



UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS DE QUIXADÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO
MESTRADO ACADÊMICO EM COMPUTAÇÃO

SAMUEL SANCHES DE FREITAS

**SISTEMA PARA TREINAMENTO E CLASSIFICAÇÃO DE IMAGENS DE
RADIOGRAFIAS**

QUIXADÁ

2022

SAMUEL SANCHES DE FREITAS

SISTEMA PARA TREINAMENTO E CLASSIFICAÇÃO DE IMAGENS DE
RADIOGRAFIAS

Dissertação apresentada ao Curso de Mestrado Acadêmico em Computação do Programa de Pós-Graduação em Computação da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Computação. Área de Concentração: Ciência da Computação

Orientadora: Prof. Dra. Paulyne Matthews Jucá

QUIXADÁ

2022

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

F938s Freitas, Samuel Sanches de.
Sistema para treinamento e classificação de imagens de radiografias / Samuel Sanches de Freitas. – 2022.
69 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Campus de Quixadá,
Programa de Pós-Graduação em Computação, Quixadá, 2022.
Orientação: Profa. Dra. Paulyne Matthews Jucá.

1. Radiografia torácica. 2. Aprendizagem profunda. 3. Rede Neural Convolutacional. I.
Título.

CDD 005

SAMUEL SANCHES DE FREITAS

SISTEMA PARA TREINAMENTO E CLASSIFICAÇÃO DE IMAGENS DE
RADIOGRAFIAS

Dissertação apresentada ao Curso de Mestrado Acadêmico em Computação do Programa de Pós-Graduação em Computação da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Computação. Área de Concentração: Ciência da Computação

Aprovada em: 29 de Novembro de 2022

BANCA EXAMINADORA

Prof. Dra. Paulyne Matthews Jucá (Orientadora)
Universidade Federal do Ceará (UFC)

Prof. Dr. Arthur de Castro Callado
Universidade Federal do Ceará (UFC)

Prof. Dr. José Maria da Silva Monteiro Filho
Universidade Federal do Ceará (UFC)

Prof. Dr. Sidartha Azevedo Lobo de Carvalho
Universidade Federal de Pernambuco (UFPE)

AGRADECIMENTOS

Agradeço à comunidade de software aberto pelas incontáveis horas de esforço e dedicação contribuídas ao ensino e à propagação do conhecimento, sem as quais este trabalho simplesmente não seria possível. Àqueles que alimentam a comunidade com suas dúvidas, experiências e contribuições nos clareiam a mente quando os livros não são capazes.

Agradeço a professora Paulyne Matthews Jucá pela paciência na orientação e incentivo que tornaram possível a conclusão desta dissertação.

Agradeço a minha esposa Erica Eva Sousa Jucá de Freitas pela compreensão e paciência demonstrada durante o período do projeto.

Agradeço aos meus pais por todos os sacrifícios feitos ao longo dos anos. Ao meu irmão e tios. A todos estes, o meu “muito obrigado” pelo apoio incondicional, alegria, exemplo e estímulo.

Agradeço a Thiago Oliveira por todo o apoio durante a produção deste trabalho.

Em nome de Régis Melo, André Campos, Edson Patrício, Daniel Araújo, Paulo Filipe, Glauco Aquino e Alan Firmiano agradeço a todos os meus amigos, essenciais nessa etapa da minha vida.

Finalmente, à banca avaliadora, por prontamente aceitar fazer parte desse momento tão especial. Aos companheiros de trabalho, funcionários e desconhecidos que contribuíram sem perceber para que esse sonho se realizasse.

“Nada na vida pode ser obtido sem uma espécie de sacrifício.”

(Edward Elric)

RESUMO

A radiografia de tórax é o tipo de exame de imagem mais comum no mundo. Sendo muito utilizado para diagnóstico de doenças torácicas. Sendo uma das ferramentas mais valiosas para o diagnóstico de doenças se interpretada com cuidado. Nesse contexto, para ajudar os médicos a diagnosticar mais facilmente pacientes doenças no tórax, este trabalho propõe um modelo de aprendizado de máquina para realizar esse diagnóstico por meio de imagens de raio-x que rodam em uma arquitetura em nuvem. As métricas de precisão, perda e área sob a curva (AUC) são usadas para avaliar o desempenho do modelo. Este modelo proposto apresenta uma boa taxa de sucesso, atingindo 98% de acurácia para algumas doenças, Este trabalho também implementou um sistema provisionado em uma arquitetura na AWS, diferenciando-o de outros trabalhos relacionados a classificação de radiografias.

Palavras-chave: Doenças de tórax. Aprendizado de máquina. Rede Neural Convolutacional. Radiografias.

ABSTRACT

Chest X-ray is the most common type of imaging examination in the world. It is widely used for diagnosing thoracic diseases and is one of the most valuable tools when interpreted carefully. In this context, to assist doctors in diagnosing thoracic diseases more easily, this study proposes a machine learning model to perform such diagnoses using cloud-based X-ray images. Accuracy, loss, and area under the curve (AUC) metrics are used to evaluate the model's performance. The proposed model demonstrates a high success rate, achieving 98% accuracy for some diseases. Additionally, this study implemented a system provisioned in an AWS architecture, setting it apart from other works related to X-ray classification.

keywords Chest diseases. Machine learning. Convolutional Neural Network. Radiographs.

LISTA DE FIGURAS

Figura 1	– Radiografias de tórax usando o anteroposterior e lateral exemplificando a presença e distribuição de atelectasia (PERONI; BONER, 2000).	18
Figura 2	– “Atelectasia do lobo médio direito por compressão brônquica extrínseca determinada pelo aumento dos linfonodos hilares” (PERONI; BONER, 2000).	19
Figura 3	– Radiografia de tórax em Pósterio-anterior (PA) mostrando Relação Cardiotorácica (CTR) (a = distância da borda direita do coração à linha média, b = distância da borda esquerda do coração à linha média e c = diâmetro torácico máximo) (ALGHAMDI <i>et al.</i> , 2020).	20
Figura 4	– Radiografias de tórax representativas de pacientes com edema pulmonar cardiogênico e não cardiogênico. (WARE; MATTHAY, 2005).	21
Figura 5	– (A) Radiografia de tórax pósterio anterior com incidência de derrame pleural e (B) Radiografia mostrando o pulmão normal (KARKHANIS; JOSHI, 2012).	22
Figura 6	– Cancer inferido devido ao colapso lobar (HOLLINGS; SHAW, 2002).	24
Figura 7	– Radiografia com placas pleurais calcificadas bilaterais. (QURESHI; GLEESON, 2006).	25
Figura 8	– Penetração profunda revelando grande pneumotórax esquerdo em paciente ventilado mecanicamente (YARMUS; FELLER-KOPMAN, 2012).	26
Figura 9	– Radiografia de um pulmão com COVID-19.	29
Figura 10	– Exemplo da arquitetura de uma CNN com convoluções, pooling e camadas totalmente conectadas para função de ativação softmax.	34
Figura 11	– Execução da operação de MaxPooling.	34
Figura 12	– Exemplo de camada totalmente conectada.	35
Figura 13	– Gráfico ROC.	40
Figura 14	– Exemplo de raio-x usado no treinamento.	53
Figura 15	– Distribuição de imagens entre as classes.	54
Figura 16	– Distribuição de imagens entre as classes.	55
Figura 17	– Percentual de representatividade de cada classe.	56
Figura 18	– Estrutura do modelo.	56
Figura 19	– Gráfico ROC e AUC do modelo.	58
Figura 20	– Arquitetura do sistema web.	60
Figura 21	– Website para consumir o modelo.	61

Figura 22 – Aviso sobre o uso do website.	62
Figura 23 – Campo para o upload da imagem.	62
Figura 24 – Resultado da classificação.	62

LISTA DE TABELAS

Tabela 1 – Matriz de Confusão	38
---	----

LISTA DE QUADROS

Quadro 1 – Quadro com trabalhos relacionados.	50
---	----

LISTA DE ABREVIATURAS E SIGLAS

PA	Pósterio-anterior
CTR	Relação Cardiorácica
CT	Tomografia computadorizada
CoV	Coronavírus
SARS	Síndrome respiratória aguda grave
OMS	Organização Mundial de Saúde
RT-PCR	teste de reação em cadeia da polimerase com transcrição reversa
CNN	Rede Neural Convolucional
FC	Camada totalmente conectada
AUC	Área Sob a Curva
VP	Verdadeiro Positivos
FP	Falso Positivos
FN	Falso Negativo
VN	Verdadeiro Negativo
ROC	Receiver Operating Characteristics
AWS	Amazon Web Services
AWS Batch	Amazon Web Services Batch
CPU	Unidade Central de Processamento
CXR	Tomografia computadorizada

SUMÁRIO

1	INTRODUÇÃO	15
2	OBJETIVOS	16
2.1	Objetivo geral	16
2.2	Objetivos específicos	16
3	REVISÃO BIBLIOGRÁFICA	17
3.1	Doenças analisadas	17
3.1.1	<i>Atelectasia</i>	17
3.1.2	<i>Cardiomegalia ou Cardiomegaly</i>	19
3.1.3	<i>Edema</i>	20
3.1.4	<i>Efusão pleural ou Pleural effusion</i>	21
3.1.5	<i>Enfisema pulmonar ou Emphysema</i>	22
3.1.6	<i>Câncer de pulmão ou mass</i>	23
3.1.7	<i>Espessamento pleural ou Pleural thickening</i>	24
3.1.8	<i>Pneumotórax ou Pneumothorax</i>	26
3.1.9	COVID-19	27
3.1.9.1	<i>Cronologia e disseminação.</i>	27
3.1.9.2	<i>Diagnóstico de imagem na COVID-19</i>	28
3.2	Aprendizagem de máquina	30
3.2.1	<i>Tipos de Classificadores</i>	31
3.2.2	<i>Rede Neural Convolucional (CNN)</i>	32
3.2.3	<i>Tensorflow</i>	36
3.3	Avaliação de algoritmos de classificação	37
3.4	Computação em nuvem	41
3.4.1	<i>Provedores de serviços em nuvem</i>	42
3.4.2	<i>Amazon CloudFront</i>	42
3.4.3	<i>S3</i>	42
3.4.4	<i>Route 53</i>	43
3.4.5	<i>AWS Batch</i>	43
4	TRABALHOS RELACIONADOS	45
4.1	Metodologia da pesquisa	45

4.2	Principais trabalhos encontrados	45
4.3	Comparação com trabalhos relacionados	48
5	METODOLOGIA	51
5.1	Conjunto de Dados	51
5.2	Configuração do Experimento	51
5.3	Avaliação	52
6	O MODELO PROPOSTO	53
7	SISTEMA WEB	60
7.1	Arquitetura em nuvem	60
7.2	Website	61
7.3	Considerações Finais	63
	REFERÊNCIAS	64

1 INTRODUÇÃO

A radiografia de tórax é o tipo de exame de imagem mais comum no mundo, com mais de 2 bilhões de procedimentos realizados a cada ano (RAJPURKAR *et al.*, 2018). Este tipo de exame é muito utilizado para triagem, diagnóstico e manejo de doenças torácicas (RAJPURKAR *et al.*, 2018). Algumas destas doenças estão entre as principais causas de mortalidade em todo o mundo (RAJPURKAR *et al.*, 2018). A radiografia de tórax contém muitas informações sobre a condição médica de um paciente. Produz imagens de costelas, coluna, pulmões, coração, vasos sanguíneos e vias aéreas (MCADAMS *et al.*, 2006). É uma das ferramentas mais valiosas para o diagnóstico de doenças se interpretada com cuidado (RAJPURKAR *et al.*, 2018; MCADAMS *et al.*, 2006).

A radiografia de tórax pode detectar e ajudar a diagnosticar várias doenças, como Atelectasia (PERONI; BONER, 2000), Cardiomegalia (H, 2022), Edema (WARE; MATTHAY, 2005), Efusão pleural (KARKHANIS; JOSHI, 2012), Enfisema pulmonar (CAMPO *et al.*, 2018), Câncer de pulmão (BEEK *et al.*, 2015), Espessamento pleural (SAITO *et al.*, 2019), Pneumotórax (CHAN *et al.*, 2020) e outras.

Portanto, com base na eficácia do diagnóstico utilizando radiografias de tórax, este trabalho tem como objetivo estudar a viabilidade de criar um sistema que auxilie no diagnóstico por meio de um modelo de aprendizado de máquina. Ao contrário de outros trabalhos existentes que se concentram apenas na classificação de imagens como COVID-19, pneumonia e normal, o modelo proposto neste trabalho visa diagnosticar também outras doenças, como atelectasia, cardiomegalia, edema, efusão pleural, enfisema pulmonar, câncer de pulmão, espessamento pleural, pneumotórax, entre outras. Dessa forma, quando surgir uma nova doença, o modelo poderá ser treinado para reconhecê-la, auxiliando em um diagnóstico mais rápido e abrangente da população, especialmente daqueles pacientes que têm acesso a radiografias, mas não possuem acesso a profissionais especializados para interpretar esses exames. O sistema fornecerá uma sugestão inicial de possível quadro antes da avaliação final por um profissional especializado.

Como exemplo de aplicação, este trabalho também propõe uma arquitetura em nuvem que permite o treinamento e a execução escalável de um modelo, juntamente com a implementação de um website para a classificação das imagens.

2 OBJETIVOS

2.1 Objetivo geral

- Implementar um sistema para classificar imagens de radiografias de tórax provisionado em uma arquitetura em nuvem.

2.2 Objetivos específicos

- Identificação dos algoritmos de classificação de imagem disponíveis na literatura para selecionar o método a ser utilizado.
- Elaborar modelo para detectar doenças em imagens de raio-x de tórax.
- Avaliar a eficácia do modelo utilizando curva ROC e AUC.
- Propor e implementar uma arquitetura em nuvem para treinar e executar o modelo.

3 REVISÃO BIBLIOGRÁFICA

O desenvolvimento do modelo para classificação de imagens de radiografias está inserido nas áreas de aprendizado de máquina e classificação de imagens. Nesta seção, serão abordados os conceitos-chave relacionados à aprendizado de máquina, avaliação de classificadores, computação em nuvem e o processo de diagnóstico das doenças analisadas.

3.1 Doenças analisadas

Nesta seção são descritas as doenças utilizadas para treinar o modelo de classificação. A maioria destas doenças tem como principal método de diagnóstico exames de imagens principalmente o uso de radiografias.

A doença categorizada como infiltração (infiltration), não possui uma seção específica por ser uma categoria genérica utilizada em doenças pulmonares como pneumonia (PATTERSON; SPONAUGLE, 2005).

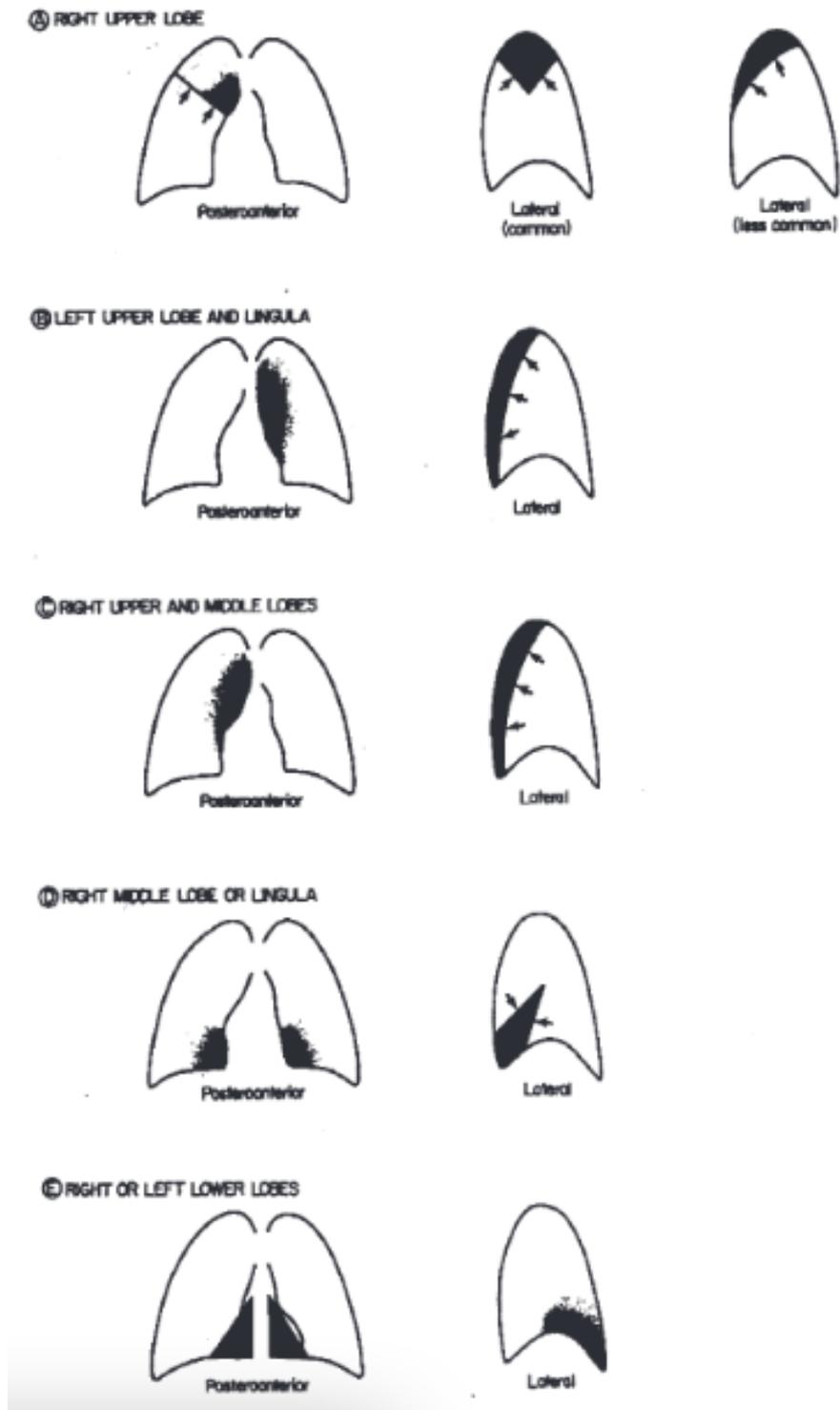
3.1.1 *Atelectasia*

O termo atelectasia descreve um estado de região colapsada e não aerada do parênquima pulmonar, que é normal (PERONI; BONER, 2000). Essa condição patológica geralmente está associada a vários distúrbios pulmonares e torácicos e representa uma manifestação da doença de base, não uma doença em si (PERONI; BONER, 2000). A atelectasia pode ocorrer de três formas: (i) obstrução das vias aéreas; (ii) compressão do parênquima por processos extratorácicos, intratorácicos, da parede torácica; e (iii) aumento da tensão superficial em alvéolos e bronquíolos (PERONI; BONER, 2000). A principal forma de diagnóstico é através de radiografias de tórax usando as projeções anteroposterior e lateral para documentar a presença de atelectasia (PERONI; BONER, 2000).

