática de módulo, onde somente os valores negativos são afetados. No entanto, na ReLU os valores negativos são transformados em nulos, ou seja, qualquer valor de *pixel* negativo que por ventura tenha aparecido nas operações de convolução e *max-pooling*, será convertido para zero.

Além dessas camadas, as redes convolucionais contém com uma camada chamada de totalmente conectada, que nada mais é do que uma rede neural comum que recebe o *output* dos mapas de características e o analisa para dar o veredito final sobre a classificação.

### 4.2.1 Transferência de conhecimento

Aplicações como esta, exigem redes neurais com arquiteturas altamente complexas, fato que se torna um grande empecilho para que desenvolvedores independentes possam desenvolver modelos como estes. Para contornar este problema, foi desenvolvida uma técnica denominada aprendizado por transferência, conhecida como *Transfer Learning*, onde parte-se de um “esqueleto” de rede neural convolucional já treinado a reconhecer certos tipos de padrões e todo o trabalho se concentra em ajustes para adequar o modelo à aplicação em análise.

No aprendizado por transferência, escolhe-se uma rede convolucional previamente treinada a reconhecer padrões similares ao da aplicação em estudo. A partir deste ponto, deve-se remodelar apenas a camada totalmente conectada da rede escolhida para que esta seja capaz de classificar os dados da aplicação.

Para atingir este resultado, as imagens de treinamento alimentam as camadas de criação de mapas de características da rede original e os mapas resultantes são inseridos em uma rede neural totalmente conectada com pesos inicializados aleatoriamente. Inicia-se então um procedimento de treinamento comum da rede totalmente conectada. Ao final

desde treinamento, tem-se uma nova rede convolucional pronta para classificar os dados de uma nova aplicação.

A Figura 10 ilustra o processo de treinamento de uma rede via aprendizado por transferência, onde uma rede previamente treinada a realizar a tarefa *A* é utilizada em um novo modelo e passa por um treinamento para aprender a realizar a tarefa *B*.

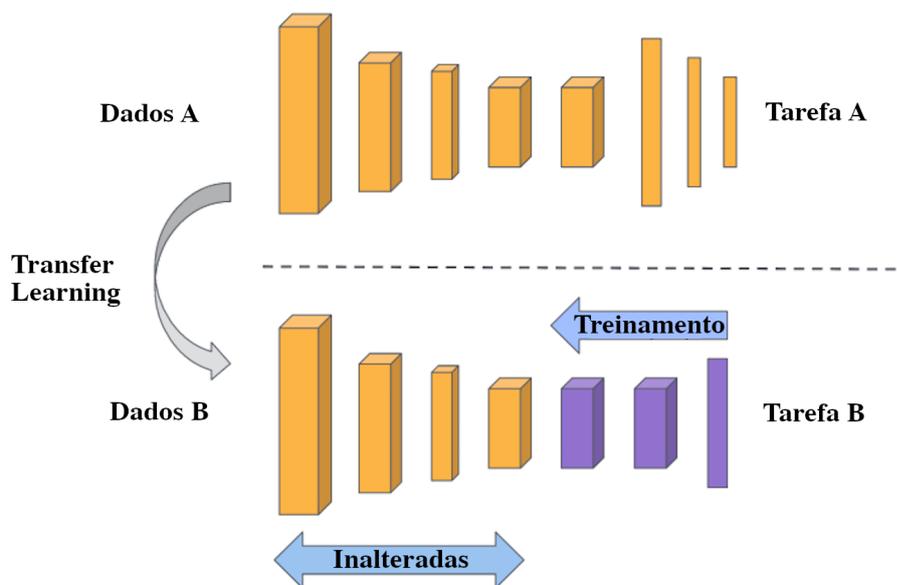


Figura 10: Ilustração do aprendizado por transferência [Fonte: (MALGONDE, 2018)]

Neste trabalho, uma rede neural convolucional foi ajustada através do aprendizado por transferência para reconhecer e classificar nódulos cancerígenos em imagens de tomografia computadorizada de pulmões. A rede escolhida para a aplicação do aprendizado por transferência é a *U-Net*, uma rede especializada em reconhecimento de padrões em imagens médicas.

### 4.2.2 Rede neural *U-Net*

A rede neural convolucional *U-Net* foi criada para trabalhos de segmentação em imagens médicas. Suas principais aplicações são reconhecimento de padrões em imagens de corrente sanguínea e tomografia. Seus mapas de características são otimizados para as morfologias mais comuns nesses tipos de imagens. Sendo assim, esta se tornou a candidata perfeita para ser a base tecnológica deste trabalho. A Figura 11 mostra a arquitetura da rede *U-Net*.

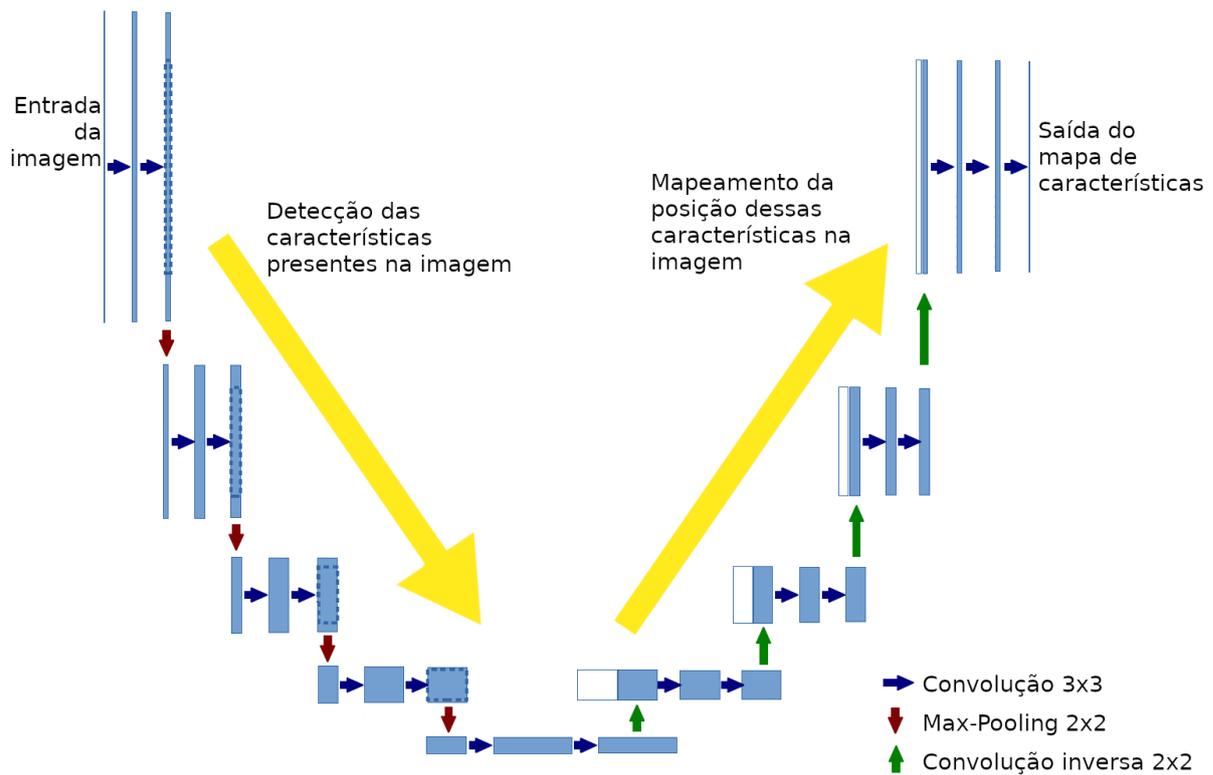


Figura 11: Arquitetura da U-Net [Fonte: (OLAF; PHILIPP; THOMAS, 2015)]

Esta rede processa as imagens em 2 fases, a obtenção das características; nesta primeira fase a imagem é escaneada com diversos filtros convolucionais para detectar o grau de pertinência da imagem com as características buscadas; e o posicionamento das características; nesta segunda fase, os mapas de característica são estendidos de volta ao tamanho original para ter como resultado um mapa final que consiga dizer o que existe na imagem e aonde está cada característica.

Em resumo, a primeira fase se preocupa em encontrar o que existe na imagem e a segunda em dizer aonde estão essas características na imagem. A rede utiliza operações de convolução com filtros 3x3, *ReLU*, *Pooling* 2x2 e convolução inversa 2x2, que trata de uma operação capaz de expandir o tamanho do mapa mantendo a proporção de distâncias entre características.

### 4.3 Considerações finais

Neste capítulo foram discutidos os principais tópicos a cerca da inteligência artificial e como ela é utilizada em aplicações reais atualmente. Foram apresentas com mais

aprofundamento as duas principais técnicas utilizadas neste trabalho, a aprendizagem profunda e o aprendizado por transferência, exemplificando a fundamentação teórica de cada método e introduzindo também a rede *U-Net*, rede neural convolucional protagonista deste estudo para a utilização do aprendizado por transferência.

No próximo capítulo serão apresentadas em detalhe a solução proposta para o problema de classificação de nódulos cancerígenos pulmonares e a metodologia técnica aplicada durante o desenvolvimento do trabalho.

# Capítulo 5

## ASPECTOS DE IMPLEMENTAÇÃO

**N**ESTE capítulo, são apresentadas todos os passos aplicados no desenvolvimento deste trabalho. São descritos os procedimentos utilizados para a obtenção das imagens médicas, bem como todos os processos de tratamento dos dados para a preparação da aplicação dos métodos de inteligência artificial. Em seguida são apresentadas as técnicas utilizadas para criar, treinar e testar o modelo classificador de nódulos cancerígenos pulmonares.

### 5.1 A metodologia

Esta aplicação trata de uma nova abordagem para o procedimento de detecção de nódulos pulmonares e visa avaliar o desempenho do treinamento de redes neurais convolucionais com inteligência de enxame, comparando-o com o treinamento realizado com algoritmos como o Gradiente Descendente.

Em resumo, a implementação desta abordagem começa pela procura de um banco de dados de imagens médicas de tomografia computadorizada de pulmões, classificados quanto à presença de tumores. Após a obtenção desses dados, foram implementadas rotinas de organização e limpeza para possibilitar a aplicação de *scripts* de leitura e tratamento. Foram aplicadas também técnicas de ampliação do volume de dados, para possibilitar um aprendizado de máquina mais sólido e robusto. Com os dados prontos, uma rede neural convolucional utilizada para classificar imagens médicas foi escolhida para a aplicação do aprendizado por transferência. Em seguida os algoritmos de inteligência de enxame foram definidos e estruturados para serem aplicados durante o treinamento. Após esta fase todos os passos necessários foram concretizados e a aplicação pode ser

desenvolvida. A Figura 12 apresenta graficamente os passos percorridos durante toda a implementação.



Figura 12: Resumo do fluxo de processos desenvolvidos na aplicação

## 5.2 Obtenção das imagens de tomografia

Após a conclusão da pesquisa bibliográfica e da apuração dos melhores conjuntos de dados para a aplicação, o banco de imagens médicas LIDC-IDRI, *Lung Image Database Consortium* foi escolhido para treinar o modelo classificador. Este banco de dados é fornecido pelo *Cancer Imaging Archive*, instituição norte americana especializada em armazenar dados de exames radiológicos com o intuito de fomentar a pesquisa por melhores técnicas de detecção e tratamento dos diversos tipos de câncer. O banco escolhido contém cerca de 250.000 imagens médicas de exames de tomografias computadorizadas de pulmões saudáveis e cancerígenos. As imagens deste banco de dados foram analisadas por quatro radiologistas especialistas e os diagnósticos dados por cada um deles estão contidos em arquivos de anotações, que são fornecidos juntos às imagens. Os dados deste banco foram coletados em um projeto que durou cerca de 8 anos e analisou milhares de pacientes. Como resultado, este projeto gerou cerca de 125 *gigabytes* de informação, que são disponibilizados publicamente na página do *Cancer Imaging Archive*.

Além do LIDC-IDRI, foi utilizado também um outro banco de dados complementar. O banco LUNA16, *Lung Nodule Analysis*, é um banco de imagens gerado a partir do LIDC-IDRI, sendo um subconjunto escolhido a partir de um trabalho de prospecção focado em escolher apenas imagens que possuíam informações completas sobre os principais atributos analisados no LIDC-IDRI. Sendo assim, o banco LUNA16 é formado apenas pelas imagens do LIDC-IDRI mais ricas em informação, para garantir um treinamento mais eficiente para o modelo classificador. É importante ressaltar que as imagens do banco fazem parte de

um conjunto de dados já classificados, o que permite um aprendizado supervisionado de alta confiança, tendo em vista que 4 especialistas diferentes deram seus diagnósticos para cada imagem do banco.

Cada exame gera uma série de imagens em seções, que varrem os pulmões de cima a baixo e fornecem um conjunto de imagens sequenciais que possibilitam uma análise tridimensional dos órgãos. Essas imagens são capturadas e armazenadas utilizando o protocolo DICOM, que estipula padrões para que as imagens possam carregar também informações sobre dados do paciente, do exame, de diagnósticos, observações, etc. A Figura 13 mostra um exemplo de imagem DICOM de um dos exames de tomografia, nela é possível ver parte das informações inseridas no formato do protocolo.

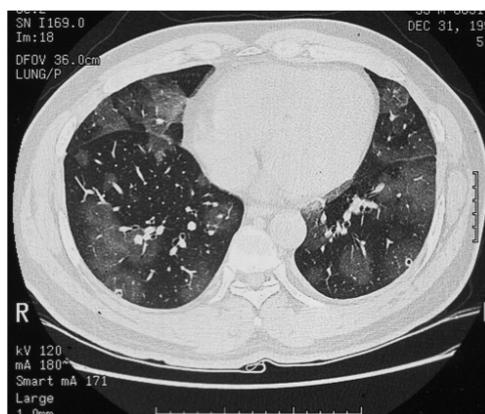


Figura 13: Imagem DICOM de exame de tomografia computadorizada de pulmão [Fonte: (KIM; AHN, 2007)]

### 5.3 Pré-processamento dos dados

A primeira rotina de pré-processamento dos dados foi a organização das imagens e a criação de listas de arquivos. Para isto foram aplicados scripts em *Python* para ler os arquivos DICOM e relacionar as imagens com as anotações, criando dois tipos de arquivos. O conteúdo de cada imagem foi salvo em um arquivo do tipo *.raw*, um formato de armazenamento de imagens geralmente chamado de negativo digital pois carrega toda a informação sem nenhum tipo de compressão e, por isso, é indicado para aplicações como esta. Após a criação de um arquivo *.raw*, o conteúdo de *metadata* desta imagem, como anotações e diagnóstico, foi salvo em um arquivo do tipo *.mhd*. Arquivos *.mhd* são muito utilizados para armazenar anotações de imagens médicas. Por se tratar de um exame de

imagens tridimensionais, cada escaneamento é composto por cerca de 200 fatias e cada fatia é uma imagem quadrada de 512 pixels de tamanho.