Para o diagnóstico de atelectasia, as radiografias de tórax usando tanto a região ântero-posterior quanto a lateral injeções são obrigatórias para documentar a presença, extensão, e distribuição das atelectasias (PERONI; BONER, 2000). Na imagem 1 são mostradas as possíveis distribuições da atelectasia. Em alguns casos, radiograficamente, pode haver problemas para diferenciar atelectasia de consolidação lobar simples ou uma massa tumoral, esses são os casos em que uma Tomografia computadorizada (CT) é útil (PERONI; BONER, 2000). Um exemplo de um caso no qual a atelectasia pode ser confundida com uma consolidação lombar é

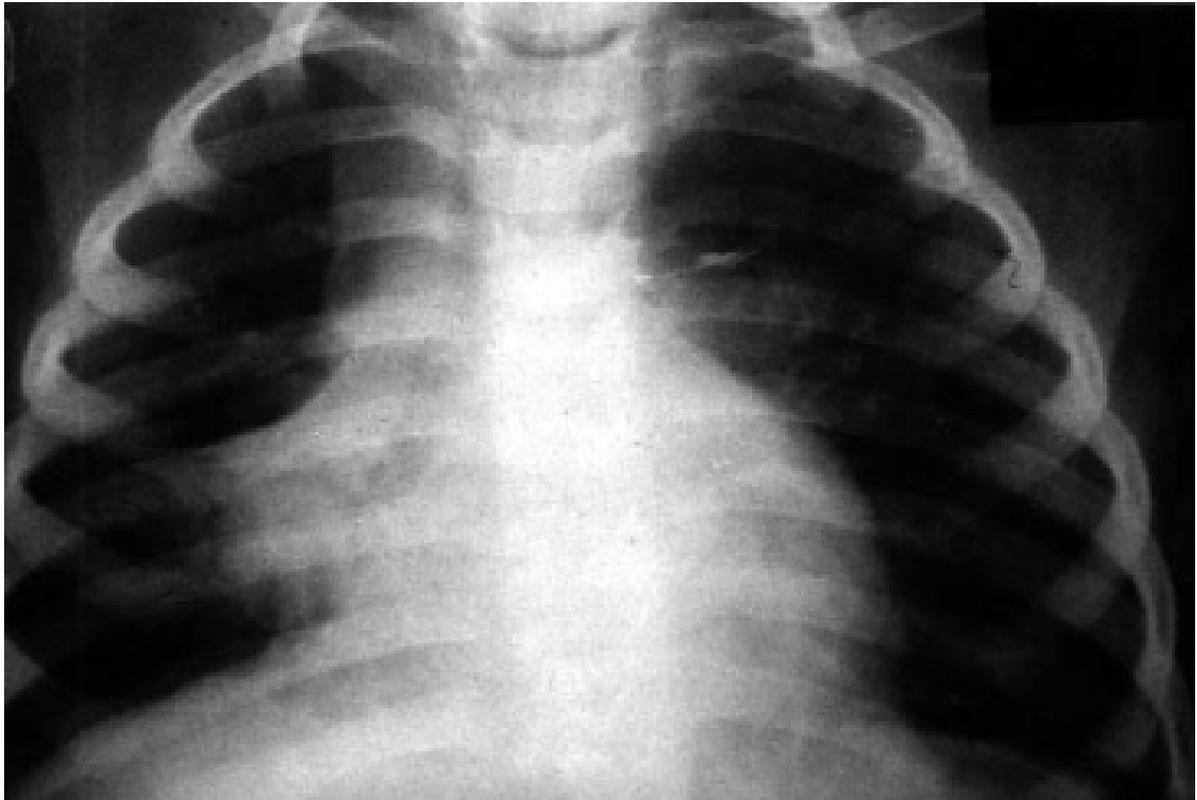
demonstrada na imagem 2.

Figura 1 – Radiografias de tórax usando o anteroposterior e lateral exemplificando a presença e distribuição de atelectasia (PERONI; BONER, 2000).



Fonte: (PERONI; BONER, 2000).

Figura 2 – “Atelectasia do lobo médio direito por compressão brônquica extrínseca determinada pelo aumento dos linfonodos hilares” (PERONI; BONER, 2000).



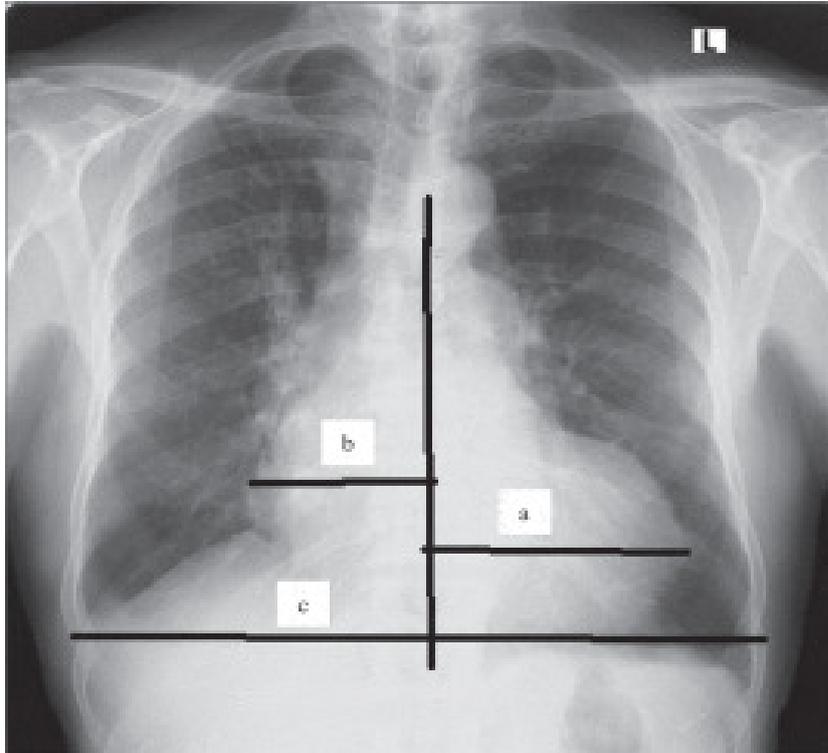
Fonte: (PERONI; BONER, 2000).

3.1.2 *Cardiomegalia ou Cardiomegaly*

Cardiomegalia é uma designação abrangente para uma variedade de condições que levam ao aumento do coração, que geralmente permanece sem diagnóstico até que os sintomas ocorram. Tornou-se cada vez mais prevalente e carrega uma alta mortalidade (H, 2022). A definição é quando o diâmetro transversal da silhueta cardíaca é maior ou igual a 50% do diâmetro transversal do tórax (índice cardiotorácico aumentado) na projeção pósterio-anterior de uma radiografia de tórax ou tomografia computadorizada (H, 2022).

O diagnóstico da cardiomegalia é feito principalmente usando técnicas de imagem que fornecem uma avaliação do tamanho do coração. Na Imagem 3, mostra o exemplo de uma radiografia com cardiomegalia. A reta A é a distância da borda direita do coração à linha média, a reta B é a distância da borda esquerda do coração à linha média e a reta C é diâmetro torácico máximo. A cardiomegalia é detectada quando a soma das retas A e B representam mais de 50% da reta C (ALGHAMDI *et al.*, 2020).

Figura 3 – Radiografia de tórax em PA mostrando CTR (a = distância da borda direita do coração à linha média, b = distância da borda esquerda do coração à linha média e c = diâmetro torácico máximo) (ALGHAMDI *et al.*, 2020).



Fonte: (ALGHAMDI *et al.*, 2020).

3.1.3 Edema

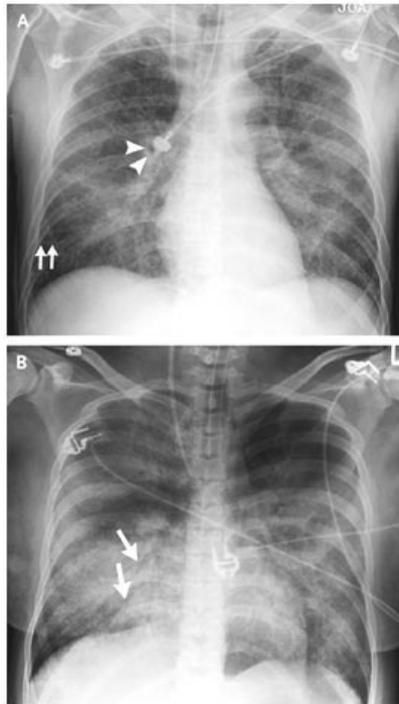
O edema pulmonar é quando há acúmulo de líquido dentro dos pulmões (WARE; MATTHAY, 2005). Os dois casos de edema pulmonar que ocorrem em humanos são: edema pulmonar cardiogênico (também denominado edema hidrostático ou hemodinâmico) e edema pulmonar não cardiogênico (também conhecido como edema pulmonar de permeabilidade aumentada, lesão pulmonar aguda ou síndrome do desconforto respiratório agudo) (WARE; MATTHAY, 2005). Embora tenham causas distintas, o edema pulmonar cardiogênico e não cardiogênico podem ser difíceis de distinguir devido às suas manifestações clínicas semelhantes (WARE; MATTHAY, 2005).

O conhecimento da causa do edema agudo de pulmão tem implicações importantes para o tratamento. Pacientes com edema pulmonar cardiogênico geralmente são tratados com diuréticos (WARE; MATTHAY, 2005). Já pacientes com edema pulmonar não cardiogênico necessitam de ventilação mecânica (WARE; MATTHAY, 2005).

As características do edema pulmonar cardiogênico e não cardiogênico podem apresentar pequenas diferenças em radiografias de tórax pósterio-anterior (WARE; MATTHAY,

2005). É importante ressaltar que a radiografia de tórax possui algumas limitações diagnósticas para o edema pulmonar, uma vez que ele pode não ser visível até que a quantidade de água nos pulmões atinja cerca de 30% do volume pulmonar (WARE; MATTHAY, 2005). Na Imagem 4, são exibidos os dois tipos de edema, no topo edema pulmonar cardiogênico e a baixo edema pulmonar não cardiogênico (WARE; MATTHAY, 2005).

Figura 4 – Radiografias de tórax representativas de pacientes com edema pulmonar cardiogênico e não cardiogênico. (WARE; MATTHAY, 2005).



Fonte: (WARE; MATTHAY, 2005).

3.1.4 *Efusão pleural ou Pleural effusion*

Um derrame pleural, ou seja, um acúmulo excessivo de líquido no espaço pleural, indica um desequilíbrio entre a formação e a remoção do líquido pleural. O acúmulo de líquido pleural não é uma doença específica, mas sim um reflexo da patologia subjacente. Os derrames pleurais acompanham uma ampla variedade de distúrbios pulmonares, pleurais e sistêmicos (KARKHANIS; JOSHI, 2012). Portanto, um paciente com derrame pleural pode se apresentar não apenas a um pneumologista, mas a um clínico geral, reumatologista, gastroenterologista, nefrologista ou cirurgião (KARKHANIS; JOSHI, 2012).

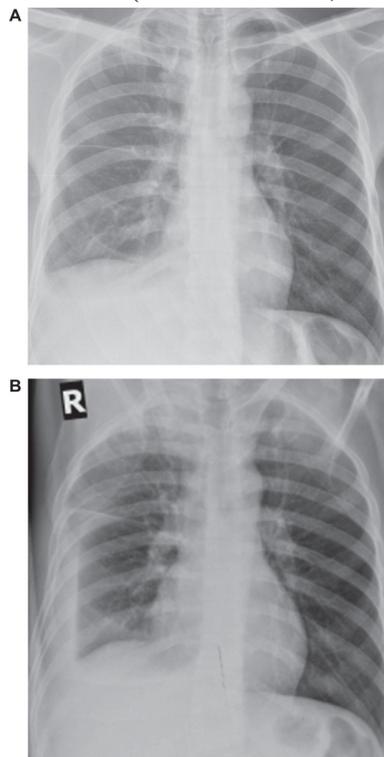
A radiografia de tórax pósterio anterior e lateral padrão continua sendo a técnica mais

importante para o diagnóstico inicial de derrame pleural (KARKHANIS; JOSHI, 2012).

As pistas radiográficas sugestivas de derrame pleural são: elevação aparente do diafragma; movimento do ápice do hemidiafragma do terço medial para o lateral; achatamento do aspecto medial do diafragma; não visualização dos vasos sanguíneos do lobo inferior abaixo do diafragma (KARKHANIS; JOSHI, 2012). O exame de radiografia é extremamente valioso para avaliação de um derrame subpulmonar (KARKHANIS; JOSHI, 2012). É muito sensível, detectando derrames tão pequenos quanto 5 mL em estudos experimentais e deve ser um exame de rotina (KARKHANIS; JOSHI, 2012).

A Imagem 5, mostra duas radiografias. Radiografia (A) com incidência de derrame pleural e Radiografia (B) mostrando um pulmão normal (KARKHANIS; JOSHI, 2012).

Figura 5 – (A) Radiografia de tórax pósterio anterior com incidência de derrame pleural e (B) Radiografia mostrando o pulmão normal (KARKHANIS; JOSHI, 2012).



Fonte: (WARE; MATTHAY, 2005).

3.1.5 *Enfisema pulmonar ou Emphysema*

O enfisema pulmonar é definido como "aumento permanente anormal dos espaços aéreos distais aos bronquíolos terminais acompanhado pela destruição de suas paredes" (MACNEE, 2008). O enfisema não pode ser curado, o tratamento consiste em aliviar os sintomas

e interromper o avanço da doença (CAMPO *et al.*, 2018). Por isso, o diagnóstico precoce é fundamental para garantir que a qualidade de vida dos pacientes não seja prejudicada (CAMPO *et al.*, 2018).

“As técnicas de diagnóstico atuais incluem exame clínico, testes de funcionalidade pulmonar e radiografia de tórax” (CAMPO *et al.*, 2018). “Os dois primeiros são usados para confirmar ou negar a presença patológica do enfisema (dependendo da expressão dos sintomas), enquanto a técnica de imagem também permite uma avaliação do grau da doença, bem como uma descrição morfológica (como a radiografia representa a estrutura e não função)” (CAMPO *et al.*, 2018).

Embora a radiografia seja rotineiramente utilizada como técnica de imagem para diagnóstico de doenças do tórax, o diagnóstico de enfisema na radiografia é um tema controverso. “Muitos estudos foram realizados para definir diretrizes que possam relacionar os sinais radiológicos observáveis de enfisema com a gravidade do enfisema” (CAMPO *et al.*, 2018). “Porém, um consenso não foi alcançado, pois diferentes observadores têm diferentes taxas de sucesso ao avaliar a gravidade do enfisema em radiografias, e cada observador usa diferentes critérios radiológicos para chegar às suas conclusões” (CAMPO *et al.*, 2018). “Todos os estudos concordam que o diagnóstico de enfisema baseado em raios-X tem uma alta variação ao relacionar sinais radiológicos com gravidade do enfisema” (CAMPO *et al.*, 2018).

3.1.6 Câncer de pulmão ou mass

“O câncer de pulmão é a causa mais comum de morte por câncer no Reino Unido, representando 6% da mortalidade nacional geral e cerca de 35.000 mortes por ano” (BEEK *et al.*, 2015). A maioria dos cânceres de pulmão está relacionada ao tabagismo e a cessação do tabagismo é a maneira mais eficaz de prevenir essa doença frequentemente fatal. A doença pode ser curada, principalmente se detectada precocemente (BEEK *et al.*, 2015). À medida que a epidemia de câncer de pulmão cresceu e se espalhou, foram exploradas formas de detectar a doença mais cedo, para melhorar a taxa de cura. As principais formas para o diagnóstico precoce são a radiografia de tórax e a tomografia computadorizada (BEEK *et al.*, 2015).

Devido à sua ampla disponibilidade, inclusive para médicos da atenção primária, a radiografia de tórax é frequentemente o primeiro exame para o diagnóstico de câncer (HOLLINGS; SHAW, 2002). O câncer de pulmão pode apresentar-se como uma massa espiculada direta, mas sua presença também pode ser inferida de outras aparências, como uma pneumonia

não resolvida ou colapso lobar (HOLLINGS; SHAW, 2002), um exemplo é mostrado na imagem 6.

Em algumas situações, nenhuma imagem adicional é necessária quando uma lesão óssea é identificada (HOLLINGS; SHAW, 2002). No entanto, a tomografia computadorizada do tórax é frequentemente necessária devido à falta de sensibilidade da radiografia na detecção de metástases (HOLLINGS; SHAW, 2002).

Figura 6 – Cancer inferido devido ao colapso lobar (HOLLINGS; SHAW, 2002).



Fonte: (HOLLINGS; SHAW, 2002).

Nas bases de dados utilizadas os termos "mass" e "nodule" se referem a câncer. Como durante a fase de coleta foram utilizados vários repositórios foi tomada a decisão de manter estas classes separadas para caso eventualmente os termos passem a definir doenças diferentes.

3.1.7 Espessamento pleural ou Pleural thickening

O espessamento pleural é um achado comum em radiografias de tórax de rotina. Normalmente envolve o ápice do pulmão, que é chamado de "capa apical pulmonar" (SAITO *et al.*, 2019). Na radiografia de tórax, a calota apical é uma densidade irregular localizada no ápice extremo e tem menos de 5mm de largura (SAITO *et al.*, 2019).

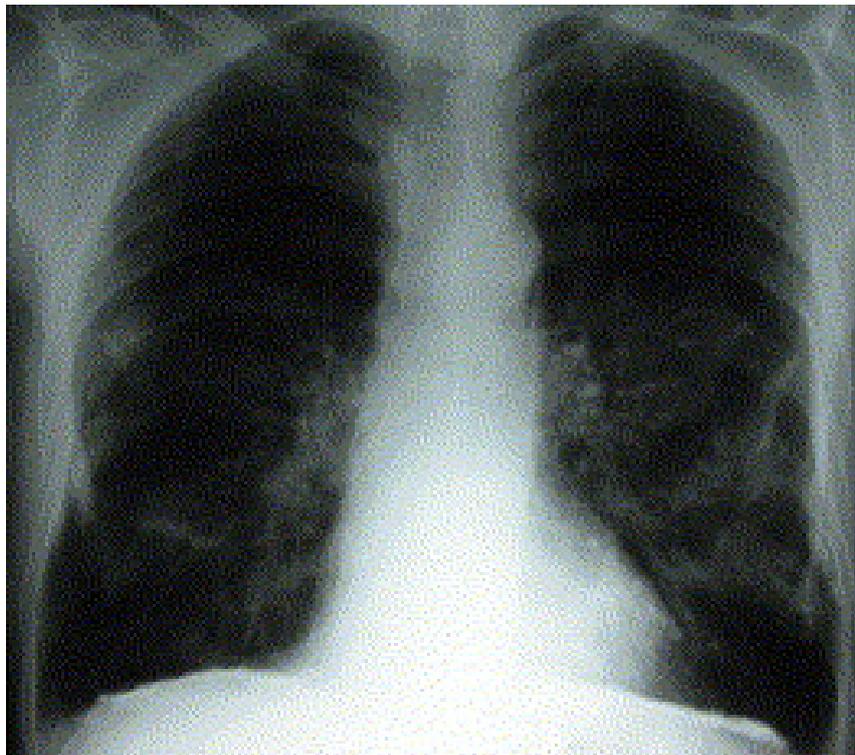
O espessamento pleural pode ser uma manifestação de várias doenças pulmonares,

incluindo infecção micobacteriana, câncer de pulmão e pneumonia intersticial idiopática (SAITO *et al.*, 2019). Portanto, os médicos devem descartar essas condições patológicas ao avaliar o espessamento pleural encontrado em uma radiografia de tórax (SAITO *et al.*, 2019).

As radiografias de um espessamento pleural são geralmente características. A radiografia torácica relatou sensibilidades que variam de 30% a 80% na detecção de placas pleurais (QURESHI; GLEESON, 2006). A detecção depende de vários fatores como, espessura e tamanho; fatores técnicos radiográficos; e a presença de calcificação (QURESHI; GLEESON, 2006). “Em certos pacientes, a detecção de placas pode ser problemática porque a gordura extrapleural proeminente ou sombras compostas da parede torácica adjacente podem imitar a formação de placas, representando uma taxa de falso-positivo de 20%” (QURESHI; GLEESON, 2006).

“Em uma radiografia torácica, as manchas geralmente aparecem como opacidades lisas com menos de 1 cm de espessura paralelas à parede torácica e envolvendo as faces pósterolaterais da sétima a 10^a costelas e a face lateral da sexta à nona costelas” (QURESHI; GLEESON, 2006). Na imagem 7, é mostrada uma radiografia com placas pleurais calcificadas bilaterais devido a um espessamento pleural.

Figura 7 – Radiografia com placas pleurais calcificadas bilaterais. (QURESHI; GLEESON, 2006).



Fonte: (QURESHI; GLEESON, 2006).

3.1.8 *Pneumotórax ou Pneumothorax*

“O pneumotórax ocorre quando o ar se acumula entre o parietal e pleura visceral, causando o colapso do parênquima pulmonar” (CHAN *et al.*, 2020). O pneumotórax traumático geralmente ocorre quando uma costela fraturada danifica o revestimento pleural ou perfura um pulmão com resultante vazamento de ar (CHAN *et al.*, 2020). O Pneumotórax traumático, sem fratura de costela ocorre quando uma força comprime o peito em uma pessoa com glote fechada de repente, aumentando a pressão intratorácica e resultando em ruptura alveolar (CHAN *et al.*, 2020).

O tamanho do pneumotórax é quantificado com base na proporção da cavidade pleural que é ocupada pelo ar, com menos de 15% da cavidade pleural classificada como pequena; 15% a 60% como moderado; e mais de 60% como grande (CHAN *et al.*, 2020; MARX *et al.*, 2013). O pneumotórax oculto, são os que não são inicialmente detectados pelo raio-x de tórax, mas são encontrados na tomografia computadorizada (CHAN *et al.*, 2020; MARX *et al.*, 2013). O pneumotórax de tensão ocorre quando o ar aprisionado significativamente desloca as estruturas mediastinais, reduzindo o fluxo sanguíneo de volta ao coração e resultando em colapso cardiopulmonar com risco de vida (CHAN *et al.*, 2020; MARX *et al.*, 2013).

Figura 8 – Penetração profunda revelando grande pneumotórax esquerdo em paciente ventilado mecanicamente (YARMUS; FELLER-KOPMAN, 2012).



Fonte: (YARMUS; FELLER-KOPMAN, 2012).

As características radiográficas de um pneumotórax podem se apresentar de várias maneiras (YARMUS; FELLER-KOPMAN, 2012). Além dos achados clássicos de colapso pulmonar dentro ou ao redor dos ápices, na posição supina, o ar pode ter uma propensão a se acumular na borda pulmonar. À medida que o ar se acumula, ele pode resultar em um sinal fosco identificando a penetração no tórax (YARMUS; FELLER-KOPMAN, 2012). Na imagem 8, é mostrado um exemplo de pneumotórax esquerdo devido a uma perfuração (YARMUS; FELLER-KOPMAN, 2012).

3.1.9 COVID-19

Nesta seção, é apresentada uma contextualização sobre a propagação da COVID-19, as características da doença, as formas de diagnóstico e como as radiografias podem auxiliar nesse processo diagnóstico.

3.1.9.1 Cronologia e disseminação.

As infecções por Coronavírus (CoV) são uma família de vírus respiratórios emergentes conhecidos por causar doenças que variam desde o resfriado comum até a síndrome respiratória aguda grave (Síndrome respiratória aguda grave (SARS)) (YIN; WUNDERINK, 2018). Os CoV são patógenos zoonóticos, o que significa que podem ser transmitidos de animais para humanos e, posteriormente, de pessoa para pessoa (LI *et al.*, 2020).

Múltiplos surtos epidêmicos ocorreram durante 2002 (SARS), com aproximadamente 800 mortes, e 2012 (Síndrome Respiratória do Oriente Médio: MERS-CoV), com 860 mortes (LI *et al.*, 2020) (BHAGAVATHULA; SHEHAB, 2020). Aproximadamente oito anos após a epidemia de MERS-CoV, o surto atual do novo coronavírus (COVID-19) na cidade de Wuhan, província de Hubei, China, emergiu como um surto global, gerando um problema significativo de saúde pública (LAI *et al.*, 2020).

Em 30 de janeiro de 2020, a Organização Mundial de Saúde (OMS) declarou o COVID-19 uma emergência de saúde pública de interesse internacional (TEAM *et al.*, 2020). Surpreendentemente, na primeira semana de Março, um número devastador de novos casos foi relatado globalmente e o COVID-19 emergiu como uma pandemia (TEAM *et al.*, 2020).

O vírus COVID-19 é principalmente transmitido de pessoa para pessoa por meio de gotículas respiratórias, contato direto e também pode haver transmissão feco-oral (SOHRABI *et al.*, 2020). Após a exposição ao vírus, o período de incubação pode variar de 2 a 14 dias

(SOHRABI *et al.*, 2020). Para combater a propagação da doença, a Organização Mundial da Saúde (OMS) e os Centros de Controle e Prevenção de Doenças (CDC) emitiram recomendações para prevenção, controle e tratamento do COVID-19 no final de janeiro (ORGANIZATION *et al.*, 2020; CONTROL *et al.*, 2020).

A próxima seção detalha o diagnóstico de imagem para a COVID-19 como é realizado atualmente.

3.1.9.2 Diagnóstico de imagem na COVID-19

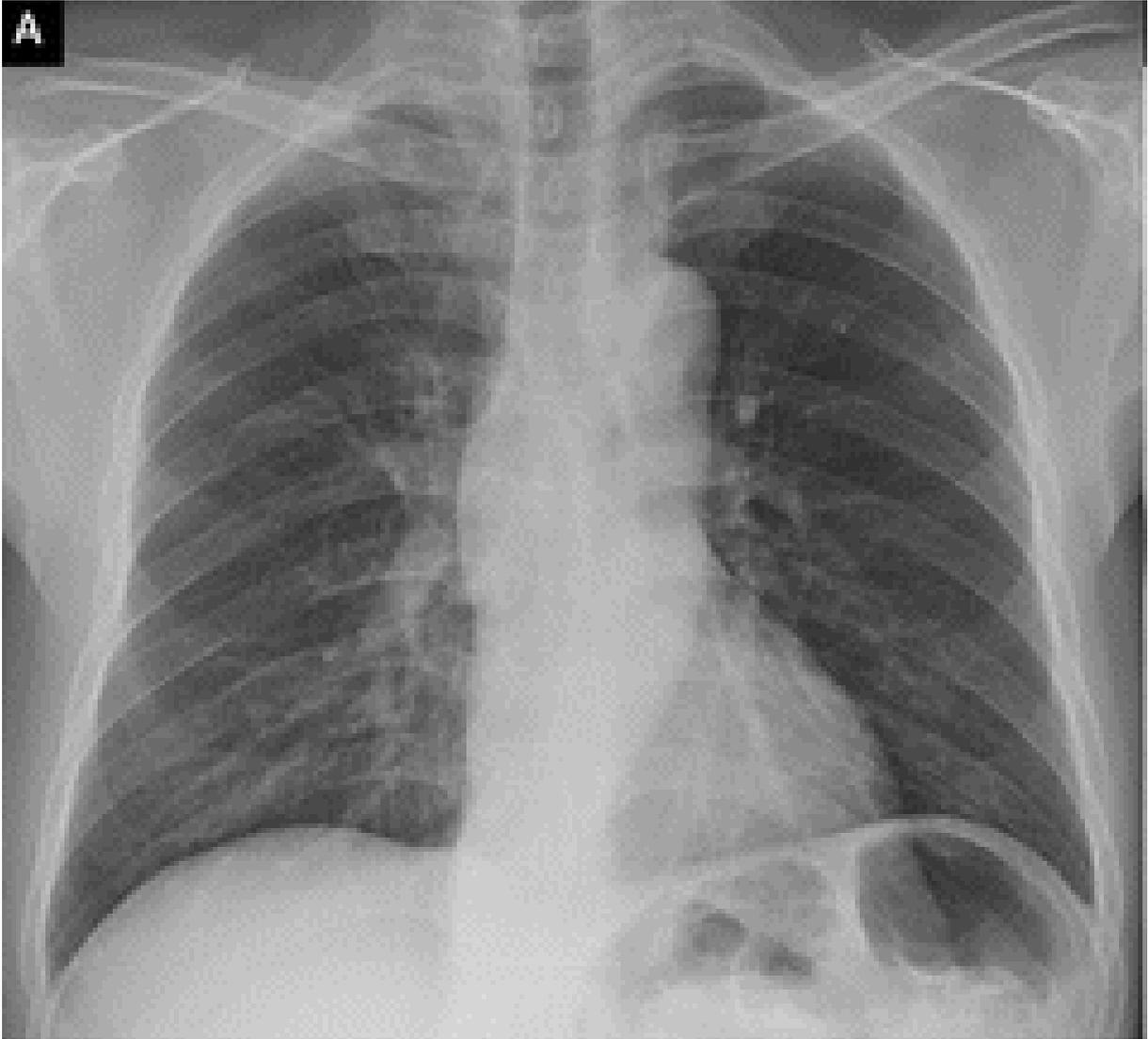
Um caso suspeito é definido como aquele com febre, dor de garganta e tosse ou que teve contato com pacientes com infecção confirmada por COVID-19 (SINGHAL, 2020). No entanto, os casos podem ser assintomáticos ou mesmo sem febre (SINGHAL, 2020). Um caso confirmado é um caso suspeito com teste molecular positivo.

Outros dados laboratoriais geralmente são utilizados para realizar o diagnóstico de COVID-19. Na COVID-19, a contagem de leucócitos geralmente é normal ou baixa. Pode haver linfopenia; uma contagem de linfócitos <1000 associada a doenças graves. A contagem de plaquetas geralmente é normal ou ligeiramente baixa. Um alto nível de procalcitonina pode indicar uma coinfeção bacteriana. Os valores de creatinina, dímero D, CPK e LDH podem estar elevados e níveis elevados estão associados a casos graves (SINGHAL, 2020). A radiografia de tórax geralmente mostra infiltrações bilaterais, algo normal no início da doença (SINGHAL, 2020). A Imagem 9 mostra um pulmão diagnosticado com COVID-19 com infiltrações bilaterais.

Atualmente, o principal teste para confirmar o diagnóstico de COVID-19 é o teste de reação em cadeia da polimerase com transcrição reversa (RT-PCR) (AI *et al.*, 2020). No entanto, a aplicação do teste RT-PCR tende a ser inadequada em muitas áreas que foram gravemente atingidas, especialmente durante o início do surto desta doença (SHI *et al.*, 2020). Esse teste possui uma taxa de acerto de apenas 71% (FANG *et al.*, 2020). Isso se deve a vários fatores, como preparação da amostra e controle de qualidade (KANNE, 2020).

Durante o atendimento clínico, equipamentos de imagem de fácil acesso, como raios-X de tórax e tomografia computadorizada de tórax, tem fornecido uma grande assistência aos médicos (KANNE, 2020; BERNHEIM *et al.*, 2020; XIE *et al.*, 2020; NARIN *et al.*, 2020; APOSTOLOPOULOS; MPESIANA, 2020; WANG; WONG, 2020). Particularmente na China, muitos casos foram identificados como suspeitos de COVID-19, após serem observadas manifestações características em tomografias computadorizadas e radiografias (KANNE, 2020).

Figura 9 – Radiografia de um pulmão com COVID-19.



Fonte: Ng *et al.* (2020).

Pacientes suspeitos de COVID-19, mesmo sem sintomas clínicos como febre e tosse, muitas vezes são hospitalizados ou colocados em quarentena para realizar testes laboratoriais adicionais. Devido à sensibilidade limitada dos testes RT-PCR atuais, é comum que esses pacientes sejam testados várias vezes, com intervalos de dias, até que um diagnóstico confiável seja obtido (WANG; WONG, 2020). Portanto, os diagnósticos baseados em imagens desempenham um papel crucial na restrição da transmissão viral e no combate ao COVID-19 (KANNE, 2020).

O fluxo de trabalho para o diagnóstico baseado em imagem para COVID-19, tomando a radiografia torácica como exemplo, inclui três estágios em geral, ou seja, 1) preparação pré-varredura, 2) aquisição de imagem e 3) diagnóstico de doença (XIE *et al.*, 2020). Na fase de preparação do pré-exame, cada sujeito é instruído e auxiliado por um técnico a posar no leito do paciente de acordo com um determinado protocolo (XIE *et al.*, 2020). No estágio de aquisição de

imagens, as imagens de radiografia são adquiridas durante uma única apnéia (XIE *et al.*, 2020). As radiografias são feitas a partir do nível da entrada torácica superior até o nível inferior com os parâmetros otimizados definidos pelo (s) radiologista (s), com base no formato do corpo do paciente (XIE *et al.*, 2020). A partir dos dados brutos adquiridos, as imagens são reconstruídas e depois transmitidas por meio de sistemas de comunicação e arquivamento de imagens (PACS) para leitura e diagnóstico subsequentes (XIE *et al.*, 2020).

Por conta do fácil acesso a exames de raios-X e da alta probabilidade de aparecer dados que indiquem COVID-19 mesmo no início da infecção, esse trabalho optou por utilizar principalmente imagens desse tipo de exame.

As imagens utilizadas por esse trabalhos fazem parte da base de dados aberta disponível em <https://github.com/ieee8023/covid-chestxray-dataset> (COHEN *et al.*, 2020) que contém imagens e diagnósticos anonimizados de pessoas de diversos países. A próxima seção apresenta os fundamentos da aprendizagem de máquina para a criação do modelo proposto por esse trabalho.

3.2 Aprendizagem de máquina

O aprendizado supervisionado é o tipo de aprendizado mais comum e estudado porque é mais fácil treinar uma máquina para aprender com dados rotulados do que com dados não rotulados (SHUKLA, 2018). O aprendizado supervisionado pode ser usado para resolver dois tipos de problemas: regressão ou classificação (SHUKLA, 2018).

Problemas de regressão, são problemas no qual o objetivo é prever valores contínuos, como tentar prever o custo de uma casa ou o clima externo em graus. Este tipo de problema não tem uma restrição de valor específica porque o valor pode ser qualquer número sem limites (SHUKLA, 2018).

Problemas de classificação, são problemas onde o objetivo é tentar classificar objetos em classes. Um exemplo de problema de classificação acontece quando se quer classificar um objeto como "carro" ou "não carro", é porque o objeto pertence apenas a uma das duas categorias específicas: "carro" e "não carro" e a resposta vai enquadrar o objeto na categoria correspondente. Isso também é chamado de problema de classificação binária (SHUKLA, 2018).

O problema tratado neste trabalho é um problema de classificação, já que o objetivo é classificar as imagens de raio-x em imagens de exames saudáveis ou em imagens que contém alguma doença. Por isso, o modelo a ser utilizado neste projeto será um classificador.

Um classificador é uma função que pega os valores de vários recursos (variáveis independentes ou preditores) em um exemplo (o conjunto de valores de variáveis independentes) e prevê a classe à qual esse exemplo pertence. Em uma imagem de raio-x, as variáveis podem ser a matriz da imagem e a classe pode ser o possível diagnóstico.

Um classificador tem vários parâmetros que devem ser aprendidos a partir dos dados de treinamento. Para isso, um conjunto de exemplos são reservados para esse propósito. O classificador treinado é um modelo de relacionamento entre os atributos do objeto e sua classe no conjunto de treinamento.

Depois de treinado, o classificador pode ser usado para determinar se os recursos usados contêm informações sobre a classe do exemplo. Esse relacionamento é testado usando o classificador treinado em um conjunto diferente de exemplos, os dados de teste. Intuitivamente, a ideia é que, se o classificador realmente capturou a relação entre recursos e classes, ele deve ser capaz de prever as classes de exemplos que não viu antes.

A seguir, são apresentados os tipos de classificadores.

3.2.1 Tipos de Classificadores

A forma como o classificador é gerado depende do tipo do classificador. Alguns tipos de algoritmos de classificação em aprendizado de máquina são: K-Nearest Neighbours, Árvores de decisão, Random Forest e Redes neurais. Cada uma dessas abordagens é apresentada a seguir.

O algoritmo K-Nearest Neighbors (KNN) é uma técnica que utiliza a proximidade como uma medida de "semelhança"(CUNNINGHAM; DELANY, 2020). Ele funciona selecionando um conjunto de pontos previamente rotulados e utiliza esses pontos para aprender como rotular novos pontos (CUNNINGHAM; DELANY, 2020). Para rotular um novo ponto, o algoritmo verifica os pontos previamente rotulados que estão mais próximos desse novo ponto (ou seja, seus vizinhos mais próximos). A proximidade é geralmente expressa em termos de uma função de dissimilaridade. Após verificar os 'k' vizinhos mais próximos, o algoritmo atribui um rótulo com base na maioria dos rótulos dos vizinhos.

Árvores de decisão é uma técnica que cria modelos na forma de uma estrutura de árvore (SAFAVIAN; LANDGREBE, 1991). O algoritmo divide um conjunto de dados em subconjuntos cada vez menores enquanto, ao mesmo tempo, uma árvore de decisão associada é desenvolvida de forma incremental (SWAIN; HAUSKA, 1977). O resultado final é uma árvore

com nós de decisão e nós folha. Um nó de decisão tem dois ou mais ramos e um nó folha representa uma classificação ou decisão (SWAIN; HAUSKA, 1977). O nó de decisão superior em uma árvore que corresponde ao melhor preditor denominado nó raiz. As árvores de decisão podem lidar com dados categóricos e numéricos (SAFAVIAN; LANDGREBE, 1991).

Random Forest ou florestas de decisões aleatórias é um método de aprendizagem de conjunto para classificação e regressão que opera construindo uma infinidade de árvores de decisão no momento do treinamento e gerando a classe que é o modo das classes (classificação) ou previsão média (regressão) das árvores individuais (RESENDE; DRUMMOND, 2018). As florestas de decisões aleatórias corrigem o hábito das árvores de decisão de se ajustar ao seu conjunto de treinamento ou seja, árvores de decisão podem sofrer de sobreajuste (overfitting) (RESENDE; DRUMMOND, 2018). As florestas de decisões aleatórias evitam o sobreajuste na maioria dos casos, pois trabalham com subconjuntos aleatórios das características e constrói árvores menores a partir de tais subconjuntos (RESENDE; DRUMMOND, 2018). Depois do treinamento, as subárvores são combinadas. Esta abordagem torna a computação mais lenta, dependendo de quantas árvores serão construídas nas florestas de decisões aleatórias (RESENDE; DRUMMOND, 2018).

As redes neurais são compostas por unidades chamadas neurônios, organizados em camadas, que transformam um vetor de entrada em uma saída (ZHANG, 2000). Cada neurônio recebe uma entrada, aplica uma função (geralmente não linear) a ela e passa a saída para a próxima camada (LIU *et al.*, 2017). Normalmente, as redes neurais são definidas como alimentação direta, o que significa que cada neurônio envia sua saída para todas as unidades na próxima camada, sem feedback para a camada anterior. As conexões entre os neurônios são ponderadas, e essas ponderações são ajustadas durante a fase de treinamento para adaptar a rede neural ao problema específico em questão (LIU *et al.*, 2017).

O classificador utilizado neste projeto será uma rede neural artificial convolucional. Esse classificador é uma rede neural criada para a classificação de imagens. Esse classificador será explicado na próxima seção.

3.2.2 Rede Neural Convolucional (CNN)

Rede Neural Convolucional (CNN) é uma arquitetura de aprendizado profundo, inspirada no mecanismo de percepção visual natural das criaturas vivas. Em 1959, (HUBEL; WIESEL, 1968) descobriram que as células do córtex visual animal são responsáveis pela

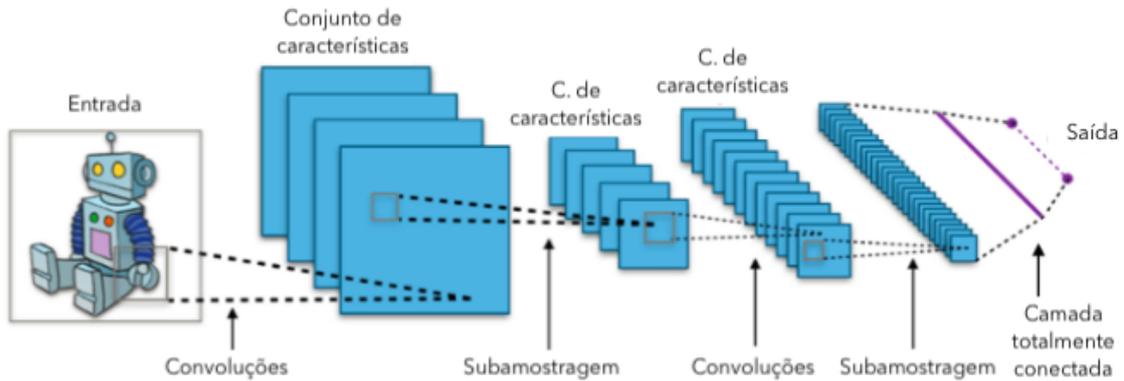
detecção de luz. Essas células atuam como filtros locais sobre o espaço de entrada e são adequadas para explorar a forte correlação espacial local presente em imagens naturais. Sendo o córtex visual animal o sistema de processamento visual mais poderoso que existe, parece natural imitar seu comportamento. Conseqüentemente, muitos modelos inspirados podem ser encontrados na literatura.