No ramo da radiologia, é muito comum a utilização da unidade *Hounsfield Units* (HU), que mede a radiodensidade da tomografia em um ponto. Com esta grandeza, é possível detectar composições químicas diferentes e assim identificar partes como ossos, ar, água, tecidos, etc. Foram levantados então os valores HU dos principais componentes encontrados nos exames da aplicação para uma possível análise sobre a influência desses dados no treinamento da rede, estes valores, retirados de (TOGA; MAZZIOTTA, 2002), são apresentados na Tabela 1.

Tabela 1: Valores *Hounsfield* de alguns tecidos encontrados em tomografias

<b>Tecido</b>	<b>Valores HU</b>
Osso denso	1000
Músculo	50
Matéria branca	45
Matéria cinzenta	40
Sangue	20
Água	0
Gordura	-100
Pulmão	-200
Ar	-1000

A rede poderia ser treinada utilizando todos os pontos de cada imagem, porém isto iria aumentar exponencialmente o esforço computacional. Para otimizar a operação do modelo, foram geradas imagens baseadas em cortes quadrados de tamanho  $50 \times 50$  *pixels*, incidentes ao redor das coordenadas apontadas nas anotações do banco de imagens, vide Figura 14.

A este ponto, o conjunto de imagens trabalhado está pronto para alimentar o modelo, no entanto a proporção entre as classes ainda é um potencial problema para o treinamento. Ao todo, o conjunto de dados possui 551.065 anotações, sendo apenas 1.351 classificadas como nódulos cancerígenos. Diante desse fato, as anotações classificadas negativamente foram reduzidas aleatoriamente e foram aplicadas rotinas de *Data Augmentation* para gerar novos dados classificados positivamente. As rotinas aplicadas foram: rotação em 90 graus, rotação em 180 graus e inversão horizontal. Após estas operações a proporção entre classes ficou em torno de 80% negativas e 20% positivas,

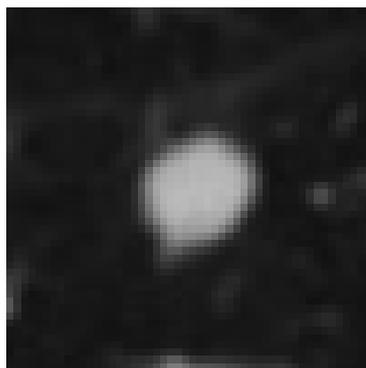


Figura 14: Corte feito ao redor das coordenadas apontadas pelos especialistas como promissoras.

o que pode ainda não ser ideal, porém aumentar demais a classe positiva poderá acarretar problemas de variância.

As últimas rotinas de pré-processamento envolveram o tratamento das imagens com um filtro de média  $3 \times 3$ , que é utilizado para reduzir a quantidade de ruído presente nas imagens. Filtros de média são matrizes quadradas de dimensão ímpar e elementos iguais entre si, com valores correspondentes ao inverso do número total de posições. Uma aplicação de filtros de média em redução de ruídos em imagens médicas é apresentada em (KOO; PARK, 1991). Outro filtro utilizado no tratamento das imagens foi um filtro *Laplaciano* de aguçamento. Filtros *Laplacianos* são utilizados para realçar regiões da imagem onde são encontradas maiores diferenças de cores entre *pixels* vizinhos (BURT; ADELSON, 1983). Com isso, estes filtros conseguem aumentar o detalhamento de regiões dentro de uma imagem, facilitando a detecção de objetos e padrões.

Com o balanceamento de classes, o pré-processamento dos dados foi finalizado. A Figura 15 exemplifica o processo de *data augmentation*, onde foram aplicadas as técnicas de rotação e inversão de imagens para ampliar o conjunto de dados.

## 5.4 Construção do modelo

Com a fase de pré-processamento dos dados concluída, iniciou-se a construção do modelo classificador da aplicação. Conforme apresentado no capítulo anterior, esta aplicação se baseia no uso da técnica de aprendizado por transferência, que faz uso de redes neurais complexas e previamente treinadas para simplificar o processo de desenvolvimento de novos modelos.

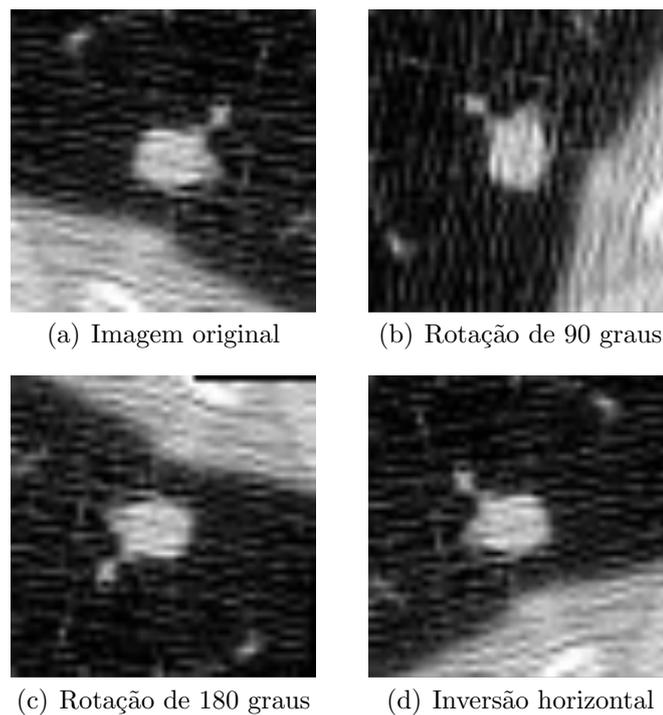


Figura 15: Transformações de *data augmentation*.

Para esta aplicação, duas redes foram cogitadas para a implementação do aprendizado por transferência, foram elas:

- ResNet: Rede Neural Convolutacional de alta complexidade, desenvolvida pela *Microsoft* para reconhecimento de padrões em imagens de temas diversos.
- U-Net: Rede Neural Convolutacional para segmentação de imagens biomédicas, desenvolvida para ser especialista em reconhecer os principais padrões em imagens de exames de imagem radiográfica e de sangue.

As redes foram cogitadas não somente por serem famosas nos ramos em que atuam, mas também por serem as duas mais utilizadas nos trabalhos relacionados que serviram de referência bibliográfica para esta aplicação. Sendo assim, seria possível realizar comparações mais sólidas entre abordagens de diferentes trabalhos.

De acordo com as referências bibliográficas, a escolha de redes treinadas a reconhecer padrões compatíveis com sua aplicação costuma produzir resultados mais satisfatórios. Levando em conta essa constatação e a escolha da maioria dos trabalhos relacionados, a U-Net foi escolhida como rede neural convolutacional base para este estudo.

Conforme apresentado no capítulo anterior, a U-Net é dividida em 2 fases, onde a primeira fase preocupa-se em encontrar o que existe na imagem e a segunda em dizer onde esses recursos estão na imagem. A rede usa operações de convolução com filtros de tamanho  $3 \times 3$ , funções de ativação ReLU e matrizes de *Max-pooling* de tamanho  $2 \times 2$ . No final do procedimento de mapeamento de características, a rede fornece um mapa de características pronto para ser analisado por uma série de camadas que compõem uma rede neural totalmente conectada.

Esta rede neural totalmente conectada é composta por 100 pesos, sendo assim ela lê os dados de entrada e gera um vetor de formato *Softmax* bidimensional, isto é, um vetor com o par  $(P, N)$  onde  $P$  e  $N$  são valores entre 0 e 1 e a soma  $P + N$  é igual a 1. Neste vetor,  $P$  representa a probabilidade de um nódulo ser cancerígeno e  $N$  de não ser, casos positivo e negativo respectivamente. Assim, a classificação será dada comparando os valores entre  $P$  e  $N$ , onde o maior valor será o vencedor e definirá a classificação.

A esta rede também foi adicionada uma camada de *Dropout*, uma técnica usada para controlar o problema de *overfitting*, causado pelo excesso de treinamento. O *Dropout* simplesmente zera uma certa quantidade de ativações de uma mapa com base em uma tomada de decisão probabilística, ou seja, caso um valor aleatório ultrapasse um valor predefinido. Essa operação simples e aparentemente contraditória é capaz de forçar a rede a ser redundante, ou seja, ela deverá ser capaz de classificar os dados corretamente mesmo sem contar com algumas de suas ativações em cada iteração. Desta forma, esta técnica se mostra uma boa aliada para combater o *overfitting*. A Figura 16 exemplifica esta operação.

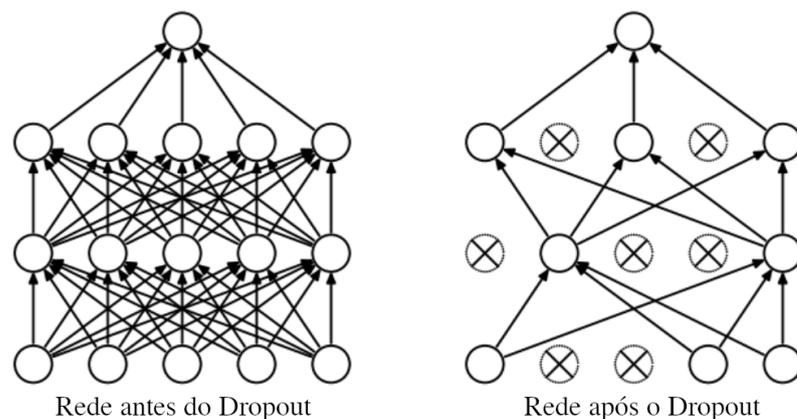


Figura 16: Operação de *Dropout* [Fonte: (SRIVASTAVA et al., 2014)]

## 5.5 Treinamento dos modelos

Devido à grande quantidade de dados necessários para o treinamento do modelo, os dados foram separados por lotes, já que o carregamento de todos os dados na memória da máquina é inviável. Para essa operação, os dados foram separados aleatoriamente para evitar problemas de lotes tendenciosos. O conjunto de treinamento possui 6878 imagens, todas devidamente classificadas pelos radiologistas especialistas.

Para as rotinas de treinamento, foi aplicada a validação *k-fold*, que divide os dados em  $k$  subconjuntos de mesmo tamanho. A partir dessa divisão,  $k - 1$  subconjuntos são utilizados para treinamento e o subconjunto restante para validação. Este processo é repetido  $k$  vezes, alternando de forma circular entre os subconjuntos. Ao final das  $k$  iterações o desempenho médio é calculado (KOHAVI, 1995). Neste trabalho, foi utilizada a validação *k-fold* com  $k$  igual a 5 e subconjuntos de 1375 imagens.

### 5.5.1 Treinamento com *backpropagation*

Com os conjuntos de treinamento e validação prontos para serem usados, o primeiro procedimento aplicado foi o treinamento sem inteligência de enxame, ou seja, usando *backpropagation* e Descida do gradiente. Para a escolha de hiper-parâmetros, como a taxa de aprendizado e o número de épocas de treinamento, foram estipulados valores possíveis para cada um deles e 100 experimentos foram executados para cada par de hiper-parâmetros.

Neste treinamento, as imagens são fornecidas ao modelo para análise e a calibração dos parâmetros da camada classificadora é feita através da otimização por descida do gradiente e pelo algoritmo *backpropagation*.

Nesta aplicação, os treinamentos com *backpropagation* serviram também para obter os hiper-parâmetros de taxa de aprendizagem e número de épocas, que foram fixados para os subsequentes treinamentos utilizando inteligência de enxame.

### 5.5.2 Treinamento utilizando algoritmos de enxame

Após o treinamento do modelo com o *backpropagation*, iniciou-se a rotina de implementação dos algoritmos de inteligência de enxame, onde estes foram selecionados, testados em funções de *benchmark* e preparados para serem utilizados no treinamento do modelo classificador de nódulos desta aplicação.

Conforme apresentado anteriormente, foram escolhidos 14 algoritmos de inteligência de enxame para testar a eficiência desta família de algoritmos no treinamento de redes neurais convolucionais para reconhecimento de nódulos cancerígenos de pulmão. Para a escolha deste algoritmos, foi levado em consideração principalmente o grau de aceitação da técnica perante a comunidade acadêmica. No entanto, alguns algoritmos mais recentes foram também selecionados para testar suas eficiências, como o GWO, introduzido em 2014.

É importante ressaltar que todos os algoritmos de enxame testados foram implantados com a topologia de estrela, em que todas as partículas são livres para se comunicar com quaisquer outras partículas.

Por serem técnicas de busca otimizada, os algoritmos de inteligência de enxame são usados para encontrar configurações ideais para valores de variáveis ou parâmetros de modelos. Desta forma, pode-se simular diversas configurações possíveis para os parâmetros de uma rede neural e, com um enxame de partículas, testar e encontrar a configuração que produz o melhor resultado de classificação.

Nesta aplicação, cada partícula do enxame é um vetor composto de 100 coordenadas, onde cada coordenada representa um peso da rede neural de classificação do modelo de aprendizagem profunda. Sendo assim, são executados experimentos de treinamento com cada partícula do enxame e seus desempenhos em relação ao erro médio quadrático são medidos. Após todo o enxame ser avaliado, são aplicadas as lógicas de inteligência de enxame de cada algoritmo para atualizar os pesos de cada partícula. Com a atualização dos pesos, inicia-se uma nova época de treinamentos. Este procedimento é mantido até que se atinja o critério de parada de treinamento. A Figura 17 apresenta uma ilustração explicativa sobre o treinamento de redes neurais com algoritmos de inteligência de enxame.