Inspirado por esta descoberta, Kunihiko Fukushima propôs o neocognitron em 1980 (FUKUSHIMA; MIYAKE, 1982), que poderia ser considerado o predecessor da CNN. Em 1990, foi publicado um artigo com a estrutura moderna das CNN (LECUN *et al.*, 1989), e posteriormente ela foi melhorada em (LECUN *et al.*, 1998). Neles os autores desenvolveram uma rede neural artificial multicamadas chamada LeNet-5, que poderia classificar dígitos escritos à mão. Como outras redes neurais, o LeNet-5 possuía múltiplas camadas e pode ser treinada com um algoritmo de retropropagação (HECHT-NIELSEN, 1992). Essa rede neural pode obter representações eficazes da imagem original, o que torna possível reconhecer padrões visuais diretamente de pixels brutos com pouco ou nenhum pré-processamento. Em um estudo paralelo de (ZHANG *et al.*, 1990), os autores usaram uma rede neural artificial invariante ao deslocamento, nomeada "SIANN" para reconhecer caracteres de uma imagem. No entanto, devido à falta de grandes dados de treinamento e poder de computação naquele momento, suas redes não puderam ter um bom desempenho em problemas mais complexos, por exemplo, imagem em grande escala e classificação de vídeo.

Uma CNN é composta por um conjunto de vários blocos: camadas de convolução, camadas de pooling (por exemplo, pooling máximo) e camadas totalmente conectadas (FC) (GU *et al.*, 2018). O desempenho de um modelo sob kernels e pesos específicos é calculado com uma função de perda por meio de propagação direta em um conjunto de dados de treinamento, e os parâmetros aprendíveis, ou seja, os kernels e pesos, são atualizados de acordo com o valor de perda por meio de uma retropropagação com algoritmo de otimização de gradiente descendente (GU *et al.*, 2018). A Imagem 10 mostra um exemplo de uma arquitetura de uma CNN.

A camada convolucional trata da convolução entre objetos. É um bloco central na arquitetura de uma CNN. Essa camada é responsável pela maior parte da carga computacional da rede. Ela realiza uma operação chamada convolução, que é uma multiplicação entre uma matriz conhecida como filtro ou kernel (conjunto de parâmetros aprendíveis) e uma matriz que representa uma parte restrita da imagem. A convolução tem como objetivo extrair recursos como bordas, cores e cantos da entrada. À medida que avançamos na rede, a camada convolucional

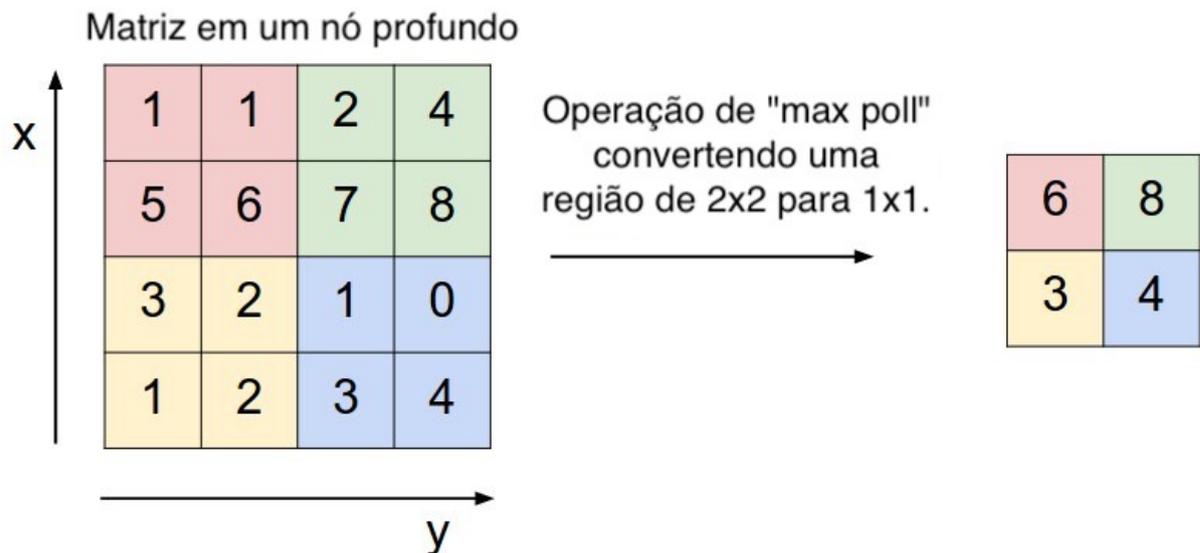
Figura 10 – Exemplo da arquitetura de uma CNN com convoluções, pooling e camadas totalmente conectadas para função de ativação softmax.



Fonte: (Towards Data Science, 2020).

começa a identificar recursos mais complexos, como formas, dígitos e partes de rosto. No final do processo de convolução, obtemos uma matriz de características com dimensões menores do que a imagem original, mas que representa características mais relevantes (GU *et al.*, 2018).

Figura 11 – Execução da operação de MaxPooling.



Fonte: Stanford (2020).

A camada de pooling serve apenas para diminuir a capacidade computacional necessária para processar os dados (GU *et al.*, 2018). Isso é feito diminuindo ainda mais as dimensões da matriz apresentada. Nesta camada, o algoritmo tenta extrair as características dominantes de uma quantidade restrita de uma região. Existem três tipos de técnicas de Pooling: MaxPooling, SumPooling e AveragePooling (GU *et al.*, 2018). Todas elas seguem o mesmo princípio e só o que muda é na maneira como calculam o valor final (GU *et al.*, 2018).

Tomando como exemplo a Imagem 11. Nela, uma matriz de tamanho (4, 4) é usada

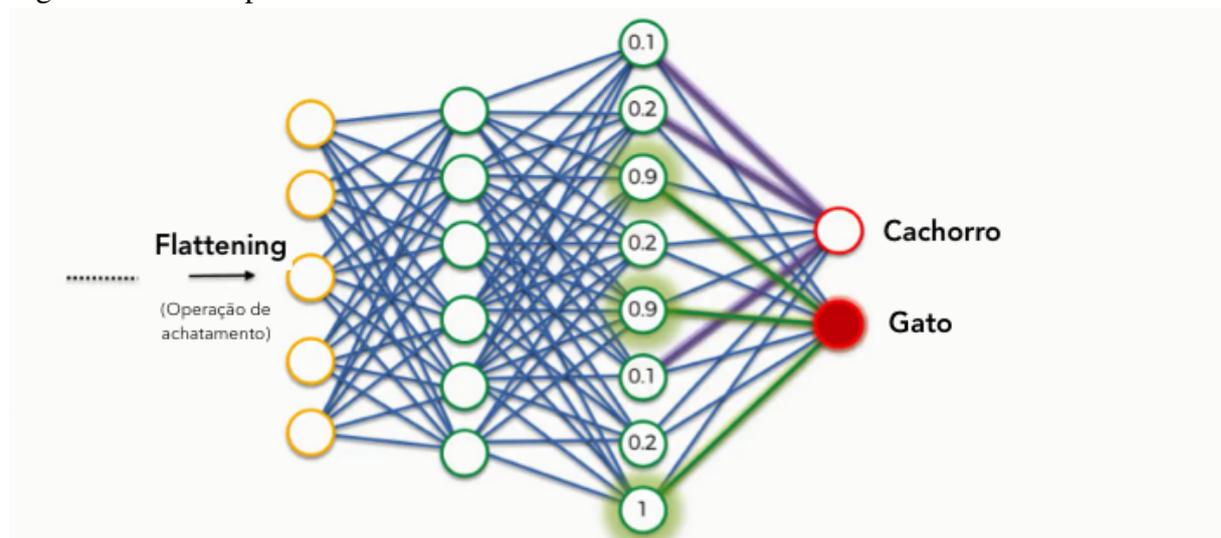
como entrada, à esquerda. Essa matriz foi dividida em quatro regiões, onde cada região representa um kernel de pooling de tamanho (2, 2). Após uma operação de pooling do tipo MaxPooling, onde o objetivo é pegar o maior entre todos os números que estão na região de pool e depois processar a região vizinha da matriz. Ao fim da etapa de pooling, a matriz foi reduzida para o tamanho de (2, 2).

Então, após a camada de pooling, temos uma matriz contendo as principais características da imagem e essa matriz tem dimensões ainda menores, o que vai ajudar no processamento das camadas seguintes.

Camada totalmente conectada (FC), é um Multi Layer Perceptron tradicional que usa uma função de ativação softmax na camada de saída. O termo “Totalmente Conectado” implica que todos os neurônios da camada anterior estão conectados a todos os neurônios da próxima camada.

A saída de uma camada convolucional e de pooling representa características de alto nível extraídas da imagem de entrada. O objetivo da camada totalmente conectada é utilizar essas características para classificar a imagem de entrada em diferentes classes, com base no conjunto de dados de treinamento. Nessa camada, todas as unidades (neurônios) estão conectadas a todas as unidades da camada anterior, permitindo uma combinação complexa das características aprendidas para a tarefa de classificação (GOODFELLOW *et al.*, 2016).

Figura 12 – Exemplo de camada totalmente conectada.



Fonte: Katyal (2020).

A função Softmax é usada para converter um vetor de pontuações arbitrárias em um vetor de probabilidades, onde a soma das probabilidades é igual a um. Ela é comumente aplicada

na camada de saída de uma rede neural totalmente conectada para produzir uma distribuição de probabilidade sobre as classes-alvo. A função Softmax transforma as pontuações em valores entre zero e um, de forma que a soma de todos os valores resultantes seja igual a um. Isso permite interpretar as saídas como probabilidades, indicando a confiança do modelo em relação a cada classe (GOODFELLOW *et al.*, 2016).

Na Imagem 12, o uso de softmax para a camada totalmente conectada fornece os valores de probabilidade para as classes "Cachorro" e "Gato". Nesta CNN, uma imagem é fornecida como entrada. Esta imagem é achatada na camada de flattening (primeira camada). O resultado dessa operação é passado para as outras camadas internas da CNN até chegar na última camada, a camada de softmax. Os nós nesta camada são as classes, e os valores nestes nós representam a probabilidade dela ser a classe correta.

A próxima seção apresenta o framework que será utilizado para dar suporte à criação do modelo nesse trabalho.

3.2.3 *Tensorflow*

A implementação de um modelo de rede neural profunda complexa não é uma atividade trivial e tem sido essencialmente restrita a especialistas. Felizmente, graças à estrutura modular das redes e às ferramentas de inferência padrão, vários frameworks que aceleram o desenvolvimento e o treinamento estão agora disponíveis (RAMPASEK; GOLDENBERG, 2016) (SHUKLA, 2018).

O TensorFlow é a mais nova ferramenta para ajudar na implementação de classificadores. Ele fornece várias melhorias, como visualização gráfica e tempo de compilação aprimorado. Os frameworks de aprendizado profundo mais amplamente usados hoje são Torch7, Theano e Caffé (RAMPASEK; GOLDENBERG, 2016).

TensorFlow, assim como Theano, usa um paradigma de programação declarativo. Isso permite que os desenvolvedores se concentrem na definição simbólica do que precisa ser computado, em vez de como exatamente e em que ordem particular esses cálculos devem ser realizados, o que é o caso na programação imperativa (RAMPASEK; GOLDENBERG, 2016). “O modelo implementado no TensorFlow pode ser otimizado para ser executado por um processador de computador ou chip gráfico. Talvez o maior diferencial no TensorFlow, seja a representação simbólica utilizada internamente que fornece uma maneira simples de otimizar muitas funções. Podem ser redes neurais ou outras funções usadas para representar matematicamente diferentes

problemas orientados por dados” (MENG *et al.*, 2017).

O TensorFlow permite que os desenvolvedores experimentem otimizações e vários algoritmos de treinamento. O TensorFlow oferece suporte a uma variedade de aplicativos, com foco no treinamento e inferência em redes neurais profundas (ABADI *et al.*, 2016).

Nesse trabalho, o framework TensorFlow será utilizado para a criação do modelo.

Depois que o modelo está criado, é necessário avaliar a qualidade do modelo gerado. A próxima seção apresenta os fundamentos da avaliação.

3.3 Avaliação de algoritmos de classificação

O treinamento e avaliação de um classificador são etapas fundamentais em qualquer projeto (COOPER *et al.*, 1997). Para isso, um conjunto de testes é utilizado para avaliar a sua eficiência. Porém, o problema de avaliar a eficiência está na interpretação dos resultados. Seu modelo pode fornecer resultados satisfatórios quando avaliado usando uma métrica, como acurácia, mas pode fornecer resultados ruins quando avaliado em comparação com outras métricas, como função de perda ou qualquer outra métrica (SOKOLOVA *et al.*, 2006). Na maioria das vezes, usamos a precisão da classificação para medir o desempenho do nosso modelo, no entanto, não é suficiente para julgar verdadeiramente o nosso modelo (MATSUBARA, 2008). As métricas discutidas nesta seção serão: acurácia, perda logarítmica, matriz de confusão e Área Sob a Curva (AUC).

Acurácia é a proporção entre o número de previsões corretas e o número total de amostras de entrada, mostrada na Equação 3.1, (FAWCETT, 2006).

$$\text{Acurácia} = \frac{\text{Número de Previsões Corretas}}{\text{Número de Previsões Feitas}} \quad (3.1)$$

A acurácia funciona bem apenas se houver um número igual de amostras pertencentes a cada classe (SOKOLOVA *et al.*, 2006). Por exemplo, considere que há 90% das amostras da classe A e 10% das amostras da classe B em nosso conjunto de treinamento. Então, nosso modelo pode facilmente obter 90% de precisão simplesmente prevendo cada amostra como pertencente à classe A. Quando o mesmo modelo é testado em um conjunto de teste com 60% das amostras da classe A e 40% das amostras da classe B, a precisão do teste cairia para 60%. A precisão de classificação é ótima, mas nos dá a falsa sensação de alcançar alta acurácia. O verdadeiro problema surge quando o custo da classificação incorreta das amostras da classe secundária é

muito alto. Se lidamos com uma doença rara, mas fatal, o custo de deixar de diagnosticar uma pessoa doente é muito maior do que o custo de enviar uma pessoa saudável para mais exames.

A perda logarítmica (loss), penaliza as classificações falsas. Funciona bem para classificação de várias classes (FAWCETT, 2006). Nela, o classificador deve atribuir probabilidade a cada classe para todas as amostras (FAWCETT, 2006). Suponha que existam N amostras pertencentes às classes M , então a perda de log é calculada como na Equação 3.2 (FAWCETT, 2006):

$$loss = \frac{-1}{N} \sum_{i=1}^N \sum_{j=1}^M y_{ij} * \log(p_{ij}) \quad (3.2)$$

Na Equação 3.2, y_{ij} , indica se a amostra i pertence à classe j ou não. p_{ij} , indica a probabilidade da amostra i pertencer à classe j .

A métrica de perda logarítmica não tem limite superior e existe no intervalo $[0, \infty]$. A perda de log mais próxima de 0, indica uma maior precisão, mas se a perda de log está longe de 0, isso indica uma menor precisão. Em geral, minimizar a perda de log fornece maior precisão para o classificador (FAWCETT, 2006) (HUANG; LING, 2005).

Matriz de confusão é uma matriz que descreve o desempenho completo de um modelo (FAWCETT, 2006) (OSI *et al.*, 2020). Por exemplo, supondo que temos um problema de classificação binária. Temos algumas amostras pertencentes a duas classes: SIM ou NÃO. Além disso, temos nosso próprio classificador que prevê uma classe para uma determinada amostra de entrada. Para testar esse modelo, seria utilizada uma matriz de confusão como mostrado na Tabela 1.

Tabela 1 – Matriz de Confusão

		Classificação	
		Positivo (SIM)	Negativo (NÃO)
Classificado Como	Positivo (SIM)	Verdadeiro Positivo (VP) Detecção correta	Falso Positivo (FP) Detecção errada
	Negativo (NÃO)	Falso Positivo (FP) Detecção errada	Verdadeiro Negativo (VN) Detecção correta

Fonte: Filho (2013).

Matriz de confusão é uma matriz que descreve o desempenho completo de um modelo (FAWCETT, 2006) (OSI *et al.*, 2020). Por exemplo, supondo que temos um problema de classificação binária. Temos algumas amostras pertencentes a duas classes: SIM ou NAO. Além

disso, temos nosso próprio classificador que prevê uma classe para uma determinada amostra de entrada. Para testar esse modelo, seria utilizada a Tabela 1

Na matriz de confusão existem 4 termos importantes:

- Verdadeiro Positivos (VP), ocorre quando o objeto classificado pelo classificador recebe a classificação correta.
- Falso Positivos (FP), ocorre quando o objeto é classificado como "SIM" pelo classificador, mas deveria ter sido "NÃO".
- Falso Negativo (FN): ocorre quando o objeto é classificado como "NÃO", mas deveria ser "SIM".
- Verdadeiro Negativo (VN): ocorre quando o objeto é classificado corretamente como "NÃO", sendo assim uma não detecção correta.

A precisão da matriz pode ser calculada tomando a média dos valores situados na "diagonal principal", mostrada na Equação 3.3

$$\text{Acurácia} = \frac{\text{Verdadeiros Positivos} + \text{Verdadeiros Negativos}}{\text{Quantidade de Dados}} \quad (3.3)$$

A análise Receiver Operating Characteristics (ROC) é uma técnica visual utilizada para avaliação de classificadores. A análise ROC é especialmente utilizada em situações onde existem classes desbalanceadas, isto é, quando uma classe tem uma proporção de exemplos maior, ou muito maior, que a outra classe (MATSUBARA, 2008).