Cada um dos algoritmos de otimização de enxame listados foi utilizado para treinar e testar o modelo classificador em 100 experimentos repetidos. Nesses experimentos, foram monitorados 4 indicadores de desempenho (precisão, precisão, sensibilidade, especificidade), utilizando o erro quadrático médio como a função de avaliação da qualidade de cada partícula.

Além disso, para avaliar possíveis melhorias no desempenho do modelo devido a mudanças nos parâmetros da rede *U-Net*, os experimentos também foram executados

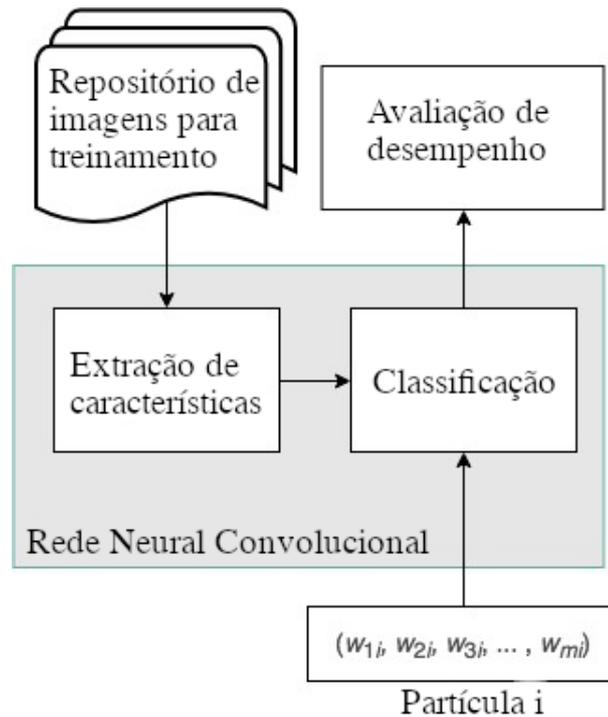


Figura 17: Treinamento de redes neurais com algoritmos de enxame

utilizando versões da *U-Net* com diferentes tamanhos de matrizes de *Pooling*:  $3 \times 3$ ,  $4 \times 4$  e  $5 \times 5$ . Também foram realizados experimentos utilizando a função tangente hiperbólica ao invés vez da *ReLU* como função da ativação da rede convolucional.

## 5.6 Considerações finais

Neste capítulo foram discutidos todos os passos referentes ao desenvolvimento e à aplicação do modelo de rede neural convolucional classificadora de nódulos pulmonares desta aplicação. Foram apresentadas as rotinas aplicadas para tratar os dados de treinamento, bem como as técnicas utilizadas para ampliar a quantidade de imagens e balancear as classes de análise.

Em seguida foram exemplificados os passos seguidos durante a construção do modelo, onde as redes cogitadas para o aprendizado por transferência foram apresentadas, os principais parâmetros das redes foram passados e testados, e as rotinas de treinamento foram explicadas. Foram apresentados também os detalhes e parâmetros utilizados na implementação de cada algoritmo de enxame, explicitando também quais métricas de desempenho foram medidas ao logo dos experimentos.

No próximo capítulo serão apresentados e discutidos todos os resultados provenientes das rotinas de treinamento e validação dos modelos desenvolvidos, comparando o desempenho de cada algoritmo e analisando a real influência do treinamento de redes neurais com algoritmos de enxame para aplicações como esta.

# Capítulo 6

## RESULTADOS DE DESEMPENHO

**N**ESTE capítulo, são apresentados e analisados os resultados alcançados com a metodologia aplicada no desenvolvimento deste trabalho. Os resultados começam pela obtenção de hiper-parâmetros para o treinamento da rede com o algoritmo *Backpropagation*. Em seguida, são apresentados os desempenhos de cada técnica de treinamento com inteligência de enxame. Ao final do capítulo, os resultados são analisados para uma definição sobre a melhor técnica aplicada.

### 6.1 Métricas de avaliação

Para a análise de performance dos modelos desenvolvidos, foram calculadas 4 métricas de desempenho para cada experimento executado. A primeira métrica computada é a acurácia, que representa a taxa de acertos totais do modelo. Esta métrica é uma das mais importantes, pois transmite uma noção geral do desempenho do modelo em qualquer situação. No entanto, para conjuntos de classes muito desbalanceados, esta métrica pode não ser eficiente sozinha. O cálculo da acurácia é apresentado pela Equação 1:

$$A = \frac{100 * (TP + TN)}{TP + TN + FP + FN}, \quad (1)$$

onde  $TP$ , é o número de classificações positivas acertadas,  $TN$ , o número de classificações negativas acertadas,  $FP$ , o número de classificações positivas equivocadas e  $FN$ , o número de classificações negativas equivocadas.

Outra métrica importante é a precisão, que calcula, dentre os casos classificados como câncer, qual porcentagem realmente tem câncer. O cálculo da precisão é apresentado pela Equação 2:

$$P = \frac{100 * TP}{TP + FP}, \quad (2)$$

onde  $TP$  e  $FP$  seguem as mesmas definições apresentadas na Equação 1.

Além da acurácia e precisão, foi medida também a especificidade de cada modelo. Esta métrica é focada em testar a capacidade do modelo em reconhecer casos sadios. Ela é calculada pela relação entre os casos classificados como sadios e os casos realmente sadios. O cálculo da especificidade é apresentado pela Equação 3:

$$E = \frac{100 * TN}{TN + FP}, \quad (3)$$

onde  $TN$  e  $FP$  seguem as mesmas definições apresentadas na Equação 1.

Por fim, a sensibilidade do modelo é calculada. Esta métrica é análoga à especificidade para os casos positivos, em outras palavras, dos casos reais de câncer, quantos por cento foram classificados como câncer. O cálculo da sensibilidade é apresentado pela Equação 4:

$$S = \frac{100 * TP}{TP + FN}, \quad (4)$$

onde  $TP$  e  $FN$  seguem as mesmas definições apresentadas na Equação 1.

## 6.2 Obtenção de hiper-parâmetros da rede

Com a estrutura o modelo classificador de nódulos pronta para aplicação e com a separação do conjunto de dados em aproximadamente 75% para treinamento e 25% para validação e teste, os primeiros experimentos visaram encontrar um par de hiper-parâmetros (taxa de aprendizado, número de épocas) ideal para o treinamento das camadas da rede, utilizando o *Backpropagation*. Com base nesses hiper-parâmetros, foi possível também estabelecer uma padronização para as comparações entre os tipos de treinamento com e sem inteligência de enxame, detalhada na seção 6.3.

Para este primeiro teste foram selecionados 4 possíveis valores de taxa de aprendizado: 0,0001; 0,001; 0,01; 0,1; e 4 opções para a quantidade de épocas de treinamento: 30, 100, 300 e 1000. Cada experimento foi executado por 100 repetições e seus desempenhos nas métricas de acurácia, precisão, sensibilidade e especificidade foram medidos e avaliados. A Tabela 17 do Apêndice A apresenta os resultados médios

obtidos para cada métrica e cada par de hiper-parâmetros, onde foram atribuídos legendas denominadas “casos” para simplificar a visualização dos dados na Figura 18.

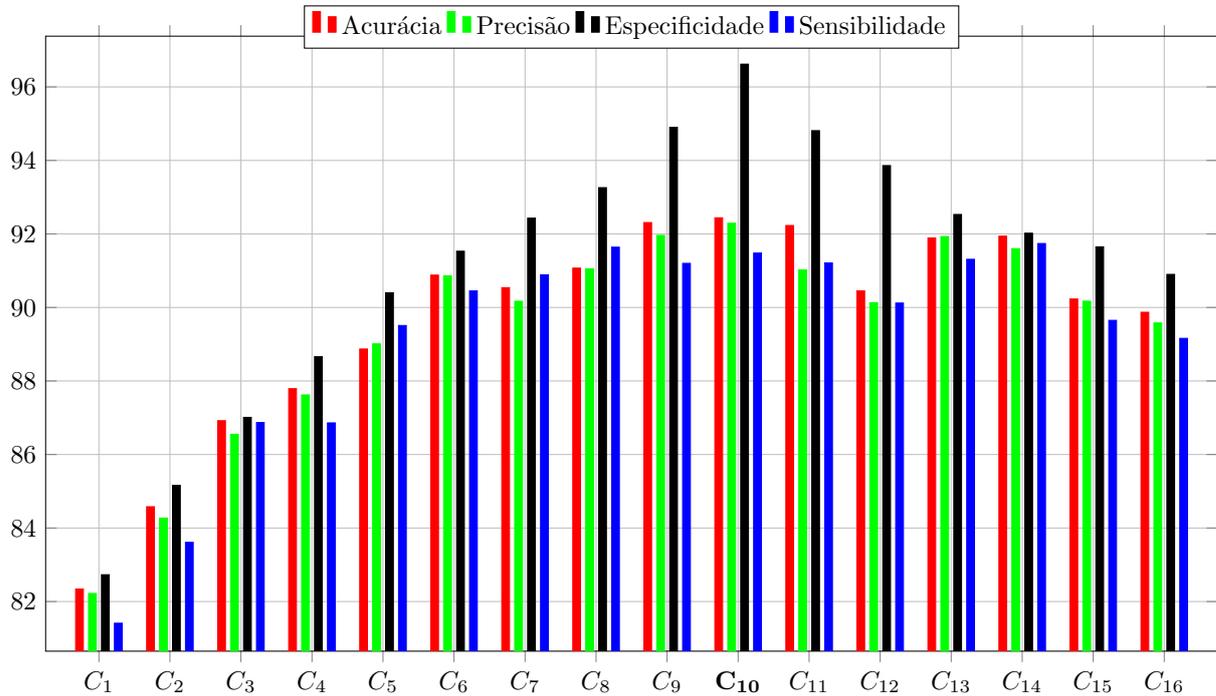


Figura 18: Desempenho dos pares de hiper-parâmetros analisado

Após os experimentos realizados no modelo que utilizou *backpropagation*, pode-se observar com uma análise detalhada, que os pesos obtidos nas camadas totalmente conectadas possuíam valores médios entre  $-2$  e  $2$ . Esta faixa de valores se mostrou uma boa métrica para estipular o espaço de busca dos algoritmos de enxame. Para obter uma boa abrangência de pontos, optou-se por uma granularidade de  $0,1$ , o que garante um espaço de 41 pontos dentro da faixa de busca para cada peso.

A Figura 18 apresenta graficamente os dados numéricos da Tabela 17, fornecendo uma análise visual onde é possível inferir uma superioridade de desempenho nos casos  $C_9$ ,  $C_{10}$  e  $C_{11}$ . Para validar esta hipótese e eleger o melhor par de hiper-parâmetros, foram realizados Testes-T de *Student*, onde os resultados aparecem na Tabela 2.

Embora os resultados do Teste-T apresentados na Tabela 2 sejam maiores do que o valor padrão de  $0,05$ , para  $95\%$  de confiança, quando comparado com os outros dois casos, o caso  $C_{10}$  apresenta resultados probabilísticos suficientemente bons para ser elegido como o par de hiper-parâmetros do modelo classificador desta aplicação. Com esta inferência, o modelo treinado com *backpropagation* foi estruturado para operar em 300 épocas de treinamento, com uma taxa de aprendizagem de  $0.001$ .

Tabela 2: Teste-T de *Student* dos casos  $C_9$ ,  $C_{10}$  e  $C_{11}$  para 95% de confiança.

Primeiro caso	Segundo caso	Resultado entre os casos
$C_{10}$	$C_{11}$	0,053
$C_{10}$	$C_9$	0,097
$C_9$	$C_{11}$	0,152

## 6.3 Resultados utilizando técnicas de enxame

Nesta seção, são apresentados os parâmetros utilizados por cada algoritmo de enxame e são avaliados os comportamentos e desempenhos de cada modelo treinado com estes algoritmos.

### 6.3.1 Parâmetros das técnicas de enxame

Considerando a rede neural totalmente conectada do modelo com seus 100 pesos, cada partícula de um enxame é construída como um vetor de 100 coordenadas, onde cada coordenada corresponde a um peso da rede. Dessa forma, cada partícula representa uma configuração de parâmetros da rede. Para o treinamento da rede com os algoritmos de enxame, foi estipulado que cada enxame iria possuir 300 indivíduos e cada experimento iria testar até um máximo de 300 iterações.

Para determinar o número de iterações de cada experimento, foi adicionado ao treinamento um mecanismo de parada antecipada, que monitora o valor do erro de validação do modelo ao fim de cada iteração. Caso o este valor de erro aumente seguidamente por duas iterações, o treinamento é interrompido. Este mecanismo é utilizado para evitar o excesso de treinamento, conhecido como *overfitting*, que pode prejudicar o desempenho do modelo. A número médio de iterações de treinamento de cada algoritmo testado é apresentado na Figura 19, onde o PSO convergiu mais rápido que os demais e o BFO foi o mais lento.

Além desses parâmetros universais, cada algoritmo de enxame possui seus próprios parâmetros específicos. Para determinar os valores de cada parâmetro, foram pesquisados na literatura os valores mais utilizados e recomendados. Diante desses valores, para cada algoritmo, as configurações de parâmetros recomendadas foram testadas por 30 repetições em 4 funções de referência. Os melhores resultados elegeram as melhores configurações

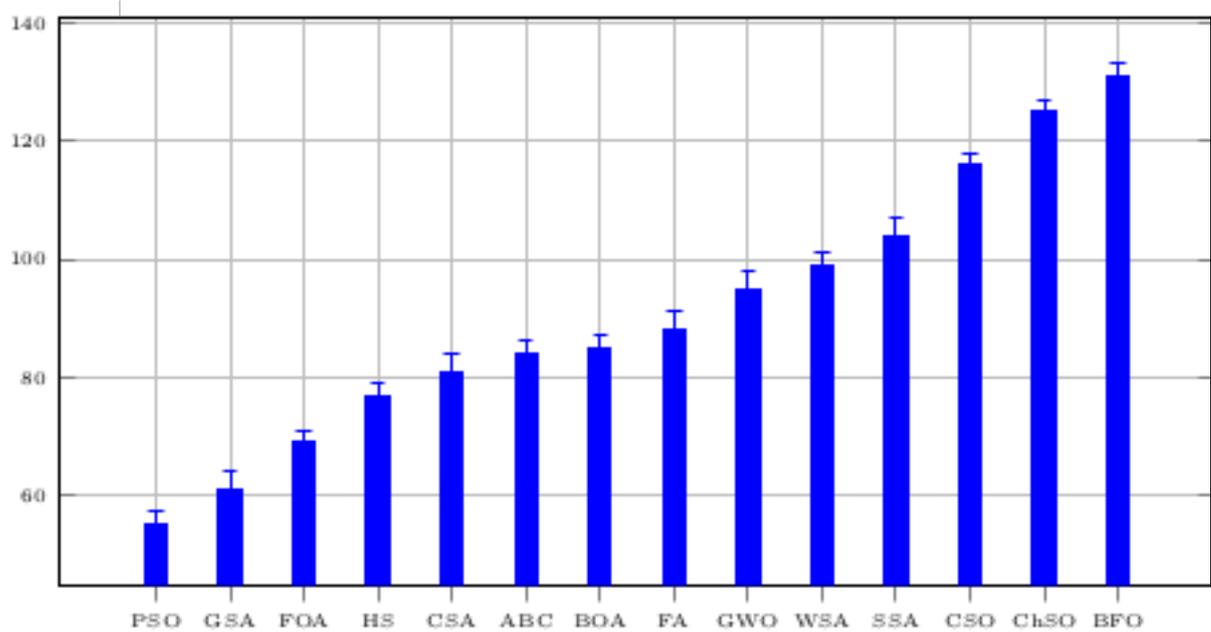


Figura 19: Quantidade de iterações de treinamento

de parâmetros para cada técnica desta aplicação. A seguir estão as descrições e valores de cada parâmetro usado pelos algoritmos testados.

No algoritmo BFO, os parâmetros utilizados são os passos de quimiotaxia, que determinam o movimento das bactérias, a distância máxima de navegação, o tamanho de cada passo e a probabilidade de eliminação de uma bactéria durante a operação.

Tabela 3: Parâmetros do BFO

Parâmetros	Valores
Passos de quimiotaxia	5
Distância de navegação	2
Tamanho do passo	1
Probabilidade de eliminação	0,05

No BOA, os parâmetros utilizados são o nível de emissão de pulso, o volume do som emitido, as frequências mínimas e máximas detectadas, a taxa de variação de volume (Alfa) e a alteração de pulso. No FOA, os parâmetros utilizados são o número de fogos de artifício normais e o número de fogos de artifício *Gaussianos*. No FA, os parâmetros utilizados são índice de atração mútua, absorção de luz, 2 parâmetros estocásticos denominados  $\alpha_1$  e  $\alpha_2$  e 2 parâmetros para ajustar a curva *Gaussiana*.

No WSA, os parâmetros utilizados são a intensidade do ultrassom na origem e a probabilidade de distorção da mensagem. No CSA, os parâmetros utilizados são a probabilidade de detecção de ovos e o número de ovos. No HS, os parâmetros usados são a taxa de ajuste do tom, a taxa de consideração da harmonia e a largura de banda. No PSO, os parâmetros utilizados são o coeficiente de inércia, o coeficiente cognitivo e o coeficiente social.

Tabela 4: Parâmetros do BOA

Parâmetros	Valores
Nível de emissão de pulso	0,9
Volume do som	0,5
Frequência mínima	0
Frequência máxima	0,02
Variação de volume	0,9
Alteração de pulso	0,9

Tabela 5: Parâmetros do FOA

Parâmetros	Valores
Número de fogos normais	50
Número de fogos <i>Gaussianos</i>	50

Tabela 6: Parâmetros do FA

Parâmetros	Valores
Índice de atração mútua	1
Índice de absorção de luz	1
$\alpha_1$	1
$\alpha_2$	0,1
Primeiro parâmetro da <i>Gaussiana</i>	0
Segundo parâmetro da <i>Gaussiana</i>	0,1

No CSO, os parâmetros utilizados são o número de gatos caçadores, o coeficiente de memória de cada gato, a contagem de dimensões para mudança de modo, o intervalo de dimensões e os coeficientes W (limite superior do peso), C (limite superior do custo).

Tabela 7: Parâmetros do WSA

<b>Parâmetros</b>	<b>Valores</b>
Intensidade do ultrassom na origem	2
Probabilidade de distorção da mensagem	0,005

Tabela 8: Parâmetros do CSA

<b>Parâmetros</b>	<b>Valores</b>
Probabilidade de detecção do ovo	0,5
Número de ovos	100

Tabela 9: Parâmetros do HS

<b>Parâmetros</b>	<b>Valores</b>
Taxa de ajuste do tom	0,5
Taxa de consideração da harmonia	0,5
Largura de banda	0,5

Tabela 10: Parâmetros do PSO

<b>Parâmetros</b>	<b>Valores</b>
Coefficiente de inércia	0,49
Coefficiente cognitivo	0,72
Coefficiente social	0,72

No SSA, o único parâmetro utilizado, além dos parâmetros universais, é o parâmetro estocástico, mostrado na Tabela 11. Vale ressaltar que os algoritmos ABC, GSA e GWO utilizam apenas os parâmetros universais.

Tabela 11: Parâmetros do SSA

<b>Parâmetros</b>	<b>Valores</b>
Parâmetro estocástico	0,4

Tabela 12: Parâmetros do CSO

Parâmetros	Valores
Número de gatos caçadores	10
Coefficiente de memória	2
Contagem de dimensões para mudança de modo	1
Intervalo de dimensões	0,1
Limite superior do peso	0,1
Limite superior do custo	1,05

### 6.3.2 Testes e resultados

Outro aspecto levado em consideração para os experimentos foi o uso de diferentes funções de ativação e tamanhos de matriz de *Pooling* na rede convolucional. Dois tipos de funções de ativação foram escolhidas: ReLU e Tangente Hiperbólica (Tanh). Estas foram escolhidas por serem as mais comumente usadas em aprendizado de máquinas.

Da mesma forma, 4 tamanhos diferentes de matrizes de *Pooling* foram elegidas para os experimentos:  $2 \times 2$ ,  $3 \times 3$ ,  $4 \times 4$  e  $5 \times 5$ . Cada algoritmo de enxame foi então avaliado ao treinar o modelo sob cada uma das 8 circunstâncias, combinando uma função de ativação com um tamanho de matriz.

Assim como no caso do treinamento com *backpropagation*, cada algoritmo foi testado repetidamente por 100 experimentos em cada circunstância de arquitetura. A Tabela 18 apresenta os valores médios de acurácia obtidos durante os experimentos para cada algoritmo de enxame e para o modelo sem otimização de enxame.

Foram também estipuladas legendas  $S_1$  a  $S_{32}$  para referenciar as circunstâncias de treinamento com diferentes funções de ativação e matrizes de *Pooling*.

A partir desses resultados, pode-se observar uma ligeira superioridade do PSO em relação aos outros algoritmos quanto à acurácia média, estando GSA e HS imediatamente abaixo em desempenho. Esta métrica nos fornece uma estimativa global da taxa de acertos de cada modelo.

A Figura 20 apresenta os resultados de acurácia para todos os algoritmos testados, utilizando ambas as funções de ativação e o tamanho  $2 \times 2$  para a matriz de *max-pooling*.

Da mesma maneira, a Tabela 19 apresenta os valores médios de precisão obtidos durante os experimentos para cada algoritmo de enxame e para o modelo sem otimização de enxame.

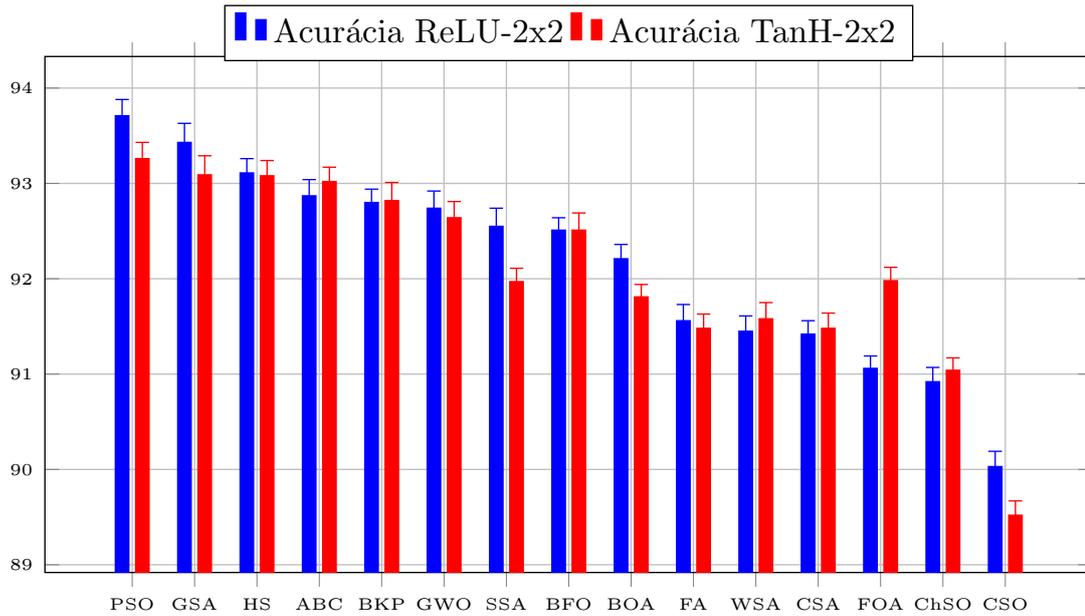


Figura 20: Acurácia média para todos os algoritmos

Com uma análise semelhante à realizada nos dados de acurácia, pode-se observar uma ligeira superioridade do GSA em relação à precisão média, estando PSO e HS imediatamente abaixo em desempenho. Esta métrica fornece uma estimativa da taxa de acertos dentre os casos classificados como cancerígenos.

A Figura 21 apresenta os resultados de precisão para todos os algoritmos testados, utilizando ambas as funções de ativação e o tamanho  $2 \times 2$  para a matriz de *max-pooling*.

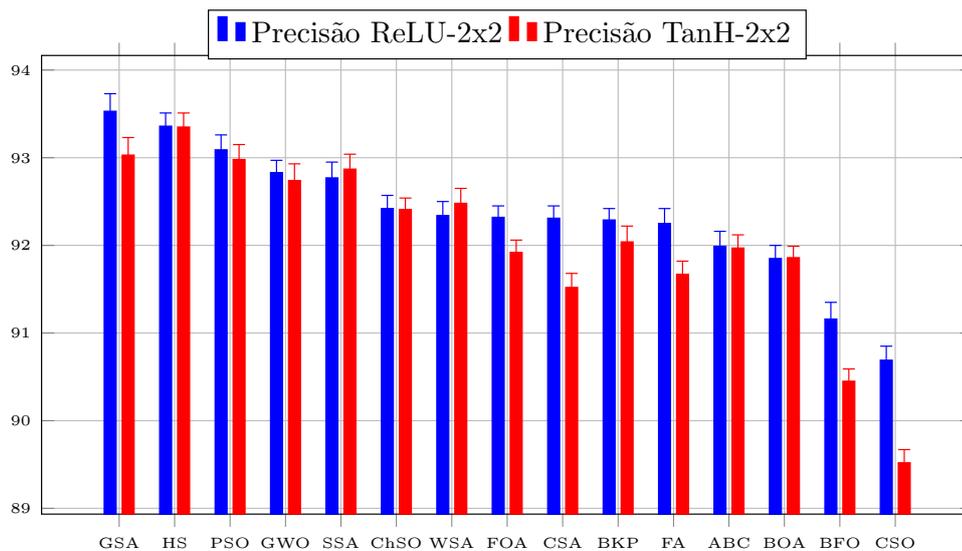


Figura 21: Precisão média para todos os algoritmos

A Tabela 20 apresenta os valores médios de sensibilidade obtidos durante os

experimentos para cada algoritmo de enxame e para o modelo sem otimização de enxame.

Em relação à sensibilidade média, pode-se observar novamente uma ligeira superioridade do PSO, estando GSA e HS imediatamente abaixo em desempenho. Esta métrica fornece uma estimativa da taxa de acertos dentre os casos realmente cancerígenos.

A Figura 22 apresenta os resultados de sensibilidade para todos os algoritmos testados, utilizando ambas as funções de ativação e o tamanho  $2 \times 2$  para a matriz de *max-pooling*.

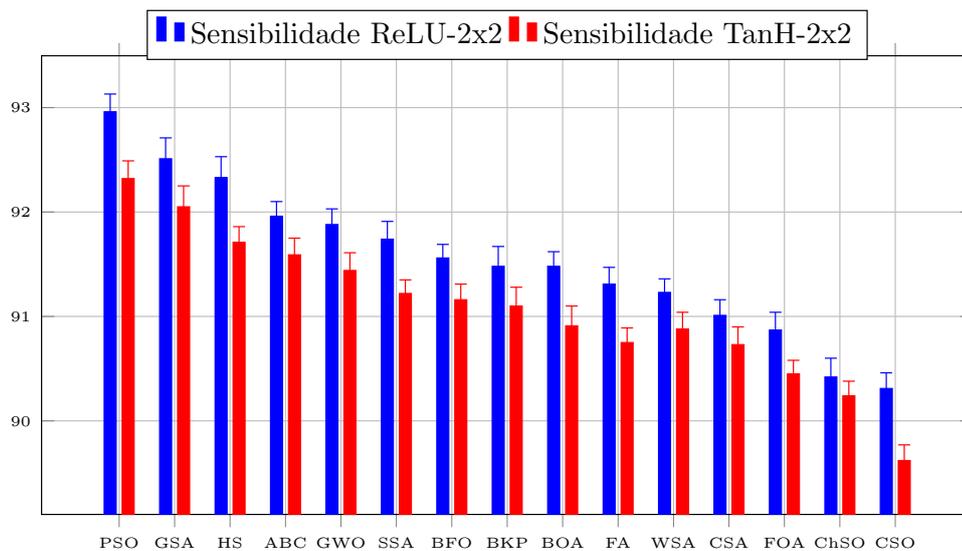


Figura 22: Sensibilidade média para todos os algoritmos

A Tabela 21 apresenta os valores médios de especificidade obtidos durante os experimentos para cada algoritmo de enxame e para o modelo sem otimização de enxame.