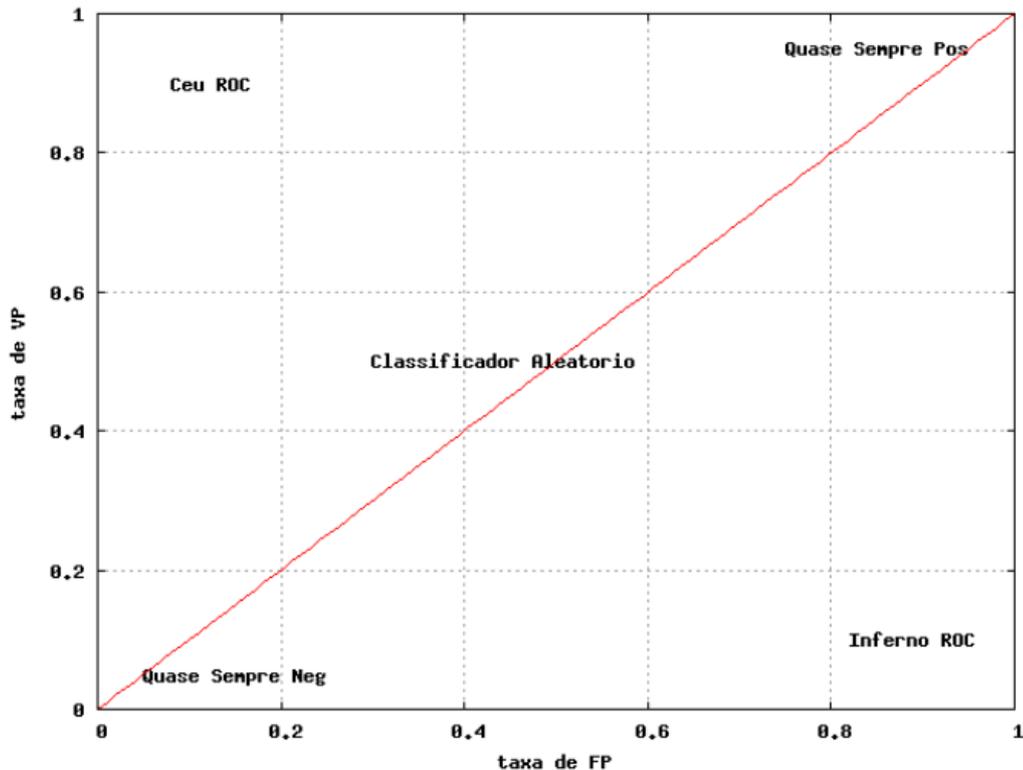
Para efetuar essa análise é utilizado um gráfico bidimensional onde os eixos X e Y representam os falsos positivos (FP) e verdadeiros positivos (VP) (MATSUBARA, 2008). Cada exemplo classificado será representado por um ponto no gráfico ROC e para cada classificador será feito um gráfico ROC. Com esses gráficos será possível comparar os classificadores e analisar qual se adapta melhor ao problema.

A Figura 13 é um exemplo de gráfico ROC. Nela é possível identificar cinco regiões importantes: Céu ROC; Inferno ROC; Quase Sempre Neg e Quase Sempre Pos, e área de classificadores aleatórios.

O “Céu ROC” ocorre quando o classificador sempre retorna o resultado esperado. Nesse quadrante ficam os pontos dos classificadores onde a classificação quase sempre é perfeita.

O “Inferno ROC” acontece quando o classificador sempre erra o resultado esperado. Nesse quadrante, ficam os pontos que representam os classificadores que nunca acertam a classificação.

Figura 13 – Gráfico ROC.



Fonte: Matsubara (2008).

O quadrante “Quase Sempre Neg”, é a região que representa os classificadores que quase sempre classificam os exemplos como negativos. Assim, o número de exemplos classificados errados é baixo, assim como o número de exemplos classificados corretamente.

O quadrante “Quase Sempre Pos”, é a região que representa os classificadores que quase sempre classificam os exemplos como positivos. Assim, quase todos os exemplos positivos são classificados corretamente e quase todos os negativos são classificados incorretamente.

A área sob a curva (AUC) é uma das métricas mais utilizadas para avaliação (SOKOLOVA *et al.*, 2006). A AUC resume à probabilidade (entre 0 e 1) de o classificador classificar um exemplo positivo escolhido aleatoriamente mais alto do que um exemplo negativo escolhido aleatoriamente (FAWCETT, 2006). Um valor AUC mais alto representa a superioridade de um classificador e vice-versa (FAWCETT, 2006).

As métricas de acurácia e perda serão utilizadas durante o treinamento do classificador. As métricas de AUC e ROC serão utilizadas para avaliar o classificador após o treinamento, principalmente para identificar em quais classes o classificador teve os melhores resultados.

3.4 Computação em nuvem

Computação em nuvem é um modelo que envolve o conceito de serviços sob demanda, o que significa usar os recursos da nuvem sob demanda e podemos dimensionar os recursos conforme a demanda (NARULA *et al.*, 2015; YEVGE *et al.*, 2022).

A computação em nuvem é classificada de duas maneiras: com base na localização da computação em nuvem com base no tipo de serviços oferecidos (NARULA *et al.*, 2015; YEVGE *et al.*, 2022).

Com base na localização, elas são classificadas como:

- Nuvem Privada: São infraestruturas alocadas para utilização dedicada por uma determinada organização. Nuvens privadas têm maior custo e segurança em comparação com nuvens públicas (NARULA *et al.*, 2015; YEVGE *et al.*, 2022).
- Nuvem Pública: Refere-se a uma infraestrutura de computação hospedada e gerenciada por provedores de serviços em nuvem. Nesse modelo, os usuários não possuem visibilidade ou controle direto sobre a infraestrutura. A base de computação normalmente é compartilhada entre algumas empresas (NARULA *et al.*, 2015; YEVGE *et al.*, 2022).
- Nuvem híbrida: esse tipo de nuvem é econômica e escalável. Quando combinamos o uso de nuvens públicas e privadas, ela é chamada de nuvem híbrida (NARULA *et al.*, 2015; YEVGE *et al.*, 2022).

Com base no tipo de serviço prestado, elas são classificadas como:

- Infraestrutura como Serviço (IaaS): Utilizando os princípios da computação em nuvem, são oferecidos serviços relacionados a hardware. Isso inclui serviços de armazenamento ou servidores virtuais (NARULA *et al.*, 2015; YEVGE *et al.*, 2022).
- Plataforma como Serviço (PaaS): A plataforma de desenvolvimento na nuvem é oferecida por eles. Fornecedores distintos fornecem plataformas distintas (NARULA *et al.*, 2015; YEVGE *et al.*, 2022).
- Software as a Service (SaaS): Serviços de software completos são oferecidos na nuvem. O aplicativo de software pode ser acessado pelos usuários hospedados pelo fornecedor da nuvem com base no pagamento por uso (NARULA *et al.*, 2015; YEVGE *et al.*, 2022).

3.4.1 Provedores de serviços em nuvem

Empresas como Google, Amazon e Microsoft estão na corrida para fornecer os melhores serviços em nuvem para seus clientes (GUPTA *et al.*, 2021). Como o mercado dos EUA e do Reino Unido é dominado pela Amazon e pela Microsoft, eles são considerados os principais gigantes da tecnologia do mundo (GUPTA *et al.*, 2021). Atualmente, a Amazon Web Services (AWS) tem a maior participação de mercado em comparação com o Microsoft Azure e o Google Cloud Platform (GCP) (GUPTA *et al.*, 2021).

Neste trabalho, foi escolhido a AWS como provedor da infraestrutura, por terem fornecido um cupom de uso para pesquisa.

3.4.2 Amazon CloudFront

O Amazon CloudFront é um serviço web para entrega de conteúdo. Ele se integra a outros produtos da Amazon Web Services para oferecer aos desenvolvedores e empresas uma forma fácil de distribuir conteúdo com baixa latência, altas velocidades de transferência e baixo custo (VARIA *et al.*, 2014).

O Amazon CloudFront pode ser usado para entregar todo o site, incluindo conteúdo dinâmico, estático e de streaming usando uma rede global de pontos de presença. As solicitações de objetos são roteadas automaticamente para o ponto de presença mais próximo, para que o conteúdo seja entregue com o melhor desempenho possível. O Amazon CloudFront é otimizado para funcionar com outros Amazon Web Services, como Amazon S3 e Amazon EC2. O Amazon CloudFront também funciona perfeitamente com qualquer servidor de origem, que armazena as versões originais e definitivas de seus arquivos (VARIA *et al.*, 2014). Assim como outros Amazon Web Services, não há valor mensal mínimo para usar o Amazon CloudFront, você paga apenas pelos dados que são entregues pelo serviço (VARIA *et al.*, 2014).

Este serviço será usado como camada de cache para reduzir a latência de acesso ao site e ao modelo e reduzir o custo de provisionamento.

3.4.3 S3

S3 é o serviço de armazenamento para uso geral fornecido pela Amazon, onde os dados dos clientes são organizados por meio de objetos armazenados em buckets (PERSICO *et al.*, 2016). Um bucket é uma unidade lógica de armazenamento identificada exclusivamente e

pertencente a um dos locais em que o provedor implantou sua infraestrutura de armazenamento. Os custos para o cliente dependem da classe de armazenamento (padrão, acesso pouco frequente ou arquivamento de longo prazo) e da região da nuvem na qual o bucket é colocado, de acordo com o uso (PERSICO *et al.*, 2016). Mais detalhadamente, o custo é calculado como a soma de três cotas, dependendo do tamanho do objeto armazenado, do número de solicitações de download e do volume de tráfego transferido (PERSICO *et al.*, 2016).

Esse serviço será usado para armazenar os dados do website, imagens para geração do modelo e o modelo após o treinamento.

3.4.4 Route 53

O Amazon Route53 é um Domain Name System (DNS) escalável (AMAZON, 2022c). O AWS Route53 é usado para executar 3 funções principais em qualquer combinação: registro de um domínio, roteamento de DNS e realização de determinadas verificações de integridade (AMAZON, 2022c).

Neste trabalho, o Route 53 será usado para redirecionar o domínio acessado para conectar ao sistema do Amazon CloudFront acessando os arquivos públicos.

3.4.5 AWS Batch

O Amazon Web Services Batch (AWS Batch) é um serviço da Amazon que permite criar ambientes de computação para executar cargas de trabalho em lote (AMAZON, 2022a).

O AWS Batch provisiona dinamicamente a quantidade e tipo ideais de recursos de computação (por exemplo, instâncias otimizadas para Unidade Central de Processamento (CPU) ou memória) com base no volume (quantidade de execuções) e nos requisitos específicos das execuções (memória e processamento necessários) (AMAZON, 2022a).

O AWS Batch pode provisionar dois tipos de instâncias: Instância Spot e Instância sob demanda. “Com instâncias sob demanda, você paga pela capacidade computacional por segundo, sem qualquer compromisso de longo prazo. Você tem pleno controle sobre o ciclo de vida dela — você decide quando executar, interromper, hibernar, iniciar, reiniciar ou encerrá-la” (AMAZON, 2022b). Com instância Spot, é possível utilizar o excedente de uma instância sob demanda ou seus recursos ociosos (AMAZON, 2022d). A grande vantagem desse tipo de instância é o custo até 95% menor em relação a uma instância sob demanda (AMAZON, 2022d). A desvantagem de usar instâncias spot é o encerramento antecipado com um aviso de

dois minutos (AMAZON, 2022d).

Neste trabalho, o AWS Batch será utilizado para realizar o treinamento do modelo. A fim de reduzir os custos de provisionamento, serão utilizadas instâncias do tipo spot, que são instâncias de computação com preços mais baixos, disponíveis quando há capacidade não utilizada na nuvem da Amazon.

Para lidar com o término antecipado das instâncias spot e evitar a perda do progresso do treinamento, será implementado um mecanismo de salvamento do estado do modelo. A cada processamento de uma imagem durante o treinamento, o estado atual do modelo será salvo no serviço de armazenamento S3 da Amazon. Dessa forma, caso uma instância spot seja interrompida, o sistema poderá recuperar o último estado do modelo armazenado e retomar o treinamento a partir desse ponto.

Essa abordagem permite que o treinamento do modelo seja mais resiliente a interrupções de instâncias spot e evita a perda de progresso já alcançado. Ao iniciar um novo treinamento, o sistema irá carregar o último estado armazenado no S3, garantindo a continuidade do processo de treinamento a partir desse ponto.

4 TRABALHOS RELACIONADOS

Nesta seção, serão apresentados alguns trabalhos relacionados à detecção de COVID-19 e diagnóstico de imagens utilizando técnicas de aprendizado de máquina.

4.1 Metodologia da pesquisa

Os termos utilizados para a pesquisa foram formados incorporando termos e sinônimos alternativos usando a expressão booleana 'OR' e combinando termos de pesquisa principais usando 'AND'. Os seguintes termos gerais de pesquisa foram usados para identificação dos trabalhos: "(COVID-19 OR SARS-COV-2 OR PNEUMONIA OR ATELECTASIS OR CARDIOMEGALY OR EDEMA OR PLEURAL-EFFUSION OR EFFUSION OR EMPHYSEMA OR MASS OR PLEURAL-THICKENING OR PNEUMOTHORAX) AND (model OR algorithm) AND (detection OR diagnosis) AND (machine learning OR neural network OR CNN) AND (chest OR breast) AND (x-ray OR CT OR computer tomography OR Tomografia computadorizada (CXR))".

Após a identificação dos termos da pesquisa, foram selecionados os portais digitais relevantes e importantes. Os quatro bancos de dados eletrônicos a seguir foram usados para a pesquisa:

- Google scholar
- ScienceDirect
- ACM Digital Library
- SpringerLink

A pesquisa foi restrita a trabalhos publicados após o mês de Maio de 2020 para evitar trabalhos desatualizados. Após determinar quais bancos de dados eletrônicos pesquisar, foi realizada uma pesquisa inicial para identificar os estudos primários candidatos. Também foram incluídos estudos que foram considerados importantes.

4.2 Principais trabalhos encontrados

No artigo "Covid-19 screening on chest x-ray images using deep learning based anomaly detection", os autores comparam a Tomografia Computadorizada (TC) de tórax e o teste RT-PCR para a detecção de COVID-19. Como a TC de tórax pode ser utilizada para o diagnóstico precoce de pneumonias virais, os autores compararam as anomalias encontradas

após realizar este exame em pacientes com COVID-19 (ZHANG *et al.*, 2020).

Os pacientes foram acompanhados por 15 dias. Os pacientes acompanhados precisavam: (a) apresentar um histórico de viagens ou histórico residencial em áreas endêmicas ou contato com indivíduos dessas áreas com febre ou sintomas respiratórios em 14 dias e (b) apresentar febre ou sintomas respiratórios agudos de causa desconhecida (ZHANG *et al.*, 2020).

Durante o processo, foram realizados diariamente exames de TC de tórax e exames RT-PCR (o exame de referência atual para a confirmação de COVID-19). Eventualmente, todos os pacientes foram diagnosticados com COVID-19 através do teste RT-PCR (ZHANG *et al.*, 2020).

Este estudo, foi realizado com uma amostra de 81 pacientes. Destes pacientes, sessenta e cinco (65) (80,3%) foram diagnosticados com COVID-19 através dos exames RT-PCR e TC de tórax. Quinze (15) pacientes (18,5%), foram diagnosticados com COVID-19 apenas através do exame de TC de tórax. Um (1) paciente (1,2%) não foi diagnosticado com COVID-19 pela TC de tórax, sendo confirmado apenas pelo exame RT-PCR (ZHANG *et al.*, 2020).

Em "Automated deep transfer learning-based approach for detection of COVID-19 infection in chest X-rays", os autores tentam demonstrar uma forma mais barata e rápida para detectar e diagnosticar COVID-19 em comparação ao exame RT-PCR (DAS *et al.*, 2020).

A principal forma de detecção de COVID-19 é o exame RT-PCR. No entanto, os kits de RT-PCR são caros e levam de 6 a 9 horas para confirmar a infecção no paciente. Mas devido à alta possibilidade do exame RT-PCR fornecer resultados falso-negativos, geralmente são utilizadas análises de exames de imagens como imagens radiológicas, como raios-X de tórax e tomografia computadorizada (TC), para detectar e diagnosticar COVID-19 (DAS *et al.*, 2020).

Para evitar a possibilidade de classificar os pacientes como falso-negativos e reduzir o custo do exame, os autores criaram um modelo de aprendizagem profunda para detectar e diagnosticar COVID-19 utilizando radiografias de tórax. A razão por trás da escolha das radiografias é que as máquinas de raios X são mais facilmente encontradas em hospitais, já que as máquinas de raios X são mais baratas do que as de tomografia computadorizada (DAS *et al.*, 2020).

Geralmente, a COVID-19 revela algumas anomalias radiológicas que podem ser facilmente detectadas em radiografias de tórax por especialistas. Como os radiologistas precisam analisar essas anomalias manualmente, isso se torna uma atividade demorada e sujeita a falhas. Automatizar isso vai baratear o diagnóstico de COVID-19 e reduzir os erros causados por falha

humana (DAS *et al.*, 2020).

A análise automática de radiografias de tórax foi feita por meio de um modelo de aprendizado profundo, uma CNN. Esse modelo foi feito utilizando uma rede neural adaptada para classificação de imagens e depois treinada com uma grande quantidade de dados de radiografias de tórax (DAS *et al.*, 2020).

Com a utilização desse modelo foi possível superar a menor sensibilidade do RT-PCR, através da classificação das imagens de raios-X de tórax (DAS *et al.*, 2020).

No trabalho "Diagnosis of Coronavirus Disease (COVID-19) from Chest X-ray images using modified XceptionNet", um modelo de aprendizado profundo para detecção de COVID-19 a partir de imagens de tórax é proposto. Um grande número de métodos baseados em aprendizado profundo está disponível para analisar imagens de tórax, mas quando essas imagens apresentam tanto anomalias causados pela COVID-19 com por outras pneumonias, esses modelos tendem a ter uma baixa taxa de acerto (SINGH *et al.*, 2020).

Os métodos atuais não são capazes de classificar as imagens com COVID-19 e pneumonia com precisão pela semelhança das anomalias. A rede proposta é capaz de classificar pneumonia e COVID-19 em classes separadas, apesar destas anomalias (SINGH *et al.*, 2020).

Para superar essa limitação, os autores criaram um modelo de aprendizagem profunda com 121 camadas baseada na rede neural ResNet, uma CNN especializada na classificação de imagens. A rede neural proposta foi chamada de DenseNet-121, e foi treinada com um banco de 100.000 imagens de raios-X de tórax de vista frontal com 14 doenças. A DenseNet-121 classificava a imagem em três classes: COVID-19, Influenza-A-viral-pneumonia e casos saudáveis (SINGH *et al.*, 2020).

O modelo projetado neste artigo é capaz de identificar com sucesso todos os casos de COVID-19 e pneumonia. O modelo proposto apresentou melhorias significativas em relação ao método de teste RT-PCR com uma acurácia de 95,80% e com precisão de 96,16% (SINGH *et al.*, 2020).

Outros trabalhos analisados são demonstrados no Quadro 1. Nestes trabalhos, na maioria o objetivo foi detectar a doença COVID-19 em imagens de Raio-X (9 artigos), em imagens de CT (2 artigos). Apenas dois artigos (RAJPURKAR *et al.*, 2018) e (PUNN; AGARWAL, 2020) tentaram diferenciar entre imagens de pacientes saudáveis, com COVID-19 e outras enfermidades. Nestes trabalhos, em ambos os casos foram usadas imagens de Raio-X para a classificação.

No trabalho (NARIN *et al.*, 2020), os autores fizeram a classificação de imagens de Raio-X e CT para a detecção de COVID-19 e outras pneumonias. Para essa classificação os autores criaram três modelos. Mas em todos os modelos utilizados a acurácia e precisão foram próximas, comparando as imagens de Raio-X e CT. As imagens de Raio-X foram um pouco melhor. Os autores acham que isso ocorreu por conta de que existia uma maior quantidade de imagens desse tipo na base utilizada para o treinamento.

Outro trabalho de destaque, é o (MAJKOWSKA *et al.*, 2020). O objetivo deste trabalho foi desenvolver modelos de aprendizagem profunda para detectar anomalias clinicamente significativas em radiografias de tórax. Para o treinamento e validação dos modelos foram usadas as imagens do repositório ChestX-ray14 (MAJKOWSKA *et al.*, 2020; WANG *et al.*, 2017). Essas mesmas imagens serão utilizadas neste trabalho.

Em todos os trabalhos analisados foram utilizados modelos baseados em CNN para o diagnóstico das radiografias e imagens de CT. Isso reafirma a recomendação da utilização de um modelo de uma CNN para a análise de imagens.

4.3 Comparação com trabalhos relacionados

Diversos autores utilizam abordagens diferentes para a detecção de COVID-19 ou outros problemas em exames baseados em imagens. No entanto, este trabalho se destaca, pelo reconhecimento de múltiplas doenças na mesma imagem e por utilizar imagens de raio-x como entrada para o modelo.