Por fim, em relação à especificidade média, pode-se observar uma ligeira superioridade do HS, estando GSA e PSO imediatamente abaixo em desempenho. Esta métrica fornece uma estimativa da taxa de acertos dentre os casos não cancerígenos.

A Figura 23 apresenta os resultados de especificidade para todos os algoritmos testados, utilizando ambas as funções de ativação e o tamanho  $2 \times 2$  para a matriz de *max-pooling*. Além das métricas percentuais de desempenho, o tempo médio de execução de cada experimento foi medido em minutos e está apresentado na Figura 24, onde o PSO foi o algoritmo mais rápido e o BFO o mais lento.

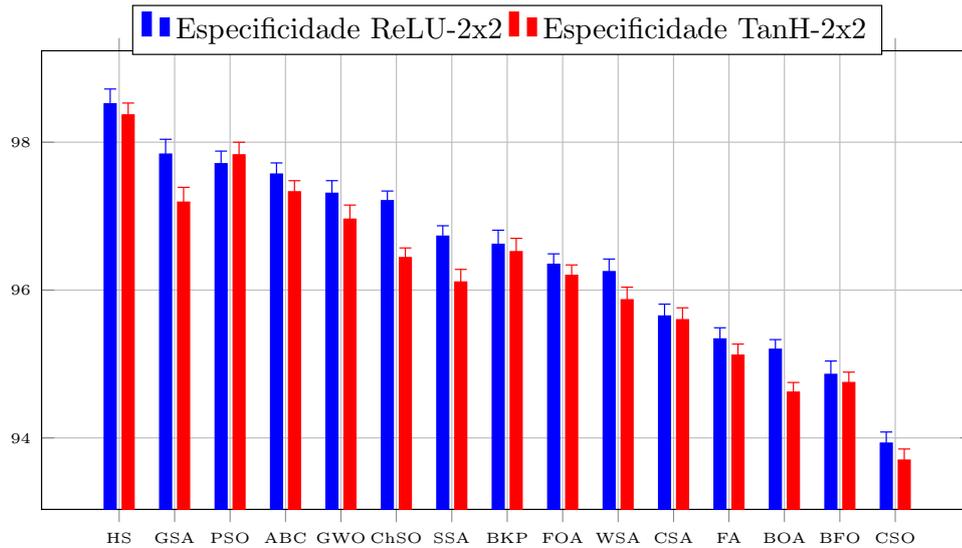


Figura 23: Especificidade média para todos os algoritmos

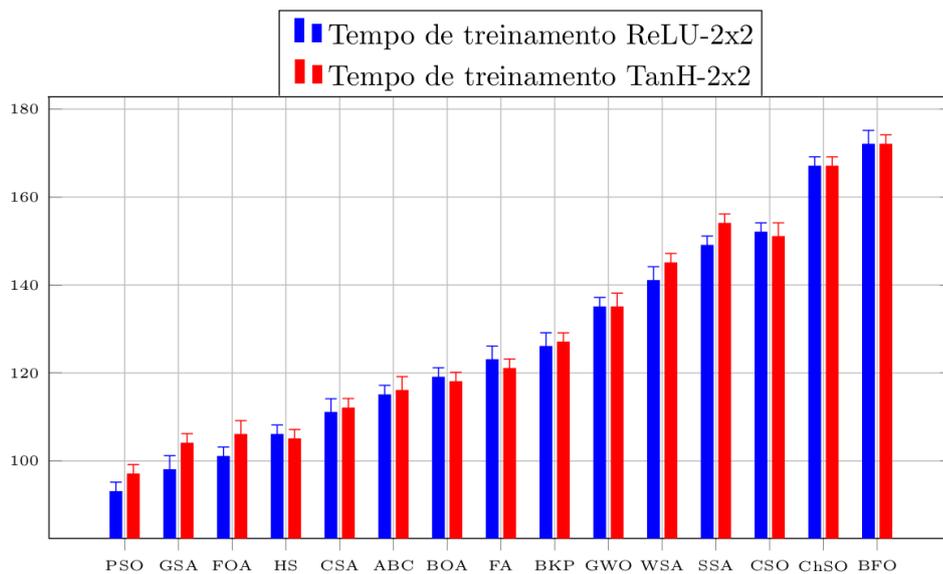


Figura 24: Tempo médio de treinamento

## 6.4 Análises de desempenho

Nesta seção são apresentadas as principais análises realizadas para inferir conclusões sobre os resultados obtidos.

Diante dos resultados obtidos durante os experimentos, pode-se verificar a real eficácia do uso de algoritmos de inteligência de enxame no treinamento de redes neurais convolucionais para a detecção de nódulos pulmonares, onde pelo menos 7 dos 14 algoritmos testados obtiveram performances superiores ao modelo treinado com

*Backpropagation* em todas as métricas de desempenho registradas. Em comparação com um dos trabalhos relacionados a este, foi possível observar o mesmo comportamento verificado em (JAYANTHI; PREETHA, 2016), onde um modelo de rede neural treinado com PSO para reconhecer câncer de mama obteve desempenho vastamente superior ao mesmo modelo treinado sem otimização de enxame.

Dentre os algoritmos testados, o PSO, o GSA e o HS obtiveram desempenhos de claro destaque, pois se mantiveram entre os 3 melhores modelos nas 4 métricas testadas. A partir destas constatações, foi iniciada uma análise específica com estes 3 algoritmos para inferir a melhor configuração de hiper-parâmetros de função de ativação e matriz de *Pooling*. Para esta análise, foram considerados os modelos treinados com PSO, GSA, HA, *backpropagation* (BKP) e os valores de média entre os 14 algoritmos utilizados.

As Figuras 25, 26, 27 e 28 apresentam, respectivamente, as análises detalhadas de acurácia, precisão, sensibilidade, especificidade e tempo de treinamento, para a comparação dos hiper-parâmetros testados.

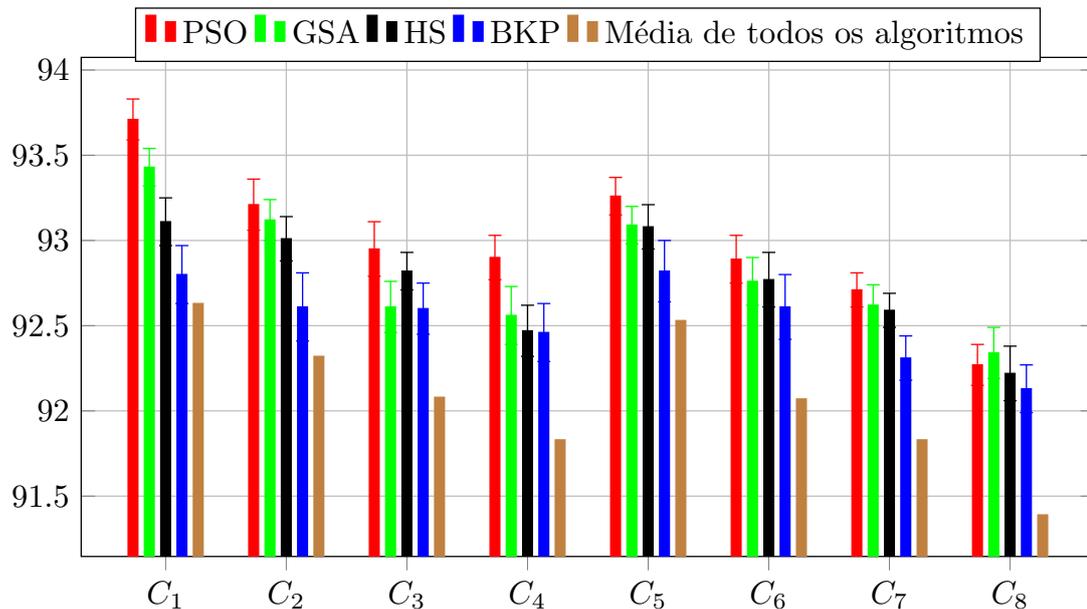


Figura 25: Acurácia média

Um dos primeiros fatores observados nesta análise foi a influência do tamanho das matrizes de max-pooling. Foi possível observar uma redução de desempenho à medida que o tamanho da matriz aumentou. Assim, pode-se concluir, para esta aplicação, que o aumento do tamanho da matriz de *Pooling* diminui o desempenho dos modelos. Esse comportamento pode estar relacionado ao aumento da perda de informação nos mapas de

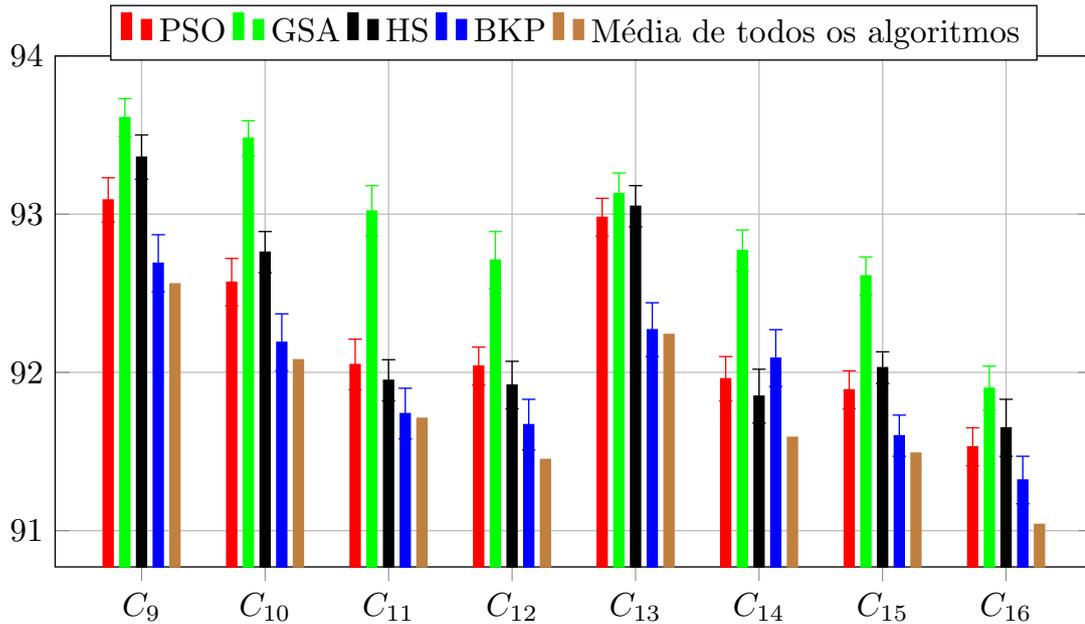


Figura 26: Precisão média

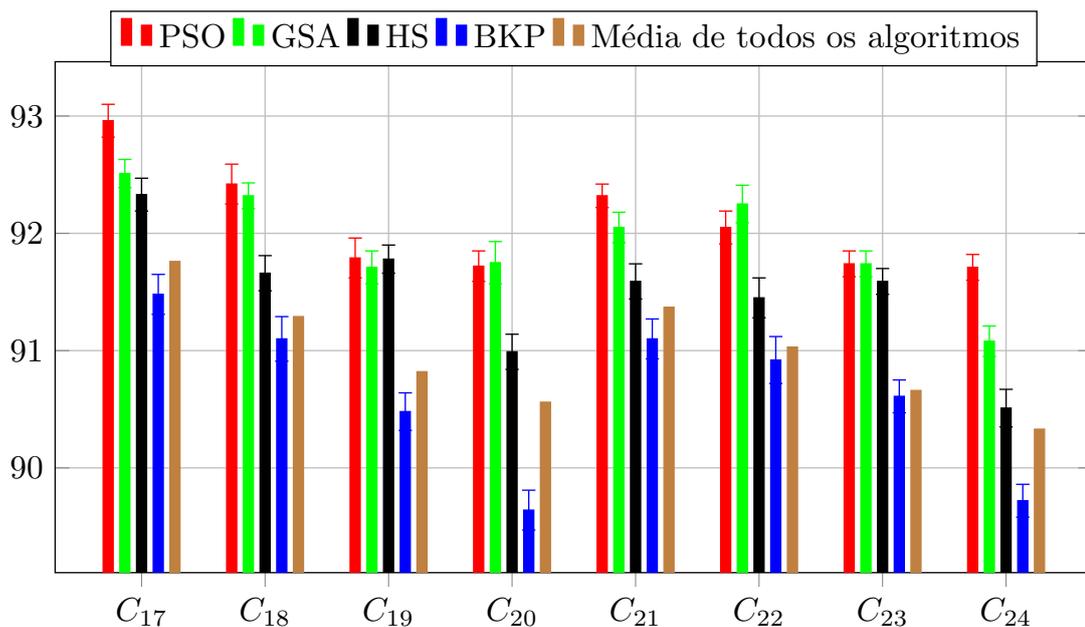


Figura 27: Sensibilidade média

características, pois quanto maior a matriz, mais *pixels* são removidos e menores são os mapas gerados.