Nos trabalhos de (ZHANG *et al.*, 2020), (DAS *et al.*, 2020), (ABBAS *et al.*, 2020) e (ZHENG *et al.*, 2020), os autores realizam uma classificação binária apenas classificando a imagem em "COVID-19" e "Não COVID-19".

Nos trabalhos de (NARIN *et al.*, 2020) e (PUNN; AGARWAL, 2020), embora os autores utilizem exemplos de várias doenças o modelo proposto por eles retorna apenas três (3) classes: "COVID-19" e "Não COVID-19" e "Outra". Dos trabalhos analisados, o único que realiza uma classificação multiclasse é o de (RAJPURKAR *et al.*, 2018). O modelo proposto neste trabalho realiza a classificação em várias classes como "Pneumonias", "Derrame pleural", "Massas pulmonares", "nódulos" e outras, mas não inclui COVID-19 que é a principal doença analisada neste trabalho.

Observando as limitações dos trabalhos encontrados, no modelo proposto ele tem como objetivo classificar imagens de raio-x para detectar vários tipos de doenças como COVID-

19, Atelectasia, Pneumonias, Derrame pleural e várias outras.

Quadro 1 – Quadro com trabalhos relacionados.

Título	Doenças Verificadas	Tipo de Exame	Referência
Covid-19 screening on chest x-ray images using deep learning based anomaly detection	COVID-19	Raio X	(ZHANG <i>et al.</i> , 2020)
Automated deep transfer learning-based approach for detection of COVID-19 infection in chest X-rays	COVID-19	Raio X	(DAS <i>et al.</i> , 2020)
Diagnosis of Coronavirus Disease (COVID-19) from Chest X-ray images using modified XceptionNet	COVID-19 e Pneumonia	CT e Raio X	(SINGH <i>et al.</i> , 2020)
Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks	COVID-19, Pneumonia, SARS e MERS	CT e Raio X	(NARIN <i>et al.</i> , 2020)
Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network	COVID-19	Raio X	(ABBAS <i>et al.</i> , 2020)
A Novel Medical Diagnosis model for COVID-19 infection detection based on Deep Features and Bayesian Optimization	COVID-19 e Pneumonia	Raio X	(NOUR <i>et al.</i> , 2020)
Automated diagnosis of COVID-19 with limited posteroanterior chest X-ray images using Fine-tuned deep neural networks	COVID-19, Pneumonia e Tuberculosis	Raio X	(PUNN; AGARWAL, 2020)
Deep learning-based Detection for COVID-19 From chest CT using weak label	COVID-19	Raio X	(ZHENG <i>et al.</i> , 2020)
Deep learning for chest Radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists	Pneumonia, nódulos, Derrame pleural, e Massas pulmonares	Raio X	(RAJPURKAR <i>et al.</i> , 2018)
Este Trabalho	COVID-19, Pneumonias, Derrame pleural, Massas pulmonares, nódulos e outras.	Raio X	

5 METODOLOGIA

Nesta seção, é mostrada a metodologia utilizada para realizar a classificação das imagens de raio-x.

5.1 Conjunto de Dados

Neste estudo, imagens de raios-X de tórax de vários pacientes com doenças pulmonares, foram obtidas do repositório GitHub de código aberto compartilhado pelo Dr. Joseph Cohen (COHEN *et al.*, 2020). Estes repositórios consistem em imagens de raio-x e tomografias de tórax principalmente de pacientes com síndrome da angústia respiratória aguda (ARDS), COVID-19, síndrome respiratória do Oriente Médio (MERS), pneumonia e síndrome respiratória aguda grave (SARS).

Os experimentos foram realizados com base em um conjunto de dados composto por imagens de raios-X. Esse conjunto de dados incluiu 708 radiografias de pacientes com COVID-19 e 3.775 radiografias de outras doenças. Todas as imagens foram redimensionadas para o tamanho de 128x128 pixels. Esse redimensionamento é realizado automaticamente na primeira camada do modelo, independentemente da imagem de entrada utilizada. Ao utilizar o classificador para esse redimensionamento, o modelo é capaz de encontrar a melhor maneira de realizar essa operação.

5.2 Configuração do Experimento

A linguagem de programação Python foi usada para treinar os modelos de aprendizagem de transferência profunda propostos. Para facilitar a implementação foi utilizado o Tensorflow apresentado na seção 3.2.3. A maioria dos experimentos foram realizados em um computador Macbook executando o MacOS 10.15.6, com processador Intel i7 Dual-Core de 1,7 GHz. Memória 8GB 1600 MHz DDR3 e placa gráfica Intel HD Graphics 5000 1536 MB.

Na terceira fase, ocorre o treinamento do modelo. Essa fase envolve a divisão dos dados em duas partes: conjunto de treinamento e conjunto de teste. O conjunto de treinamento corresponde a 75% do conjunto total de dados, enquanto o conjunto de teste representa 25%. Essa proporção foi escolhida com base em sua eficiência comprovada em estudos como (RAJPURKAR *et al.*, 2018) e (ZHENG *et al.*, 2020) O framework Keras (utilizado para o treinamento) utilizará essa proporção para embaralhar os dois conjuntos antes de cada ciclo de treinamento.

Para o desenvolvimento do modelo, o projeto foi dividido em quatro fases: (i) coleta dos dados; (ii) pré-processamento dos dados; (iii) treinamento; (iv) classificação; (v) validação e avaliação.

5.3 Avaliação

Para realizar a avaliação do modelo, foi reservado 25% do conjunto total de dados. Essa separação tem como objetivo garantir que o modelo não tenha apenas memorizado os dados, evitando assim o overfitting, que ocorre quando o modelo aprende a classificar os exemplos específicos do conjunto de treinamento, mas não generaliza bem para novos dados.

Para avaliação de desempenho, adotamos as métricas de Precisão (ACC) e a área sob a curva (AUC). Elas foram explicadas na Seção 3.3.

6 O MODELO PROPOSTO

O projeto foi dividido em cinco fases: (i) coleta dos dados; (ii) pré-processamento dos dados; (iii) treinamento; (iv) classificação; (v) validação e avaliação.

Na primeira fase do projeto, são coletadas as imagens que serão utilizadas para o treinamento e avaliação dos classificadores. Essas imagens são obtidas a partir do repositório do Dr. Joseph Cohen (COHEN *et al.*, 2020), porém, apenas as imagens de raio-x serão utilizadas no estudo. Um exemplo de uma imagem coletada pode ser observado na Figura 14. É importante ressaltar que todas as imagens já estão previamente rotuladas com seus respectivos diagnósticos.

Figura 14 – Exemplo de raio-x usado no treinamento.

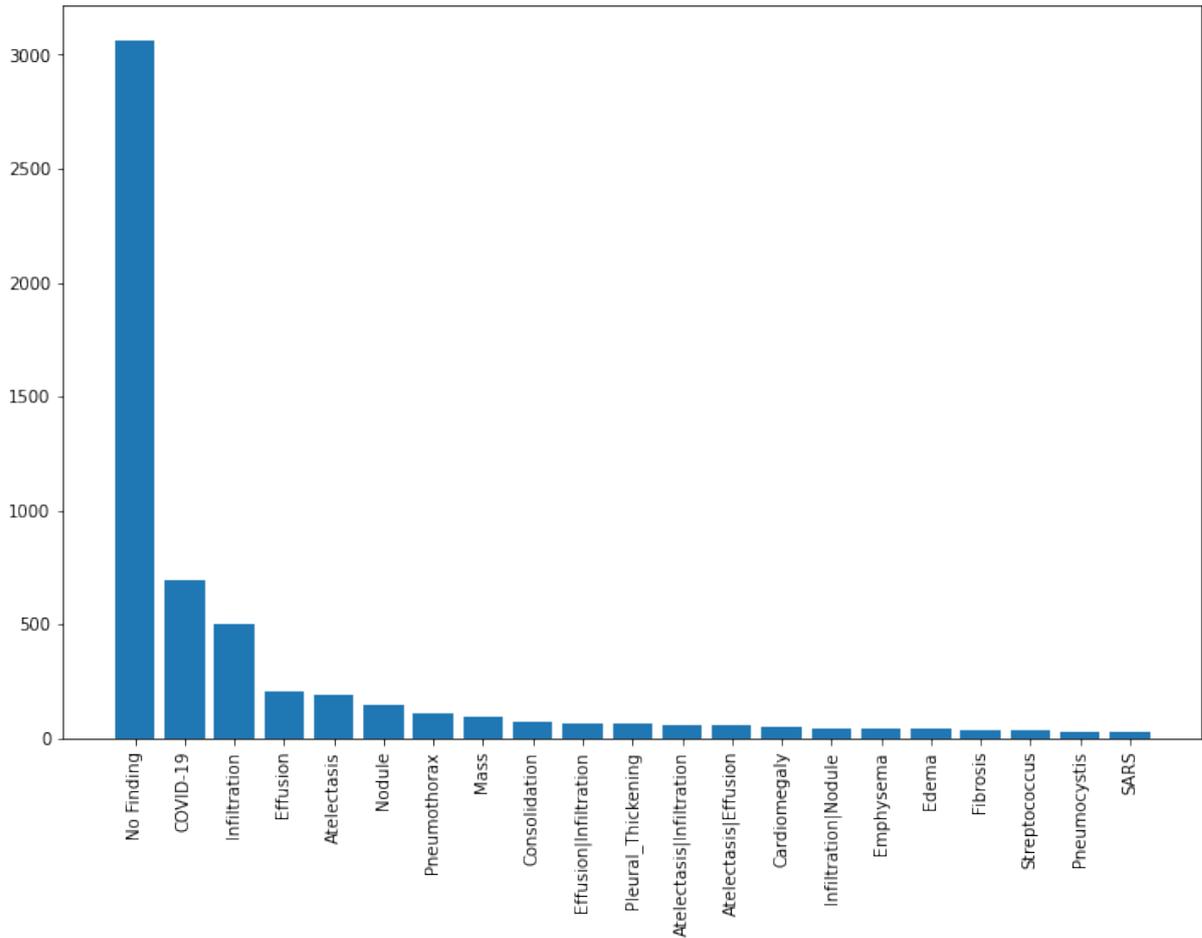


Fonte: (COHEN *et al.*, 2020).

Na segunda fase do projeto, é realizada a etapa de pré-processamento dos dados. Nessa etapa, são aplicadas operações para tratar classes que possuem um número reduzido de exemplos, além de melhorar a aprendizagem das classes desbalanceadas restantes. O objetivo principal é equilibrar a distribuição dos dados entre as diferentes classes, visando aprimorar o desempenho e a capacidade de generalização do modelo. A Figura 15 apresenta a distribuição da quantidade de imagens entre as diferentes classes.

As técnicas de balanceamento de classes são necessárias quando o conjunto de dados apresenta um desequilíbrio na quantidade de exemplos entre as classes, o que pode

Figura 15 – Distribuição de imagens entre as classes.



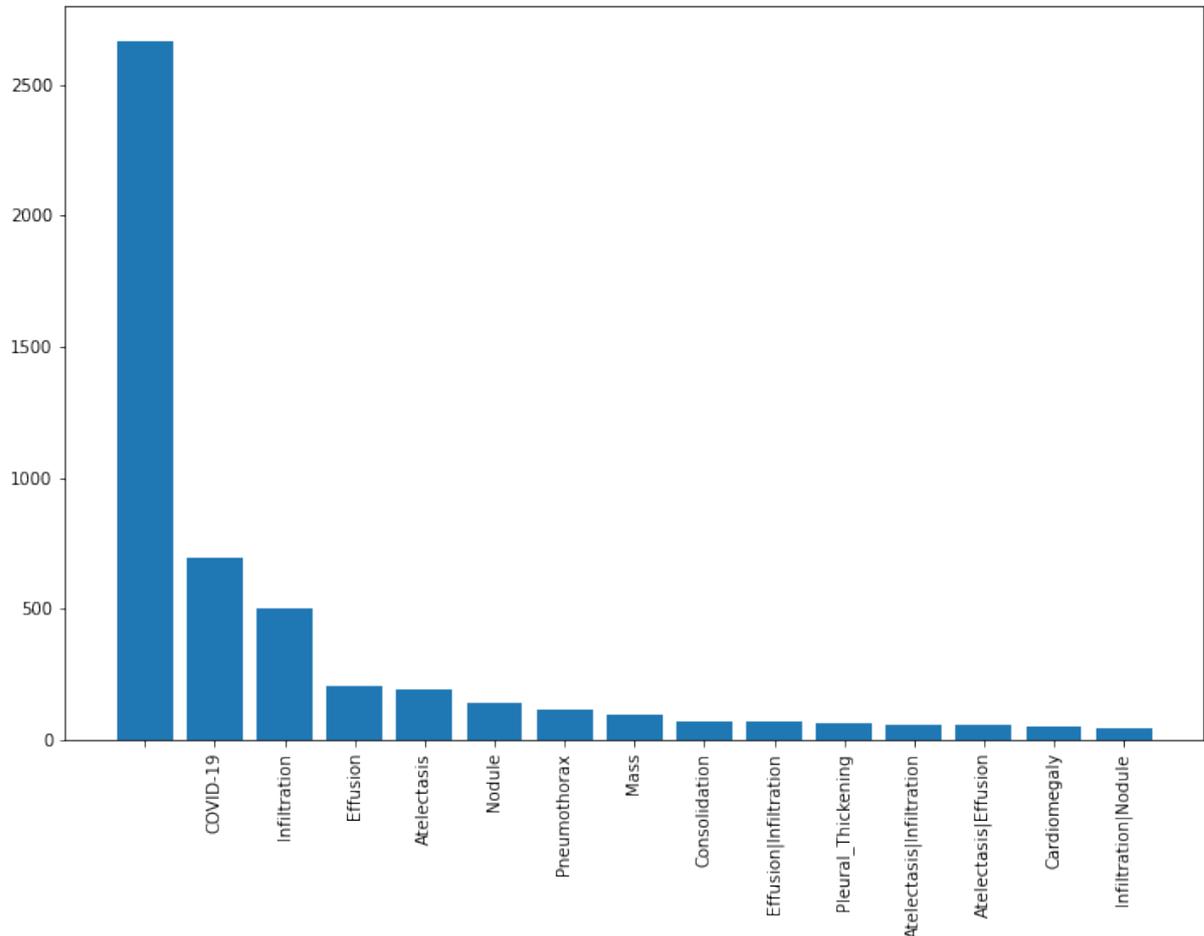
Fonte: O Autor.

levar o classificador a aprender de forma enviesada ou dar mais peso a algumas classes. No contexto desta pesquisa, o conjunto de dados utilizado é altamente desequilibrado, havendo uma quantidade significativamente maior de imagens de pacientes infectados com coronavírus em comparação com outras classes. Portanto, é necessário aplicar técnicas de balanceamento de classe para facilitar o processo de aprendizagem, garantindo que o modelo não seja tendencioso em relação às classes minoritárias.

Neste trabalho, foi utilizada a técnica de classe ponderada para lidar com o desequilíbrio de classes. Nessa abordagem, busca-se equilibrar os dados ajustando os pesos atribuídos a cada classe durante o cálculo da perda. Normalmente, todas as classes têm pesos iguais, mas em casos em que existem classes minoritárias com poucos exemplos, pode ser necessário atribuir um peso maior a essas classes para compensar o desequilíbrio. Os pesos de classe calculados são então incorporados à função objetivo (função de perda) do modelo de aprendizado profundo, a fim de penalizar mais fortemente as previsões incorretas relacionadas às amostras das classes minoritárias. Essa abordagem ajuda a aumentar a importância das classes minoritárias durante o

treinamento e a melhorar o desempenho do modelo nessas classes. A Imagem 16 mostra como ficou a nova distribuição de imagens entre as classes.

Figura 16 – Distribuição de imagens entre as classes.

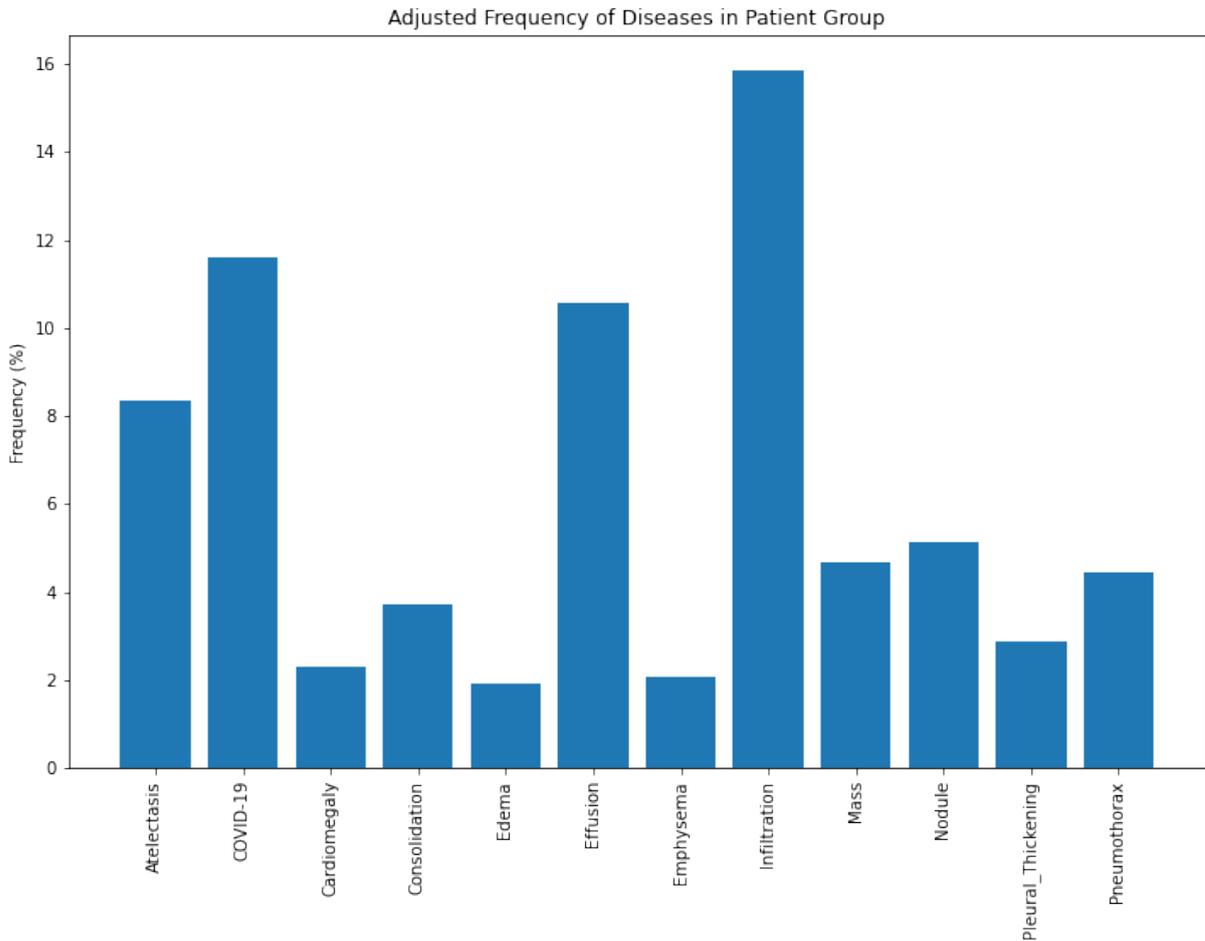


Fonte: O Autor.

Na terceira fase, é feito o treinamento do modelo. A fase de treinamento consiste em primeiro dividir os dados em duas partes e, em seguida, treinar o modelo. A primeira parte é a parte de treinamento que apresenta 75% do conjunto de dados, enquanto a segunda parte é a parte de teste que apresenta 25% do conjunto de dados.