Estes resultados mostram também que os modelos treinados com a função Tanh obtiveram desempenho levemente inferior aos modelos ReLU. Esse comportamento pode estar relacionado à relação entre o desempenho e a linearidade das operações matemáticas aplicadas em redes neurais convolucionais, pois é sabido que funções

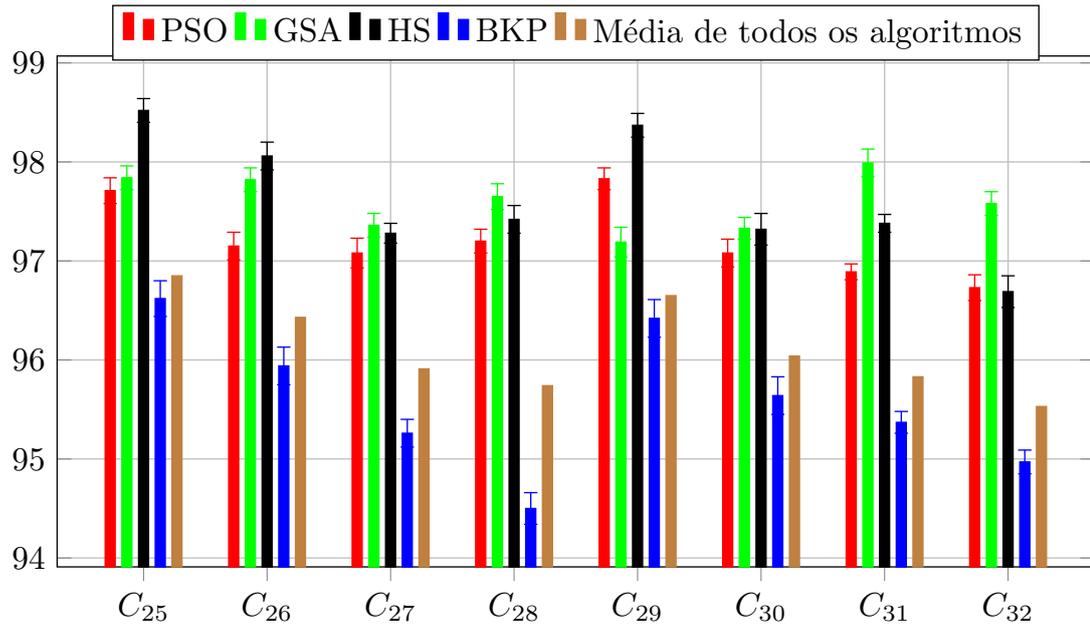


Figura 28: Especificidade média

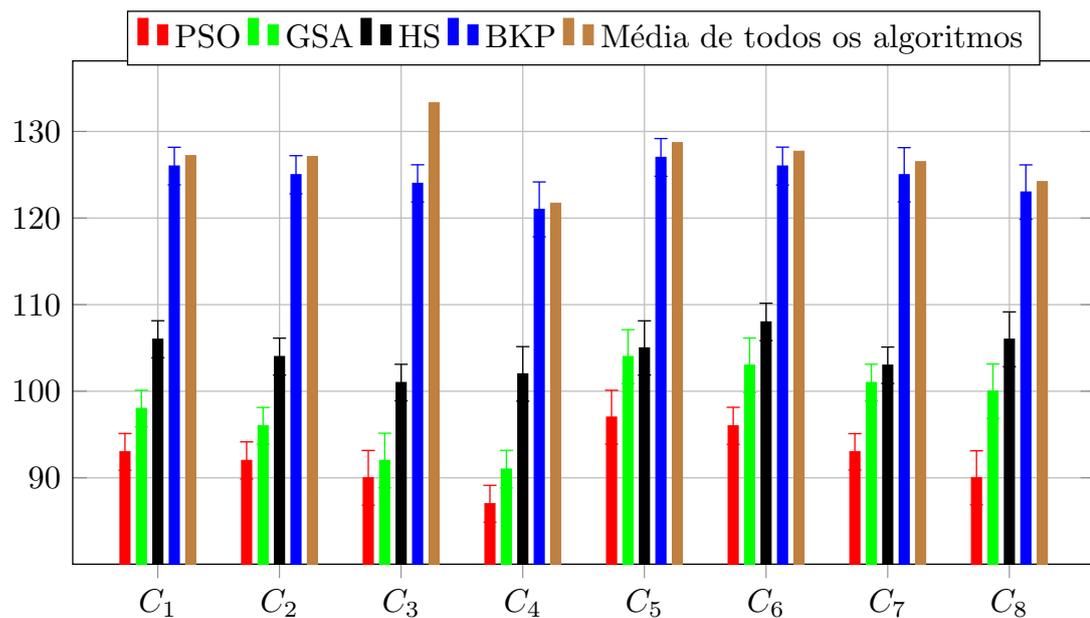


Figura 29: Tempo médio de treinamento

ReLU são normalmente utilizadas para quebrar parte da linearidade dessas operações, aumentando o desempenho do modelo.

A partir das Figuras 25, 26, 27 e 28, é possível visualizar a superioridade de desempenho dos modelos ReLU sobre os de Tanh. Também é notável que quanto menor o tamanho da matriz de *Pooling*, melhor o desempenho do modelo. Assim, os melhores resultados foram obtidos utilizando o tamanho de matriz  $2 \times 2$  e a função de ativação

ReLu.

Utilizando os melhores resultados obtidos pelos algoritmos PSO, GSA, HS e *Backpropagation*, foi feita também a comparação com o desempenho médio obtido pelos especialistas em radiologia na classificação das imagens. Na métrica de acurácia, o desempenho médio dos especialistas foi de 90,47%, em precisão foi de 83,40%, em sensibilidade 78,01% e em especificidade 94,71%, conforme apresentados na Figura 30.

Estes resultados mostram que os especialistas tendem a ter mais facilidade em reconhecer casos sadios e que cometem menos erros falso positivos do que falso negativos. Contudo, os modelos desenvolvidos com inteligência de enxame e com *backpropagation* foram superiores em todas as métricas de desempenho calculadas.

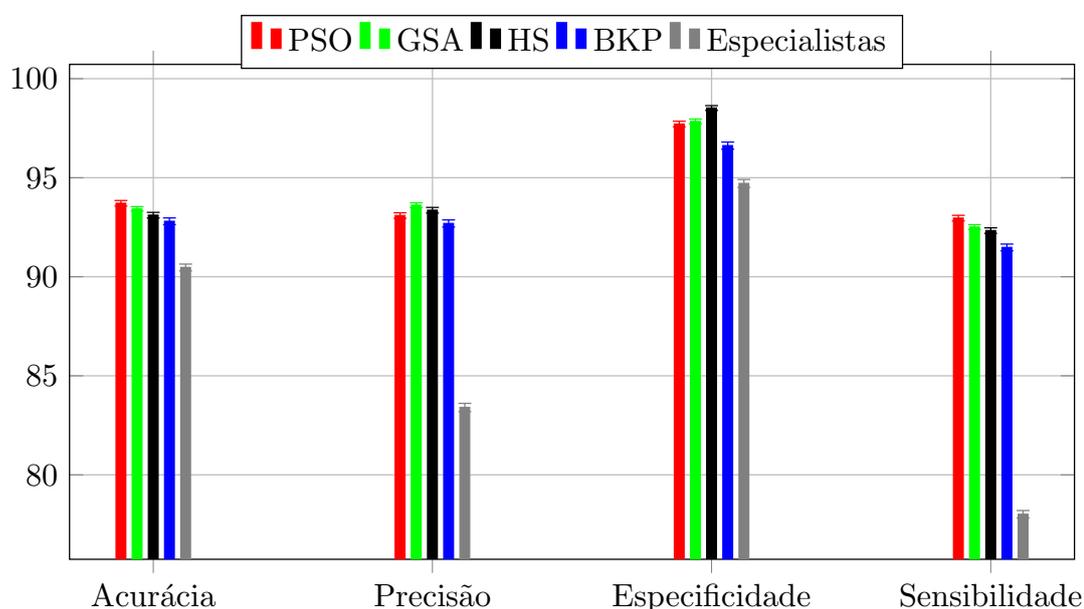


Figura 30: Comparação de desempenho entre algoritmos e especialistas

## 6.5 Validação estatística

Nesta seção serão apresentados e discutidos os métodos estatísticos utilizados para validar as conclusões inferidas a partir dos resultados obtidos.

A tomada de decisão sobre o melhor algoritmo de inteligência de enxame utilizado para treinar o modelo deve levar em consideração as 4 métricas utilizadas, para isso foi feita uma validação estatística por meio de Testes-T de *Student* para verificar quais conclusões podem ser inferidas com 95% de confiança (valores abaixo de 0,05).

A Tabela 13 mostra o teste estatístico realizado em cima da métrica de acurácia, buscando validar a conclusão da superioridade de desempenho do PSO. Ambos os resultados foram inferiores a 0,05, confirmando a premissa e elegendo o PSO como o melhor algoritmo sob a métrica de acurácia.

Tabela 13: Teste-T de *Student* para a acurácia.

Algoritmo 1	Algoritmo 2	Resultado
PSO	GSA	0,0056
PSO	HS	0,0057

A Tabela 14 mostra o teste estatístico realizado em cima da métrica de precisão, buscando validar a conclusão da superioridade de desempenho do GSA. Ambos os resultados foram inferiores a 0,05, confirmando a premissa e elegendo o GSA como o melhor algoritmo sob a métrica de precisão.

Tabela 14: Teste-T de *Student* para a precisão.

Algoritmo 1	Algoritmo 2	Resultado
GSA	PSO	0,0001
GSA	HS	0,0011

A Tabela 15 mostra o teste estatístico realizado em cima da métrica de sensibilidade, buscando validar a conclusão da superioridade de desempenho do PSO. Nesta validação, o PSO se mostrou superior ao HS, porém não foi possível descartar a hipótese nula quando comparando o PSO ao GSA com 95% de confiança. No entanto, para este segundo caso obteve-se um valor de teste suficientemente próximo a 0,05, o que permite ainda sim um alto grau de confiança. Sendo assim, pode-se confirmar a premissa e eleger o PSO como o melhor algoritmo sob a métrica de sensibilidade.

Tabela 15: Teste-T de *Student* para a sensibilidade.

Algoritmo 1	Algoritmo 2	Resultado
PSO	GSA	0,0668
PSO	HS	0,0012

A Tabela 16 mostra o teste estatístico realizado em cima da métrica de especificidade, buscando validar a conclusão da superioridade de desempenho do HS. Nesta validação, o HS se mostrou superior ao PSO, porém não foi possível descartar a hipótese nula quando comparando o HS ao GSA com 95% de confiança. Sendo assim, não foi possível comprovar que HS e GSA obtiveram desempenhos diferentes para 95% de confiança sob a métrica de especificidade.

Tabela 16: Teste-T de *Student* para a especificidade.

Algoritmo 1	Algoritmo 2	Resultado
HS	PSO	0,0040
HS	GSA	0,4407

A partir dessas validações é possível concluir que, nesta aplicação, o algoritmo GSA obteve o melhor desempenho em precisão, o que significa que este obtém o maior número de acertos entre os nódulos classificados como cancerígenos. Sob a especificidade, ou seja, a habilidade em assegurar que nódulos não cancerígenos sejam classificados como não cancerígenos, os algoritmos HS e GSA se destacaram, obtendo resultados estatisticamente similares. Por fim, o algoritmo PSO obteve os melhores desempenhos em acurácia e sensibilidade, o que significa que este não só tem a taxa de acertos mais alta, mas também é melhor em assegurar que os nódulos cancerígenos sejam realmente detectados como cancerígenos.

Em uma análise geral, quando se trata da detecção de uma doença grave como o câncer, a classificação negativa de um paciente com câncer é muito mais dolorosa do que a classificação positiva de um paciente sem câncer. Assim, o número de falsos negativos (pacientes com câncer classificados como sem câncer) é o fator mais importante nessa decisão, seguido pela acurácia, que mensura os acertos gerais do modelo. Considerando esses fatores, pode-se concluir que o PSO é o melhor algoritmo para o treinamento do modelo classificador de nódulos pulmonares desta aplicação.

## 6.6 Considerações finais

Neste capítulo foram apresentados e discutidos todos os resultados e inferências obtidos a partir dos experimentos realizados nesta aplicação. Os resultados indicam que para

---

aplicações como esta, o desempenho do modelo é melhorado quanto menor for a matriz de *Pooling*. Também foi possível verificar um melhor desempenho de funções de ativação ReLU em relação à Tanh.

No capítulo seguinte serão apresentadas as conclusões relevantes ao desenvolvimento e resultados desta dissertação, assim como os trabalhos futuros relacionados à proposta de detecção de nódulos cancerígenos com aprendizagem profunda e inteligência de enxame.

## Capítulo 7

# CONCLUSÕES E TRABALHOS FUTUROS

**E**STE capítulo completa esta dissertação, apresentando as principais conclusões obtidas durante o desenvolvimento do presente trabalho. São apresentadas também as possíveis direções para trabalhos futuros envolvendo aprendizagem profunda para a detecção de câncer de pulmão em exames de tomografia computadorizada.

### 7.1 Conclusões

Esta dissertação abordou o uso de aprendizagem profunda e inteligência de enxame para desenvolver um modelo capaz de reconhecer e classificar nódulos cancerígenos em exames de tomografia computadorizada de pulmões. Neste trabalho, investigou-se a real eficácia da utilização de algoritmos de inteligência de enxame no treinamento de modelos de redes neurais convolucionais. Ao todo foram elegidos 14 algoritmos de inteligência de enxame para compor os experimentos da aplicação. Como apresentado nos artigos derivados deste trabalho, (PINHEIRO; NEDJAH; MOURELLE, 2019) e (PINHEIRO; MOURELLE; NEDJAH, 2019), os resultados principais da aplicação estão na constatação da superioridade do treinamento da rede com algoritmos de enxame e na eleição do melhor algoritmo para o modelo, o PSO. Além do treinamento com inteligência de enxame, para efeito de comparação, o modelo foi treinado e testado também com a metodologia padrão, utilizando o algoritmo *Backpropagation*. Visando enriquecer a pesquisa e eventualmente encontrar resultados ainda mais expressivos, investigou-se também o desempenho dos modelos quando operando sob diferentes configurações de hiper-parâmetros da rede neural convolucional. O primeiro hiper-parâmetro testado foi a função utilizada nas camadas de ativação. Embora a função ReLU seja largamente utilizada na grande maioria das

aplicações de redes neurais convolucionais, para este experimento foi decidido avaliar o desempenho do modelo ao operar também com a função tangente hiperbólica (Tanh). O segundo hiper-parâmetro testado foi o tamanho da matriz de *Max-pooling* usada na rede neural convolucional. A rede utilizada tinha como padrão uma matriz de tamanho  $2 \times 2$ , isto é, que elege o *pixel* mais relevante dentro de uma janela quadrada de 4 *pixels*. Decidiu-se testar o efeito do aumento do tamanho desta matriz para as versões  $3 \times 3$ ,  $4 \times 4$  e  $5 \times 5$ . A partir destas definições, 120 modelos diferentes foram testados, utilizando cada uma das combinações entre os 4 tamanhos de matriz de *max-pooling*, as 2 funções de ativação e os 15 algoritmos de treinamento (14 com inteligência de enxame e 1 com *backpropagation*). Destes experimentos foram extraídas as 4 principais métricas de desempenho utilizadas em aplicações de aprendizado de máquinas, a acurácia, a precisão, a sensibilidade e a especificidade.