Durante o treinamento, foram aplicadas pequenas distorções nas imagens, como rotações de até 20 graus e ampliações. Essas técnicas têm o objetivo de aumentar a robustez do modelo, permitindo que ele generalize melhor os dados para cada doença, especialmente aquelas que possuem poucos exemplos. Ao expor o modelo a diferentes variações das imagens, espera-se que ele aprenda a reconhecer os padrões relevantes de forma mais eficaz e seja capaz de lidar com variações e ruídos presentes nos dados reais. Isso contribui para melhorar o desempenho do modelo e sua capacidade de generalização para diferentes casos clínicos.

Figura 17 – Percentual de representatividade de cada classe.



Fonte: O Autor.

O modelo proposto foi baseado no modelo apresentado no trabalho (RAJPURKAR *et al.*, 2018; NARIN *et al.*, 2020), mas eles não avaliavam COVID-19 com outras doenças pulmonares. A seleção de 75% para o treinamento e 25% para o teste provou que é eficiente em vários trabalhos como (PUNN; AGARWAL, 2020; ZHENG *et al.*, 2020).

Figura 18 – Estrutura do modelo..

```

model = Sequential([
    Conv2D(16, 3, padding='same', activation='relu', input_shape=(IMG_HEIGHT, IMG_WIDTH, 1)),
    MaxPooling2D(),
    Dropout(0.2),
    Conv2D(32, 3, padding='same', activation='relu'),
    MaxPooling2D(),
    Dropout(0.2),
    Conv2D(64, 3, padding='same', activation='relu'),
    MaxPooling2D(),
    Dropout(0.2),
    Flatten(),
    Dense(512, activation='relu'),
    Dense(len(all_labels), activation = 'sigmoid')
])

multi_disease_model.compile(optimizer = 'adam', loss = 'categorical_crossentropy',
                             metrics = ['categorical_accuracy', 'mae'])

```

Fonte: O Autor.

Nesse modelo, foram adicionadas várias camadas de *Dropout*. O *Dropout* é uma técnica disponibilizada pelo TensorFlow que descarta aleatoriamente uma fração dos dados que passam por essa camada durante o treinamento. Essa abordagem é usada para tornar o modelo mais robusto, reduzindo a probabilidade de que ele memorize excessivamente os valores de entrada. Ao descartar parte dos dados em cada etapa de treinamento, o *Dropout* impede que as unidades individuais do modelo dependam demais de outras unidades específicas, incentivando uma aprendizagem mais generalizada e distribuída. Isso ajuda a evitar o *overfitting* e melhora a capacidade de generalização do modelo para novos dados.

O treinamento foi realizado utilizando lotes (*batch*) de 32 imagens por vez, durante um total de 25 épocas. O treinamento não foi executado por mais épocas, pois após a vigésima quinta época, o classificador não mostrou melhorias significativas nos valores de acurácia e função de perda. Isso indica que o modelo atingiu um ponto de convergência e não se beneficiaria de mais iterações de treinamento. O número de épocas foi determinado com base na observação do desempenho do modelo ao longo do treinamento, buscando encontrar um equilíbrio entre o tempo de treinamento e a obtenção de resultados satisfatórios.

Ao contrário da maioria dos trabalhos relacionados que se concentram em classificar as imagens em apenas duas classes (COVID-19 e não COVID-19), ou seja, modelos binários, este trabalho busca classificar outras doenças pulmonares além do COVID-19. Portanto, o modelo precisa ser capaz de realizar uma classificação multiclasse, onde cada imagem pode ser atribuída a uma das várias classes possíveis. Isso amplia a aplicabilidade e a utilidade do modelo, permitindo que ele seja utilizado para identificar e distinguir diferentes doenças pulmonares com base em imagens de raio-x.

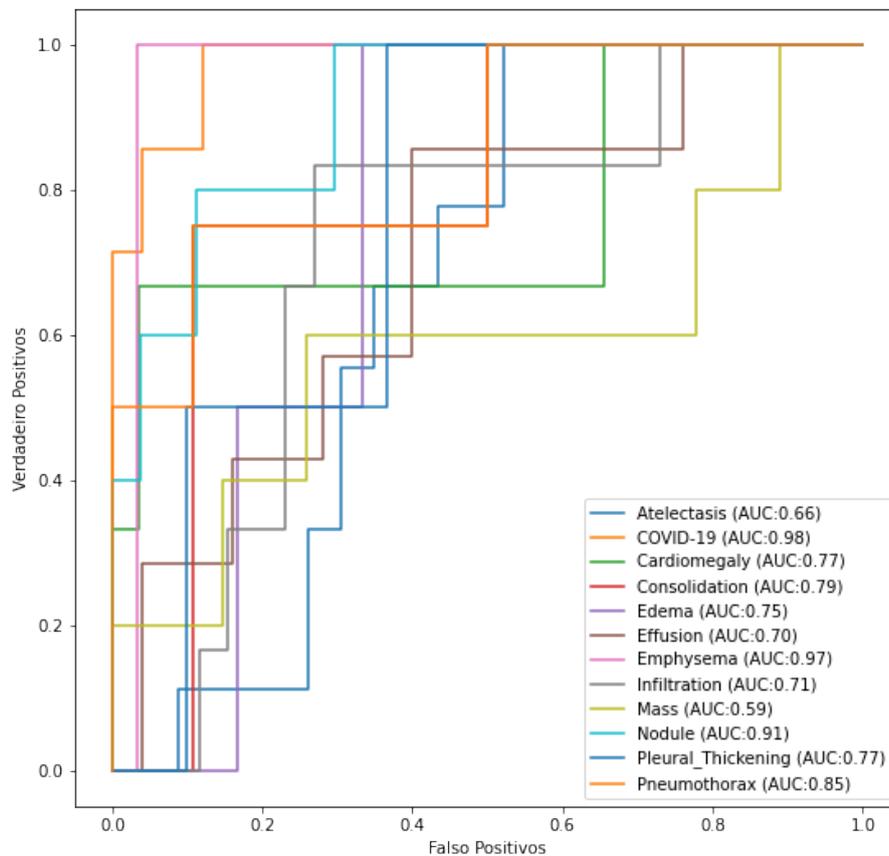
Na quarta fase, é realizada a classificação das imagens do conjunto de teste, que corresponde a 25% dos dados previamente separados. Após a classificação de todas as imagens, é feita uma análise de desempenho do classificador para cada classe, utilizando as métricas descritas na Seção 3.3.

Nessa seção são mostrados os resultados para o modelo proposto conforme descrito na Sessão 6.

A Imagem 19 mostra o gráfico ROC do classificador. Nesse gráfico, cada linha representa uma doença (classe) classificada pelo modelo, e ao lado do nome da doença na legenda de cores, é apresentado o valor AUC correspondente a cada classe. No gráfico, a região dos verdadeiros positivos está no eixo vertical, enquanto a região dos falsos positivos está no eixo

horizontal. A altura da curva representa a taxa de verdadeiros positivos, ou seja, o percentual de objetos corretamente classificados como positivos. O valor AUC representa a área sob a curva ROC e é uma medida da precisão global do classificador, representando o percentual de classificações corretas (verdadeiros positivos + verdadeiros negativos).

Figura 19 – Gráfico ROC e AUC do modelo.



Fonte: O Autor.

Este gráfico mostra que no estado atual do classificador algumas doenças já tem uma grande taxa de acerto. No estado atual do modelo, a maioria das doenças teve uma taxa de Verdadeiros Positivos e Verdadeiros Negativos maior que 60% baseado no valor AUC 3.3 dessas doenças. Um destaque foi a COVID-19 que teve uma taxa de acerto de 98%. A pior doença classificada foi o Câncer categorizado como *mass* com 59% de acerto.

O modelo proposto conseguiu superar trabalhos semelhantes a este. Por exemplo os trabalhos (ZHANG *et al.*, 2020), (DAS *et al.*, 2020), (SINGH *et al.*, 2020), (NARIN *et al.*,

2020) e (ABBAS *et al.*, 2020) que produziram um classificador especializado para identificar COVID-19 tiveram acurácia de 95%, 97%, 95%, 96% e 93% que embora altas ainda ficaram abaixo do acurácia alcançada no modelo proposto que foi de 98%.

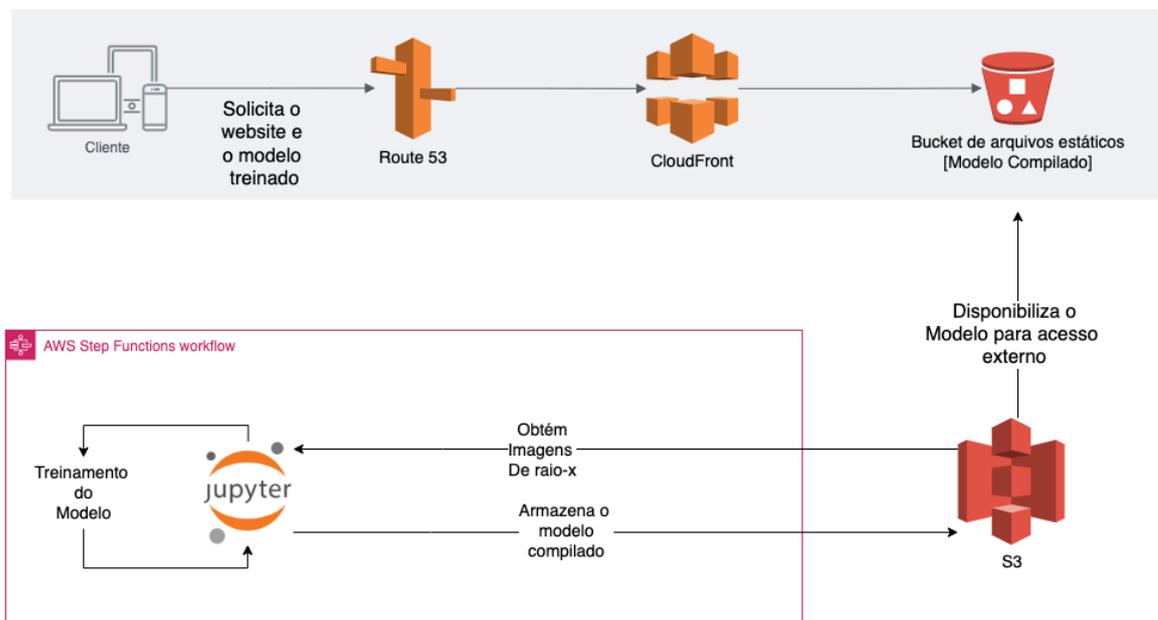
O modelo proposto conseguiu superar trabalhos semelhantes a este. Por exemplo os trabalhos (ZHANG *et al.*, 2020), (DAS *et al.*, 2020), (SINGH *et al.*, 2020), (NARIN *et al.*, 2020) e (ABBAS *et al.*, 2020) que produziram um classificador especializado para identificar COVID-19 tiveram acurácia de 95%, 97%, 95%, 96% e 93% que embora altas ainda ficaram abaixo do acurácia alcançada no modelo proposto que foi de 98%. Ao comparar com o trabalho de (RAJPURKAR *et al.*, 2018), que classifica múltiplas doenças, o modelo proposto conseguiu superá-lo nas doenças Enfisema (97% contra 70% do artigo) e Nódulos (91% contra 85% do artigo). Mas ficou abaixo ao comparar com as doenças Atelectasia (66% contra 86% do artigo) e Massas (59% contra 86% do artigo). Nas outras doenças as acurácias foram aproximadas exceto para COVID-19 que não foi analisado no artigo.

7 SISTEMA WEB

Como um dos objetivos do projeto é criar um modelo que possa evoluir constantemente com novos dados e seja executável em um sistema acessível aos profissionais de saúde para auxiliar no diagnóstico das doenças, foi desenvolvido um website para validar essa ideia. Neste capítulo, descreveremos a arquitetura para o provisionamento do sistema e o funcionamento do website.

7.1 Arquitetura em nuvem

Figura 20 – Arquitetura do sistema web.



Fonte: O Autor.

A arquitetura responsável pelo treinamento do modelo é executada no serviço AWS Batch. Nesse serviço, o sistema coleta as imagens obtidas na fase de coleta que devem estar em um sistema de armazenamento fornecido pela AWS chamado S3. Com essas imagens, são executadas as etapas de pré-processamento e treinamento do modelo. Como está sendo utilizado o TensorFlow, o modelo gerado possui extensão ".h5", o formato padrão usado pelo framework. Como o modelo nesse formato possui um tamanho relativamente grande para o carregamento web (acima de 300Mb), foi necessário convertê-lo para um formato em camadas, facilitando o carregamento através de navegadores. Após essa conversão, o modelo treinado é armazenado no AWS S3, pronto para ser acessado pelo website. O desenho dessa parte da arquitetura pode ser

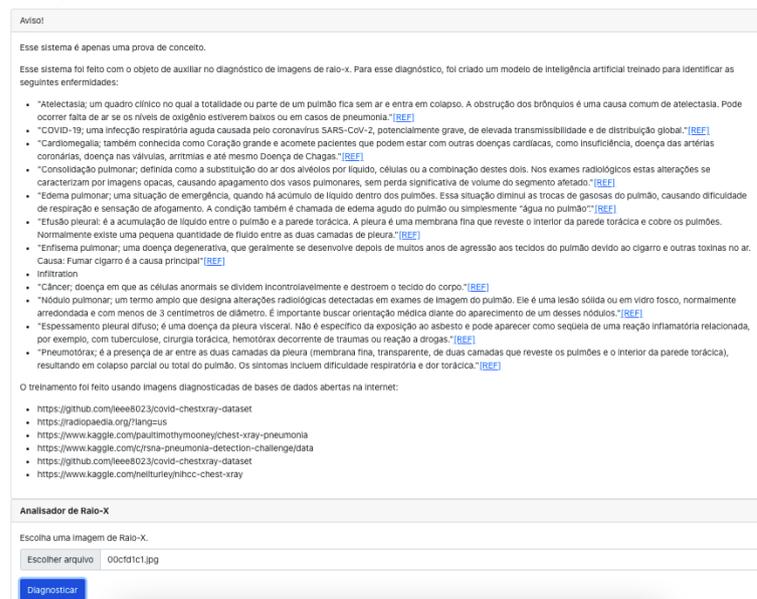
observada no inferior da Figura 20.

A infraestrutura para a disponibilização do sistema web está localizada na parte superior da Figura 20. Após o armazenamento do modelo no AWS S3, os arquivos do modelo são duplicados e armazenados juntamente com os arquivos do website em um bucket (uma espécie de pasta ou diretório no AWS S3). Esses arquivos do bucket podem ser configurados para permitir acesso público de leitura. Em seguida, esse bucket é conectado ao CloudFront da AWS, que é uma ferramenta de cache para arquivos estáticos. Dessa forma, é possível reduzir o custo e o tempo de acesso aos arquivos pelo sistema web. O CloudFront disponibiliza os arquivos por meio de uma URL gerada aleatoriamente pela AWS. Para facilitar a manutenção do sistema web, o serviço AWS Route 53 (sistema de DNS da AWS) é utilizado para mapear a URL gerada pelo CloudFront para uma URL mais amigável e acessível externamente.

7.2 Website

O website implementado para realizar a classificação dos modelos utiliza o processamento no lado do cliente, permitindo assim transferir parte do processamento de classificação do servidor para o cliente. Essa abordagem ajuda a reduzir os custos de provisionamento na nuvem e aumenta a escalabilidade do sistema. A Imagem 21, mostra o website desenvolvido.

Figura 21 – Website para consumir o modelo.



Fonte: O Autor.

O website foi separado em três sessões. Logo no topo existe um aviso de que ele é apenas uma prova de conceito e comenta as doenças que são analisadas e as bases de dados que

foram utilizadas durante o treinamento. A Imagem 22 mostra o texto com o aviso. A segunda parte é o um campo de entrada para o envio da imagem, mostrado na imagem 23. A terceira seção mostra o resultado da classificação como mostrado na Imagem 24.

Figura 22 – Aviso sobre o uso do website.

Aviso!

Esse sistema é apenas uma prova de conceito.

Esse sistema foi feito com o objeto de auxiliar no diagnóstico de imagens de raio-x. Para esse diagnóstico, foi criado um modelo de inteligência artificial treinado para identificar as seguintes enfermidades:

- "Atelectasia; um quadro clínico no qual a totalidade ou parte de um pulmão fica sem ar e entra em colapso. A obstrução dos brônquios é uma causa comum de atelectasia. Pode ocorrer falta de ar se os níveis de oxigênio estiverem baixos ou em casos de pneumonia." [REF]
- "COVID-19; uma infecção respiratória aguda causada pelo coronavírus SARS-CoV-2, potencialmente grave, de elevada transmissibilidade e de distribuição global." [REF]
- "Cardiomegalia; também conhecida como Coração grande e acomete pacientes que podem estar com outras doenças cardíacas, como insuficiência, doença das artérias coronárias, doença nas válvulas, arritmias e até mesmo Doença de Chagas." [REF]
- "Consolidação pulmonar; definida como a substituição do ar dos alvéolos por líquido, células ou a combinação destes dois. Nos exames radiológicos estas alterações se caracterizam por imagens opacas, causando apagamento dos vasos pulmonares, sem perda significativa de volume do segmento afetado." [REF]
- "Edema pulmonar; uma situação de emergência, quando há acúmulo de líquido dentro dos pulmões. Essa situação diminui as trocas de gasosas do pulmão, causando dificuldade de respiração e sensação de afogamento. A condição também é chamada de edema agudo do pulmão ou simplesmente "água no pulmão" [REF]
- "Efusão pleural; é a acumulação de líquido entre o pulmão e a parede torácica. A pleura é uma membrana fina que reveste o interior da parede torácica e cobre os pulmões. Normalmente existe uma pequena quantidade de fluido entre as duas camadas de pleura." [REF]
- "Enfisema pulmonar; uma doença degenerativa, que geralmente se desenvolve depois de muitos anos de agressão aos tecidos do pulmão devido ao cigarro e outras toxinas no ar. Causa: Fumar cigarro é a causa principal" [REF]
- Infiltration
- "Câncer; doença em que as células anormais se dividem incontrolavelmente e destroem o tecido do corpo." [REF]
- "Nódulo pulmonar; um termo amplo que designa alterações radiológicas detectadas em exames de imagem do pulmão. Ele é uma lesão sólida ou em vidro fosco, normalmente arredondada e com menos de 3 centímetros de diâmetro. É importante buscar orientação médica diante do aparecimento de um desses nódulos." [REF]
- "Espessamento pleural difuso; é uma doença da pleura visceral. Não é específico da exposição ao asbesto e pode aparecer como seqüela de uma reação inflamatória relacionada, por exemplo, com tuberculose, cirurgia torácica, hemotórax decorrente de traumas ou reação a drogas." [REF]
- "Pneumotórax; é a presença de ar entre as duas camadas da pleura (membrana fina, transparente, de duas camadas que reveste os pulmões e o interior da parede torácica), resultando em colapso parcial ou total do pulmão. Os sintomas incluem dificuldade respiratória e dor torácica." [REF]

O treinamento foi feito usando imagens diagnosticadas de bases de dados abertas na internet:

- <https://github.com/ieee8023/covid-chestxray-dataset>
- <https://radiopaedia.org/?lang=us>
- <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>
- <https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/data>
- <https://github.com/ieee8023/covid-chestxray-dataset>
- <https://www.kaggle.com/heiturley/nihcc-chest-xray>

Fonte: O Autor.

Figura 23 – Campo para o upload da imagem.

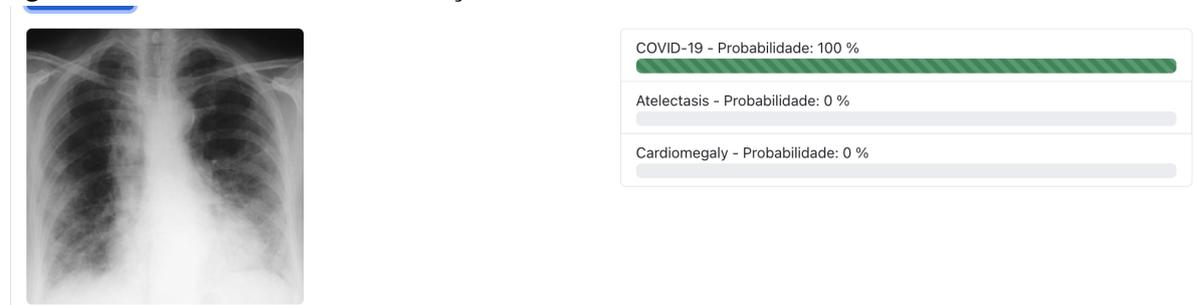
Analizador de Raio-X

Escolha uma imagem de Raio-X.