Este estudo foi motivado pela crescente demanda de desenvolvimento de sistemas capazes de detectar o câncer de pulmão com máximo desempenho, pois sabe-se que trata-se de uma das mais mortíferas modalidades de uma doença cuja detecção e tratamento são processos de alta complexidade. Seu objetivo foi investigar diferentes abordagens para o desenvolvimento de um modelo de sistema de diagnóstico auxiliado por computador (Sistema CAD), onde visou-se obter desempenhos expressivos ao utilizar algoritmos de inteligência de enxame e aprendizagem profunda. Como apresentado nos artigos derivados deste trabalho, (PINHEIRO; NEDJAH; MOURELLE, 2019) e (PINHEIRO; MOURELLE; NEDJAH, 2019), os resultados principais da aplicação estão na constatação da superioridade do treinamento da rede com algoritmos de enxame e na eleição do melhor algoritmo para o modelo.

Inicialmente buscou-se um banco de dados de imagens médicas de tomografia computadorizada de pulmão que fosse robusto o suficiente para conter um grande número de imagens e informações complementares de expressiva relevância, como descrito no início do Capítulo 5. Esta pesquisa resultou na eleição do banco LIDC (Cancer Imaging Archive, 2014) como base principal de dados e no banco LUNA16 (COLIN et al., 2016) como base auxiliar. Em seguida criou-se um ambiente virtual usando *Python*, principal linguagem utilizada para análise de dados e aprendizado de máquinas na atualidade. O ambiente foi montado para inicialmente agir sobre as imagens da base de dados, aplicando rotinas de pré-processamento como a criação de lotes de imagens para otimização de esforço

computacional, a geração de imagens focadas nos pontos de interesse apontados pelos radiologistas especialistas e a aplicação de técnicas de *data augmentation* para aumentar o conjunto de treinamento e balancear mais a presença das classes (nódulos cancerígenos ou sádios). Com os dados prontos o modelo foi submetido a uma rotina de busca de hiper-parâmetros, utilizando o treinamento com *backpropagation* e realizando 16 experimentos, com 4 taxas de aprendizados diferentes e 4 opções de quantidade de épocas de treinamento. Após essa rotina foi elegido o melhor par de hiper-parâmetros para serem utilizados em toda a aplicação. A este ponto, foram iniciados os experimentos de treinamento com os algoritmos de inteligência de enxame, executando 100 repetições para cada um das 14 técnicas sob cada uma das 8 configuração de função de ativação e matriz de *max-pooling*.

A partir destes experimentos, pode-se concluir a real eficácia da utilização de algoritmos de inteligência de enxame para o treinamento de modelos classificadores de câncer de pulmão utilizando o aprendizado por transferência. Dos 14 algoritmos testados, em média 7 se mostraram superiores ao modelo treinado com o *backpropagation* e cada métrica de performance registrada.

Dadas as avançadas técnicas disponíveis e a quantidade massiva de dados para treinamento, sistemas CAD atuais, em geral conseguem atingir acurácias altíssimas, superiores a 90%. Sendo assim, a disputa por melhores técnicas se dá por pequenas diferenças percentuais, onde a validação é garantida por testes estatísticos aplicados em experimentos com a maior quantidade possível de repetições. Neste cenário, foi possível extrair conclusões e validar os resultados pois testes-t de *Student* foram aplicados aos experimentos, comprovando com 95% de confiança que: Em relação ao modelo treinado com o *backpropagation*, o modelo que utilizou PSO foi cerca de 1% superior em acurácia e 1,5% em sensibilidade. Da mesma forma o GSA se mostrou superior em cerca de 1,2% em precisão e o HS em quase 2% em especificidade. Além dos resultados medidos por estas métricas, conforme apresentado no Capítulo 7, os modelos treinados com inteligência de enxame foram satisfatoriamente superiores em questão de tempo de execução, fornecendo mais um argumento para a constatação da superioridade do treinamento de redes convolucionais com inteligência de enxame.

A partir de uma análise focada em eleger o melhor par de hiper-parâmetros “função de ativação e tamanho da matriz de *max-pooling*”, foi possível concluir que, para esta aplicação, o aumento do tamanho da matriz de *max-pooling* reduz o desempenho dos

modelos. Fato que pode estar atrelado ao aumento da perda de informação nos mapas de características, pois a medida que a matriz cresce, mais *pixels* são removidos destes mapas. Concluiu-se também que os modelos treinados com a função Tanh foram levemente inferiores aos que utilizaram a ReLU. Este comportamento pode estar relacionado à linearidade da função Tanh, que é apontada por grande parte da literatura como um problema para o treinamento de redes convolucionais. Este fato justifica a grande aplicabilidade da função ReLU em modelos como este, pois esta utiliza operações não lineares durante a ativação. Assim, os melhores desempenhos foram obtidos utilizando o tamanho de matriz  $2 \times 2$  e a função de ativação ReLU.

Ao fim dos experimentos e dos testes estatísticos, foi conduzida uma análise sobre a natureza do problema para eleger o melhor modelo classificador de nódulos cancerígenos pulmonares. Ao listar os possíveis casos de diagnóstico, temos os casos *TP*, *TN*, *FP* e *FN*. Os casos TP e TN tratam de acertos do modelo, já os casos FP e FN, são erros do modelo. Quando se fala em diagnóstico de câncer, um caso FP é um erro relativamente leve, pois, embora o paciente possa ser submetido a tratamentos desnecessários, cedo ou tarde algum outro exame poderá reverter o mal entendido e constatar que este paciente não possui a doença. Por outro lado, um caso FN em geral é gravíssimo, pois o câncer tem suas chances de cura reduzidas drasticamente conforme o avanço da doença, o diagnóstico precoce é fundamental para maximizar a possibilidade de cura. A partir destes fatos, pode-se constatar que, para esta aplicação, o modelo treinado com o PSO foi o melhor, pois obteve não só a maior taxa de acertos gerais, mas também foi mais eficiente em encontrar os pacientes que de fato estavam doentes.

## 7.2 Trabalhos Futuros

Nesta seção são citadas algumas possíveis modificações à metodologia proposta, com o intuito de melhorar o desempenho dos modelos desenvolvidos neste trabalho. Também são levantadas propostas para futuros trabalhos na área de diagnóstico de câncer utilizando aprendizagem profunda e inteligência de enxame.

Dentre as possibilidades de modificações da metodologia, um dos principais pontos a se testar é o uso de diferentes mapas de extração de características na rede neural convolucional, experimentos que podem ser de alta complexidade, porém podem trazer melhoras expressivas ao desempenho do modelo.

Visando ainda ampliar a verificação da superioridade do treinamento de redes convolucionais com inteligência de enxame, outros algoritmos poderiam ser testados, como por exemplo o algoritmo de otimização inspirado nos pássaros de *Hitchcock*, desenvolvido por (MORAIS; MOURELLE; NEDJAH, 2018).

Para ampliar a gama de técnicas utilizadas e fornecer mais uma comparação aos algoritmos de inteligência de enxame, outros trabalhos poderiam ser conduzidos utilizando algoritmos genéticos para realizar o treinamento das redes neurais convolucionais. Outro ponto que poderia ser desenvolvido como uma comparação a este trabalho é o uso de redes convolucionais estado da arte como a *ResNet* (HE et al., 2016) e a *GoogLeNet* (SZEGEDY et al., 2015) para a aplicação do aprendizado por transferência.

Além das modificações, para trabalhos futuros, este tipo de aplicação poderia ser expandido para a detecção e classificação de nódulos de outros tipos de câncer, como o de mama, pele ou fígado.

# REFERÊNCIAS

- ALIHODZIC, A.; TUBA, M. Framework for bat algorithm optimization metaheuristic. *Recent Researches in Medicine, Biology and Bioscience*, p. 157–162, 08 2013.
- BURT, P.; ADELSON, E. The laplacian pyramid as a compact image code. *IEEE Transactions on communications, IEEE*, v. 31, n. 4, p. 532–540, 1983.
- Cancer Imaging Archive. *The Lung Image Database Consortium image collection (LIDC-IDRI)*. 2014. <https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI>. Accessed: 2018-02-02.
- CHON, A.; BALACHANDRA, N.; LU, P. *Deep convolutional neural networks for lung cancer detection*. Stanford University, USA, 08 2017.
- CHU, S.-C.; TSAI, P.-W. Computational intelligence based on the behavior of cats. *International Journal of Innovative Computing, Information and Control*, v. 3, n. 1, p. 168–176, 02 2007.
- COLIN, J. et al. *Lung Nodule Analysis 2016 (LUNA16)*. 2016. <https://luna16.grand-challenge.org/>. Accessed: 2018-02-02.
- CUEVAS, E.; CIENFUEGOS, M. A new algorithm inspired in the behavior of the social-spider for constrained optimization. *Expert Systems with Applications: An International Journal*, v. 41, p. 412–425, 02 2014.
- DHAWARE, B. U.; PISE, A. C. Lung cancer detection using bayasein classifier and fcm segmentation. In: IEEE. *Automatic Control and Dynamic Optimization Techniques (ICACDOT), International Conference on*. Pune, India, 2016. p. 170–174.
- FARAG, A. A. et al. Feature fusion for lung nodule classification. *International journal of computer assisted radiology and surgery*, Springer, v. 12, n. 10, p. 1809–1818, 2017.

- GEEM, Z. W.; KIM, J. H.; LOGANATHAN, G. A new heuristic optimization algorithm: Harmony search. *Simulation Journal*, v. 76, n. 2, p. 60–68, 2001.
- GOLUB, G. H.; LOAN, C. F. V. An analysis of the total least squares problem. *SIAM journal on numerical analysis*, SIAM, v. 17, n. 6, p. 883–893, 1980.
- HE, K. et al. Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. Las Vegas, USA: 1, 2016. v. 1, n. 1, p. 770–778.
- HUBEL, D. H.; WIESEL, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, Wiley Online Library, v. 160, n. 1, p. 106–154, 1962.
- JAYANTHI, S.; PREETHA, K. Breast cancer detection and classification using artificial neural network with particle swarm optimization. *International Journal of Advanced Research in Basic Engineering Sciences and Technology*, International Journal of Advanced Research in Basic Engineering Sciences and Technology, v. 2, 2016.
- KARABOGA, D.; BASTURK, B. A powerful and efficient algorithm for numerical function optimization: Artificial bee colony ABC algorithm. *J. of Global Optimization*, Kluwer Academic Publishers, Hingham, MA, USA, v. 39, n. 3, p. 459–471, nov. 2007. ISSN 0925-5001.
- KELNAR, D. *The fourth industrial revolution: a primer on Artificial Intelligence (AI)*. 2016. <https://medium.com/mmc-writes/the-fourth-industrial-revolution-a-primer-on-artificial-intelligence-ai/ff5e7fffcae1>. Accessed: 2018-02-02.
- KENNEDY, J. Particle swarm optimization. *Encyclopedia of machine learning*, Springer, p. 760–766, 2010.
- KHAJEHZADEH, M.; ESLAMI, M. Gravitational search algorithm for optimization of retaining structures. *Indian Journal of Science and Technology*, v. 5, n. 1, 2012.
- KIM, J.; AHN, K. *Imaging findings of central nervous system vasculitis associated with Goodpasture's Syndrome: a case report*. 2007. [https://openi.nlm.nih.gov/detailedresult.php?img=PMC2627458\\_kjr-8-545-g001&req=4](https://openi.nlm.nih.gov/detailedresult.php?img=PMC2627458_kjr-8-545-g001&req=4). Accessed: 2018-02-02.

- KOHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1995. (IJCAI'95), p. 1137–1143. ISBN 1-55860-363-8. Disponível em: <<http://dl.acm.org/citation.cfm?id=1643031.1643047>>.
- KOO, J. I.; PARK, S. B. Speckle reduction with edge preservation in medical ultrasonic images using a homogeneous region growing mean filter (hrgmf). *Ultrasonic imaging*, Elsevier, v. 13, n. 3, p. 211–237, 1991.
- KUAN, K. et al. Deep learning for lung cancer detection: Tackling the kaggle data science bowl 2017 challenge. *ArXiv preprint:1705.09435*, 2017.
- MALGONDE, S. *Transfer learning using Tensorflow*. 2018. <https://medium.com/@subodh.malgonde/transfer-learning-using-tensorflow-52a4f6bcde3e>. Accessed: 2018-02-02.
- MENDES, J. V.; CECILIO, A. M. Tendências regionais da mortalidade por câncer no estado de são paulo 2000 a 2010. *BEPA. Boletim Epidemiológico Paulista (Online)*, scielosp, v. 9, p. 24 – 45, 08 2012.
- MENDES, R. *Population Topologies and Their Influence in Particle Swarm Performance*. Tese (Doutorado) — University of Minho, 2004.
- MENG LIU, G.; ZHANG. A new bio-inspired algorithm: Chicken swarm optimization. *Advances in Swarm Intelligence. ICSI 2014. Lecture Notes in Computer Science*, v. 8794, p. 86–94, 2014.
- MESSAY, T.; HARDIE, R. C.; TUINSTRA, T. R. Segmentation of pulmonary nodules in computed tomography using a regression neural network approach and its application to the lung image database consortium and image database resource initiative dataset. *Medical image analysis*, v. 22 1, p. 48–62, 2015.
- MIAH, M. B. A.; YOUSUF, M. A. Detection of lung cancer from ct image using image processing and neural network. *2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)*, Dhaka, Bangladesh, n. 2, p. 1–6, 2015.

- MIRJALILI, S.; LEWIS, A. The whale optimization algorithm. *Advances in Engineering Software*, v. 95, p. 51–67, 05 2016.
- MIRJALILI, S.; MIRJALILI, S. M.; LEWIS, A. Grey wolf optimizer. *Advances in Engineering Software*, v. 69, p. 46 – 61, 2014. ISSN 0965-9978.
- MORAIS, R. G.; MOURELLE, L. M.; NEDJAH, N. Hitchcock birds inspired algorithm. In: SPRINGER. *International Conference on Computational Collective Intelligence*. Bristol, UK, 2018. p. 169–180.
- NAITO, N. et al. Kikuchi-fujimoto disease associated with community acquired pneumonia showing intrathoratic lymphadenopathy without cervical lesions. *SpringerPlus*, v. 4, n. 1, p. 693, Nov 2015. ISSN 2193-1801.
- OLAF, R.; PHILIPP, F.; THOMAS, B. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. 2015. <https://lmb.informatik.uni-freiburg.de/people/ronneber/u-net/>. Accessed: 2018-02-02.
- PARVEEN, S. S.; KAVITHA, C. Detection of lung cancer nodules using automatic region growing method. In: *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*. Tiruchengode, India: 4, 2013. v. 4, n. 4, p. 1–6.
- PASSINO, K. M. Bacterial foraging optimization. *International Journal of Swarm Intelligence*, IGI Global, Hershey, PA, USA, v. 1, n. 1, p. 1–16, jan. 2010. ISSN 1947-9263.
- PINHEIRO, C. A. de P.; NEDJAH, N.; MOURELLE, L. de M. Detection and classification of pulmonary nodules using deep learning and swarm intelligence. *Multimedia Tools and Applications*, Mar 2019. ISSN 1573-7721.
- PINHEIRO, C. de P.; MOURELLE, L. M.; NEDJAH, N. Lung nodule diagnosis via deep learning and swarm intelligence. In: SPRINGER. *International Conference on Computational Science*. Faro, Portugal, 2019. p. 169–180.
- RAJABIOUN, R. Cuckoo optimization algorithm. *Applied Soft Computing*, v. 11, p. 5508–5518, 12 2011.

- RITTHIPAKDEE THAMMANO, P.; JITKONGCHUEN. Firefly mating algorithm for continuous optimization problems. *Hindawi Journal*, v. 2017, jul. 2017. ISSN 1687-5265.
- SILVA, V. A. d. et al. Hamartoma pulmonar como diagnóstico diferencial de metástase de carcinoma de mama: relato de caso. *Jornal Brasileiro de Pneumologia*, scielo, v. 33, p. 738 – 742, 12 2007. ISSN 1806-3713.
- SIVAKUMAR, S.; CHANDRASEKAR, C. Lung nodule detection using fuzzy clustering and support vector machines. *International Journal of Engineering and Technology*, v. 5, n. 1, p. 179–185, 2013.
- SRIVASTAVA, N. et al. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, v. 15, p. 1929–1958, 2014.
- SUYKENS, J. A.; VANDEWALLE, J. Least squares support vector machine classifiers. *Neural processing letters*, Springer, v. 9, n. 3, p. 293–300, 1999.
- SZEGEDY, C. et al. Going deeper with convolutions. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. Boston, USA: 28, 2015. p. 1–9.
- TAN, Y.; ZHU, Y. Fireworks algorithm for optimization. In: *Proceedings of the First International Conference on Advances in Swarm Intelligence - Volume Part I*. Beijing, China: Springer-Verlag, 2010. p. 355–364.
- TOGA, A. W.; MAZZIOTTA, J. C. Ct angiography and ct perfusion imaging. In: *Brain Mapping: The Methods (Second Edition)*. Second edition. San Diego, USA: Academic Press, 2002. p. 427 – 484. ISBN 978-0-12-693019-1.
- VEREMYOV, R. *Implementing the human brain: Exploring the potential of convolutional neural networks*. 2018. <https://distillery.com/blog/implementing-human-brain-exploring-potential-convolutional-neural-networks/>. Accessed: 2018-02-02.
- YUVARANI, P. Image denoising and enhancement for lung cancer detection using soft computing technique. In: *IET. Sustainable Energy and Intelligent Systems (SEISCON 2012), IET Chennai 3rd International on*. Chennai, India, 2012. p. 1–4.

# APÊNDICE A – Resultados dos experimentos

Este apêndice apresenta os valores obtidos nos experimentos de treinamento e teste dos modelos classificadores de nódulos cancerígenos.

## A.1 Obtenção de hiper-parâmetros da rede

Tabela 17: Impacto dos hiper-parâmetros da camada totalmente conectada da rede nas métricas de desempenho

Caso	Épocas	Tx. apren.	Acurácia	Precisão	Especif.	Sensib.
$C_1$	30	0,0001	82,40	82,30	82,81	81,50
$C_2$	30	0,001	84,65	84,35	85,24	83,70
$C_3$	30	0,01	87,00	86,62	87,07	86,93
$C_4$	30	0,1	87,87	87,69	88,72	86,93
$C_5$	100	0,0001	88,94	89,10	90,48	89,60
$C_6$	100	0,001	90,97	90,93	91,60	90,52
$C_7$	100	0,01	90,60	90,24	92,51	90,96
$C_8$	100	0,1	91,14	91,12	93,33	91,72
$C_9$	300	0,0001	92,39	92,04	94,97	91,26
$C_{10}$	300	0,001	92,50	92,37	96,70	91,54
$C_{11}$	300	0,01	92,31	91,10	94,88	91,28
$C_{12}$	300	0,1	90,52	90,21	93,95	90,20
$C_{13}$	1000	0,0001	91,96	92,00	92,60	91,37
$C_{14}$	1000	0,001	92,00	91,67	92,11	91,83
$C_{15}$	1000	0,01	90,30	90,24	91,74	89,72
$C_{16}$	1000	0,1	89,95	89,66	90,98	89,23

## A.2 Métricas de desempenho

Tabela 18: Acurácia média para as funções de ativação e matrizes de *max-pooling*

Algoritmo	ReLU				Tanh			
	$S_1$	$S_2$	$S_3$	$S_4$	$S_5$	$S_6$	$S_7$	$S_8$
	<b>2×2</b>	<b>3×3</b>	<b>4×4</b>	<b>5×5</b>	<b>2×2</b>	<b>3×3</b>	<b>4×4</b>	<b>5×5</b>
<b>PSO</b>	<b>93,78</b>	<b>93,29</b>	<b>93,02</b>	<b>93,01</b>	<b>93,33</b>	<b>92,98</b>	<b>92,79</b>	<b>92,35</b>
GSA	93,52	93,20	92,70	92,63	93,16	92,83	92,70	92,42
HS	93,17	93,07	92,89	92,53	93,16	92,85	92,66	92,28
ABC	92,95	92,47	92,38	92,40	93,10	92,22	92,37	91,69
GWO	92,80	92,22	92,30	91,73	92,73	91,47	92,20	91,62
SSA	92,62	91,90	91,79	91,60	92,05	91,95	91,43	90,76
BFO	92,57	92,54	92,27	91,56	92,59	92,36	91,48	91,04
SOE	92,88	92,70	92,68	92,55	92,91	92,67	92,37	92,22
BOA	92,30	91,90	91,36	91,40	91,89	91,09	90,90	90,57
FA	91,62	91,20	90,77	90,00	91,56	90,88	90,31	89,18
WSA	91,53	91,18	91,05	91,08	91,66	90,58	90,59	91,17
CSA	91,51	91,19	90,88	90,36	91,57	90,60	90,51	89,98
FOA	91,14	90,70	90,58	90,57	91,04	90,35	90,54	90,54
ChSO	90,99	91,04	90,74	90,14	91,10	91,19	90,08	89,40
CSO	90,10	89,97	89,86	90,05	89,60	89,61	89,37	89,88
<b>Média</b>	<b>92,71</b>	<b>92,40</b>	<b>92,16</b>	<b>91,94</b>	<b>92,62</b>	<b>92,16</b>	<b>91,95</b>	<b>91,43</b>

Tabela 19: Precisão média para as funções de ativação e matrizes de *max-pooling*

Algoritmo	ReLU				Tanh			
	$S_9$	$S_{10}$	$S_{11}$	$S_{12}$	$S_{13}$	$S_{14}$	$S_{15}$	$S_{16}$
	<b>2×2</b>	<b>3×3</b>	<b>4×4</b>	<b>5×5</b>	<b>2×2</b>	<b>3×3</b>	<b>4×4</b>	<b>5×5</b>
PSO	93,16	92,35	92,13	92,11	93,05	91,75	91,57	91,30
<b>GSA</b>	<b>93,59</b>	<b>93,69</b>	<b>93,50</b>	<b>93,59</b>	<b>93,10</b>	<b>92,85</b>	<b>92,69</b>	<b>92,17</b>
HS	93,44	92,64	92,01	91,98	93,42	91,94	92,11	91,21
ABC	92,08	91,31	91,41	91,32	92,05	90,95	91,41	91,32
GWO	92,89	92,16	91,53	91,13	92,81	92,18	91,60	90,66
SSA	92,86	92,54	92,17	91,54	92,95	92,01	91,77	91,25
BFO	91,24	91,29	90,81	90,29	90,51	90,89	90,61	89,57
SOE	92,36	92,28	91,80	91,13	92,12	92,35	91,58	91,08
BOA	91,92	91,65	91,48	90,91	91,95	91,73	91,20	91,06
FA	92,33	91,54	91,46	91,06	91,73	91,14	90,91	90,43
WSA	92,40	91,94	91,28	90,48	92,55	91,18	91,32	90,45
CSA	92,39	92,11	91,68	91,63	91,59	91,28	91,22	91,64
FOA	92,38	91,84	91,60	90,89	92,51	91,42	91,20	90,88
ChSO	92,49	91,71	91,01	91,10	91,98	91,26	90,24	90,76
CSO	90,76	90,80	90,59	90,05	90,39	90,02	90,30	89,77
<b>Média</b>	<b>92,56</b>	<b>92,08</b>	<b>91,71</b>	<b>91,45</b>	<b>92,24</b>	<b>91,59</b>	<b>91,49</b>	<b>91,04</b>

Tabela 20: Sensibilidade média para as funções de ativação e matrizes de *max-pooling*

Algoritmo	ReLU				Tanh			
	$S_{17}$	$S_{18}$	$S_{19}$	$S_{20}$	$S_{21}$	$S_{22}$	$S_{23}$	$S_{24}$
	$2 \times 2$	$3 \times 3$	$4 \times 4$	$5 \times 5$	$2 \times 2$	$3 \times 3$	$4 \times 4$	$5 \times 5$
<b>PSO</b>	<b>93,04</b>	<b>92,50</b>	<b>91,87</b>	<b>91,81</b>	<b>92,39</b>	<b>92,11</b>	<b>91,81</b>	<b>91,80</b>
GSA	92,60	92,41	91,77	91,82	92,13	92,33	91,80	91,16
HS	92,39	91,74	91,87	91,08	91,68	91,51	91,67	90,59
ABC	91,95	91,18	90,36	90,27	91,79	91,14	90,12	89,80
GWA	91,83	91,96	91,91	91,27	90,98	91,12	91,43	90,64
SSA	92,03	91,48	91,03	90,90	91,52	91,65	90,18	90,62
BFO	90,49	90,43	90,31	90,32	90,32	89,71	89,73	90,39
SOE	91,57	91,17	90,56	89,72	91,19	90,99	90,69	89,79
BOA	91,31	91,30	90,79	90,90	90,52	90,57	90,67	90,32
FA	91,09	90,56	90,04	89,73	91,22	90,67	90,19	89,33
WSA	90,95	90,78	90,65	90,24	90,80	89,96	90,07	90,13
CSA	91,40	91,18	91,31	90,55	90,96	91,02	91,12	90,53
FOA	91,56	90,96	90,38	90,40	90,82	90,32	89,88	90,39
ChSO	91,62	91,11	90,37	90,02	91,30	90,48	90,27	89,95
CSO	90,39	89,99	89,73	88,95	89,68	89,69	89,62	88,86
<b>Média</b>	<b>91,76</b>	<b>91,29</b>	<b>90,82</b>	<b>90,56</b>	<b>91,37</b>	<b>91,03</b>	<b>90,66</b>	<b>90,33</b>

Tabela 21: Especificidade média para as funções de ativação e matrizes de *max-pooling*

Algoritmo	ReLU				Tanh			
	$S_{25}$	$S_{26}$	$S_{27}$	$S_{28}$	$S_{29}$	$S_{30}$	$S_{31}$	$S_{32}$
	$2 \times 2$	$3 \times 3$	$4 \times 4$	$5 \times 5$	$2 \times 2$	$3 \times 3$	$4 \times 4$	$5 \times 5$
PSO	97,76	97,19	97,12	97,27	97,89	97,13	96,93	96,78
GSA	97,89	97,86	97,99	97,68	97,23	97,38	98,05	97,63
<b>HA</b>	<b>98,56</b>	<b>98,08</b>	<b>97,31</b>	<b>97,46</b>	<b>98,39</b>	<b>97,36</b>	<b>97,41</b>	<b>96,74</b>
ABC	97,62	96,78	95,95	95,28	97,38	96,48	95,42	94,78
GWA	97,37	97,15	96,48	95,90	96,97	96,72	96,33	95,56
SSA	96,77	96,82	96,92	96,65	96,17	96,97	96,84	95,78
BFO	94,92	95,04	94,65	94,78	94,79	95,00	94,44	94,41
SOE	96,66	96,00	95,31	94,56	96,56	95,67	95,41	94,52
BOA	95,23	95,23	94,65	94,61	94,67	94,73	94,26	94,66
FA	95,38	95,02	94,50	93,96	95,18	94,52	94,42	94,02
WSA	96,30	95,69	95,16	94,36	95,93	95,23	94,67	94,36
CSA	95,70	95,81	95,06	95,12	95,65	95,49	94,97	94,61
FA	96,39	95,77	95,44	95,28	96,25	95,17	94,94	95,21
ChSO	97,25	96,68	96,17	95,66	96,48	96,00	95,35	95,26
CSO	93,97	93,96	93,74	93,13	93,73	93,49	93,18	92,39
<b>Média</b>	<b>96,89</b>	<b>96,47</b>	<b>95,96</b>	<b>95,79</b>	<b>96,68</b>	<b>96,07</b>	<b>95,89</b>	<b>95,58</b>