Escolher arquivo

Fonte: O Autor.

Figura 24 – Resultado da classificação.



Fonte: O Autor.

O website está sendo disponibilizado no endereço "<https://model.samueelfreitas.com.br/index.html>" para acesso.

7.3 Considerações Finais

O objetivo geral deste trabalho é desenvolver um modelo eficiente para auxiliar no diagnóstico de doenças utilizando imagens de radiografias. Para alcançar esse objetivo, foi realizada uma revisão das técnicas de aprendizado de máquina e avaliação de algoritmos de classificação, bem como uma revisão das doenças que serão classificadas pelo modelo proposto. Essa revisão foi realizada para garantir a viabilidade e relevância deste trabalho.

Para o treinamento inicial do modelo, foi realizado um processo de coleta manual de imagens a partir das fontes mencionadas na Seção 5.1. Essas imagens foram submetidas a uma etapa de pré-processamento e, posteriormente, foram utilizadas para o treinamento e validação do modelo.

O modelo obteve uma taxa de acerto aceitável, por conta de ter sido utilizada uma abordagem mais conservadora, onde o algoritmo tem uma alta taxa de acerto, porém uma baixa cobertura. Para trabalhos futuros, seria proposto aprimorar o modelo adicionando outras doenças e outros dados, como sintomas informados para assim melhorar a taxa de acerto e reduzir os falso-positivos.

O objetivo principal do trabalho, que era criar um modelo para auxiliar no diagnóstico de radiografias e torná-lo acessível através de um site, foi alcançado. Para trabalhos futuros, uma proposta seria automatizar a coleta de imagens utilizadas no treinamento, permitindo que o modelo esteja constantemente atualizado com novas variantes das doenças classificadas. Isso garantiria um modelo mais robusto e adaptado às novas situações clínicas que possam surgir.

REFERÊNCIAS

- ABADI, M.; BARHAM, P.; CHEN, J.; CHEN, Z.; DAVIS, A.; DEAN, J.; DEVIN, M.; GHEMAWAT, S.; IRVING, G.; ISARD, M. *et al.* Tensorflow: A system for large-scale machine learning. In: **12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16)**. [S. l.: s. n.], 2016. p. 265–283.
- ABBAS, A.; ABDELSAMEA, M. M.; GABER, M. M. Classification of covid-19 in chest x-ray images using detrac deep convolutional neural network. **arXiv preprint arXiv:2003.13815**, 2020.
- AI, T.; YANG, Z.; HOU, H.; ZHAN, C.; CHEN, C.; LV, W.; TAO, Q.; SUN, Z.; XIA, L. Correlation of chest ct and rt-pcr testing in coronavirus disease 2019 (covid-19) in china: a report of 1014 cases. **Radiology**, Radiological Society of North America, p. 200642, 2020.
- ALGHAMDI, S. S.; ABDELAZIZ, I.; ALBADRI, M.; ALYANBAAWI, S.; ALJONDI, R.; TAJALDEEN, A. Study of cardiomegaly using chest x-ray. **Journal of Radiation Research and Applied Sciences**, Elsevier, v. 13, n. 1, p. 460–467, 2020.
- AMAZON. **Amazon Batch**. 2022. <https://aws.amazon.com/pt/batch/>. Accessed: 2022-06-30.
- AMAZON. **Amazon On-Demand Instances**. 2022. https://docs.aws.amazon.com/pt_br/AWSEC2/latest/UserGuide/ec2-on-demand-instances.html. Accessed: 2022-06-30.
- AMAZON. **Amazon Route53**. 2022. <https://aws.amazon.com/pt/route53/>. Accessed: 2022-06-30.
- AMAZON. **Amazon Spot Instances**. 2022. https://docs.aws.amazon.com/pt_br/AWSEC2/latest/UserGuide/using-spot-instances.html. Accessed: 2022-06-30.
- APOSTOLOPOULOS, I. D.; MPESIANA, T. A. Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. **Physical and Engineering Sciences in Medicine**, Springer, p. 1, 2020.
- BEEK, E. J. V.; MIRSADRAEE, S.; MURCHISON, J. T. Lung cancer screening: Computed tomography or chest radiographs? **World journal of radiology**, Baishideng Publishing Group Inc, v. 7, n. 8, p. 189, 2015.
- BERNHEIM, A.; MEI, X.; HUANG, M.; YANG, Y.; FAYAD, Z. A.; ZHANG, N.; DIAO, K.; LIN, B.; ZHU, X.; LI, K. *et al.* Chest ct findings in coronavirus disease-19 (covid-19): relationship to duration of infection. **Radiology**, Radiological Society of North America, p. 200463, 2020.
- BHAGAVATHULA, A.; SHEHAB, A. The story of mysterious pneumonia and the response to deadly novel coronavirus (2019-ncov): So far. **NEMJ**, v. 1, n. 1, p. 7–10, 2020.
- CAMPO, M. I.; PASCAU, J.; ESTÉPAR, R. S. J. Emphysema quantification on simulated x-rays through deep learning techniques. In: IEEE. **2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)**. [S. l.], 2018. p. 273–276.
- CHAN, K. K.; JOO, D. A.; MCRAE, A. D.; TAKWOINGI, Y.; PREMJI, Z. A.; LANG, E.; WAKAI, A. Chest ultrasonography versus supine chest radiography for diagnosis of pneumothorax in trauma patients in the emergency department. **Cochrane Database of Systematic Reviews**, John Wiley & Sons, Ltd, n. 7, 2020.

- COHEN, J. P.; MORRISON, P.; DAO, L.; ROTH, K.; DUONG, T. Q.; GHASSEMI, M. Covid-19 image data collection: Prospective predictions are the future. **arXiv** **2006.11988**, 2020. Disponível em: <https://github.com/ieee8023/covid-chestxray-dataset>.
- CONTROL, C. for D.; PREVENTION *et al.* Update and interim guidelines on outbreak of 2019 novel coronavirus (2019-ncov). **Available online (Accessed on 31 March 2020).[Link]**, 2020.
- COOPER, G. F.; ALIFERIS, C. F.; AMBROSINO, R.; ARONIS, J.; BUCHANAN, B. G.; CARUANA, R.; FINE, M. J.; GLYMOUR, C.; GORDON, G.; HANUSA, B. H. *et al.* An evaluation of machine-learning methods for predicting pneumonia mortality. **Artificial intelligence in medicine**, Elsevier, v. 9, n. 2, p. 107–138, 1997.
- CUNNINGHAM, P.; DELANY, S. J. k-nearest neighbour classifiers—. **arXiv preprint arXiv:2004.04523**, 2020.
- DAS, N. N.; KUMAR, N.; KAUR, M.; KUMAR, V.; SINGH, D. Automated deep transfer learning-based approach for detection of covid-19 infection in chest x-rays. **IRBM**, Elsevier, 2020.
- FANG, Y.; ZHANG, H.; XIE, J.; LIN, M.; YING, L.; PANG, P.; JI, W. Sensitivity of chest ct for covid-19: comparison to rt-pcr. **Radiology**, Radiological Society of North America, p. 200432, 2020.
- FAWCETT, T. An introduction to roc analysis. **Pattern recognition letters**, Elsevier, v. 27, n. 8, p. 861–874, 2006.
- FILHO, G. M. d. O. Classificação do número de pessoas nas motos em imagens de trânsito corretamente segmentadas. 2013.
- FUKUSHIMA, K.; MIYAKE, S. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. In: **Competition and cooperation in neural nets**. [S. l.]: Springer, 1982. p. 267–285.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep learning**. [S. l.]: MIT press, 2016.
- GU, J.; WANG, Z.; KUEN, J.; MA, L.; SHAHROUDY, A.; SHUAI, B.; LIU, T.; WANG, X.; WANG, G.; CAI, J. *et al.* Recent advances in convolutional neural networks. **Pattern Recognition**, Elsevier, v. 77, p. 354–377, 2018.
- GUPTA, B.; MITTAL, P.; MUFTI, T. A review on amazon web service (aws), microsoft azure & google cloud platform (gcp) services. 2021.
- H, S. W. A. **Cardiomegaly**. 2022. <https://www.ncbi.nlm.nih.gov/books/NBK542296/>. Accessed: 2022-06-30.
- HECHT-NIELSEN, R. Theory of the backpropagation neural network. In: **Neural networks for perception**. [S. l.]: Elsevier, 1992. p. 65–93.
- HOLLINGS, N.; SHAW, P. Diagnostic imaging of lung cancer. **European Respiratory Journal**, Eur Respiratory Soc, v. 19, n. 4, p. 722–742, 2002.
- HUANG, J.; LING, C. X. Using auc and accuracy in evaluating learning algorithms. **IEEE Transactions on knowledge and Data Engineering**, IEEE, v. 17, n. 3, p. 299–310, 2005.

HUBEL, D. H.; WIESEL, T. N. Receptive fields and functional architecture of monkey striate cortex. **The Journal of physiology**, Wiley Online Library, v. 195, n. 1, p. 215–243, 1968.

KANNE, J. P. **Chest CT findings in 2019 novel coronavirus (2019-nCoV) infections from Wuhan, China: key points for the radiologist**. [S. l.]: Radiological Society of North America, 2020.

KARKHANIS, V. S.; JOSHI, J. M. Pleural effusion: diagnosis, treatment, and management. **Open access emergency medicine: OAEM**, Dove Press, v. 4, p. 31, 2012.

KATYAL, R. **Convolutional Neural Networks: Why, what and How!** 2020. Disponível em: <https://blog.usejournal.com/convolutional-neural-networks-why-what-and-how-f8f6dbebb2f9>. Acesso em: 28 jul. 2020.

LAI, C.-C.; SHIH, T.-P.; KO, W.-C.; TANG, H.-J.; HSUEH, P.-R. Severe acute respiratory syndrome coronavirus 2 (sars-cov-2) and corona virus disease-2019 (covid-19): the epidemic and the challenges. **International journal of antimicrobial agents**, Elsevier, p. 105924, 2020.

LECUN, Y.; BOSER, B.; DENKER, J.; HENDERSON, D.; HOWARD, R.; HUBBARD, W.; JACKEL, L. Handwritten digit recognition with a back-propagation network. **Advances in neural information processing systems**, v. 2, 1989.

LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, Ieee, v. 86, n. 11, p. 2278–2324, 1998.

LI, Q.; GUAN, X.; WU, P.; WANG, X.; ZHOU, L.; TONG, Y.; REN, R.; LEUNG, K. S.; LAU, E. H.; WONG, J. Y. *et al.* Early transmission dynamics in wuhan, china, of novel coronavirus–infected pneumonia. **New England Journal of Medicine**, Mass Medical Soc, 2020.

LIU, W.; WANG, Z.; LIU, X.; ZENG, N.; LIU, Y.; ALSAADI, F. E. A survey of deep neural network architectures and their applications. **Neurocomputing**, Elsevier, v. 234, p. 11–26, 2017.

MACNEE, W. Chapter 38 - chronic obstructive pulmonary disease: Epidemiology, physiology, and clinical evaluation. In: ALBERT, R. K.; SPIRO, S. G.; JETT, J. R. (Ed.). **Clinical Respiratory Medicine (Third Edition)**. Third edition. Philadelphia: Mosby, 2008. p. 491–516. ISBN 978-0-323-04825-5. Disponível em: <https://www.sciencedirect.com/science/article/pii/B9780323048255100388>.

MAJKOWSKA, A.; MITTAL, S.; STEINER, D. F.; REICHER, J. J.; MCKINNEY, S. M.; DUGGAN, G. E.; ESWARAN, K.; CHEN, P.-H. C.; LIU, Y.; KALIDINDI, S. R. *et al.* Chest radiograph interpretation with deep learning models: assessment with radiologist-adjudicated reference standards and population-adjusted evaluation. **Radiology**, Radiological Society of North America, v. 294, n. 2, p. 421–431, 2020.

MARX, J.; HOCKBERGER, R.; WALLS, R. **Rosen's Emergency Medicine-Concepts and Clinical Practice E-Book: 2-Volume Set**. [S. l.]: Elsevier Health Sciences, 2013.

MATSUBARA, E. T. **Relações entre ranking, análise ROC e calibração em aprendizado de máquina**. Tese (Doutorado) – Universidade de São Paulo, 2008.

MCADAMS, H. P.; SAMEI, E.; III, J. D.; TOURASSI, G. D.; RAVIN, C. E. Recent advances in chest radiography. **Radiology**, Radiological Society of North America, v. 241, n. 3, p. 663–683, 2006.

- MENG, C.; SUN, M.; YANG, J.; QIU, M.; GU, Y. Training deeper models by gpu memory optimization on tensorflow. In: **Proc. of ML Systems Workshop in NIPS**. [S. l.: s. n.], 2017. v. 7.
- NARIN, A.; KAYA, C.; PAMUK, Z. Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. **arXiv preprint arXiv:2003.10849**, 2020.
- NARULA, S.; JAIN, A. *et al.* Cloud computing security: Amazon web service. In: IEEE. **2015 Fifth International Conference on Advanced Computing & Communication Technologies**. [S. l.], 2015. p. 501–505.
- NG, M.-Y.; LEE, E. Y.; YANG, J.; YANG, F.; LI, X.; WANG, H.; LUI, M. M.-s.; LO, C. S.-Y.; LEUNG, B.; KHONG, P.-L. *et al.* Imaging profile of the covid-19 infection: radiologic findings and literature review. **Radiology: Cardiothoracic Imaging**, Radiological Society of North America, v. 2, n. 1, p. e200034, 2020.
- NOUR, M.; CMERT, Z.; POLAT, K. A novel medical diagnosis model for covid-19 infection detection based on deep features and bayesian optimization. **Applied Soft Computing**, Elsevier, p. 106580, 2020.
- ORGANIZATION, W. H. *et al.* Infection prevention and control during health care when novel coronavirus (ncov) infection is suspected: interim guidance, 25 january 2020. World Health Organization, 2020.
- OSI, A. A.; DIKKO, H. G.; ABDU, M.; IBRAHIM, A.; ISMA'IL, L. A.; SARKI, H.; MUHAMMAD, U.; SULEIMAN, A. A.; SANI, S. S.; RINGIM, M. Z. A classification approach for predicting covid-19 patient survival outcome with machine learning techniques. **medRxiv**, Cold Spring Harbor Laboratory Press, 2020.
- PATTERSON, H. S.; SPONAUGLE, D. N. Is infiltrate a useful term in the interpretation of chest radiographs? physician survey results. **Radiology**, Radiological Society of North America, v. 235, n. 1, p. 5–8, 2005.
- PERONI, D.; BONER, A. Atelectasis: mechanisms, diagnosis and management. **Paediatric respiratory reviews**, Elsevier, v. 1, n. 3, p. 274–278, 2000.
- PERSICO, V.; MONTIERI, A.; PESCAPE, A. On the network performance of amazon s3 cloud-storage service. In: IEEE. **2016 5th IEEE International Conference on Cloud Networking (Cloudnet)**. [S. l.], 2016. p. 113–118.
- PUNN, N. S.; AGARWAL, S. Automated diagnosis of covid-19 with limited posteroanterior chest x-ray images using fine-tuned deep neural networks. **arXiv preprint arXiv:2004.11676**, 2020.
- QURESHI, N. R.; GLEESON, F. V. Imaging of pleural disease. **Clinics in chest medicine**, Elsevier, v. 27, n. 2, p. 193–213, 2006.
- RAJPURKAR, P.; IRVIN, J.; BALL, R. L.; ZHU, K.; YANG, B.; MEHTA, H.; DUAN, T.; DING, D.; BAGUL, A.; LANGLOTZ, C. P. *et al.* Deep learning for chest radiograph diagnosis: A retrospective comparison of the cheXnext algorithm to practicing radiologists. **PLoS medicine**, Public Library of Science San Francisco, CA USA, v. 15, n. 11, p. e1002686, 2018.

RAMPASEK, L.; GOLDENBERG, A. Tensorflow: Biology's gateway to deep learning? **Cell systems**, Elsevier, v. 2, n. 1, p. 12–14, 2016.

RESENDE, P. A. A.; DRUMMOND, A. C. A survey of random forest based methods for intrusion detection systems. **ACM Computing Surveys (CSUR)**, ACM New York, NY, USA, v. 51, n. 3, p. 1–36, 2018.

SAFAVIAN, S. R.; LANDGREBE, D. A survey of decision tree classifier methodology. **IEEE transactions on systems, man, and cybernetics**, IEEE, v. 21, n. 3, p. 660–674, 1991.

SAITO, A.; HAKAMATA, Y.; YAMADA, Y.; SUNOHARA, M.; TARUI, M.; MURANO, Y.; MITANI, A.; TANAKA, K.; NAGASE, T.; YANAGIMOTO, S. Pleural thickening on screening chest x-rays: a single institutional study. **Respiratory research**, BioMed Central, v. 20, n. 1, p. 1–7, 2019.

SHI, F.; WANG, J.; SHI, J.; WU, Z.; WANG, Q.; TANG, Z.; HE, K.; SHI, Y.; SHEN, D. Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for covid-19. **IEEE reviews in biomedical engineering**, IEEE, 2020.

SHUKLA, N. **Machine learning with TensorFlow**. [S. l.]: Manning Publications Co., 2018.

SINGH, K. K.; SIDDHARTHA, M.; SINGH, A. Diagnosis of coronavirus disease (covid-19) from chest x-ray images using modified xceptionnet. **Romanian Journal of Information Science and Technology**, EDITURA ACAD ROMANE CALEA 13 SEPTEMBRIE NR 13, SECTOR 5, BUCURESTI 050711 ..., v. 23, p. S91–105, 2020.

SINGHAL, T. A review of coronavirus disease-2019 (covid-19). **The Indian Journal of Pediatrics**, Springer, p. 1–6, 2020.

SOHRABI, C.; ALSAFI, Z.; O'NEILL, N.; KHAN, M.; KERWAN, A.; AL-JABIR, A.; IOSIFIDIS, C.; AGHA, R. World health organization declares global emergency: A review of the 2019 novel coronavirus (covid-19). **International Journal of Surgery**, Elsevier, 2020.

SOKOLOVA, M.; JAPKOWICZ, N.; SZPAKOWICZ, S. Beyond accuracy, f-score and roc: a family of discriminant measures for performance evaluation. In: SPRINGER. **Australasian joint conference on artificial intelligence**. [S. l.], 2006. p. 1015–1021.

STANFORD. **CS231n Convolutional Neural Networks for Visual Recognition**. 2020. Disponível em: <https://cs231n.github.io/convolutional-networks/>. Acesso em: 28 jul. 2020.

SWAIN, P. H.; HAUSKA, H. The decision tree classifier: Design and potential. **IEEE Transactions on Geoscience Electronics**, IEEE, v. 15, n. 3, p. 142–147, 1977.

TEAM, E. E. *et al.* Note from the editors: World health organization declares novel coronavirus (2019-ncov) sixth public health emergency of international concern. **Eurosurveillance**, European Centre for Disease Prevention and Control, v. 25, n. 5, p. 200131e, 2020.

Towards Data Science. **Understanding CNN (Convolutional Neural Network)**. 2020. Disponível em: <https://towardsdatascience.com/understanding-cnn-convolutional-neural-network-69fd626ee7d4>.

VARIA, J.; MATHEW, S. *et al.* Overview of amazon web services. **Amazon Web Services**, v. 105, 2014.

WANG, L.; WONG, A. Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. **arXiv preprint arXiv:2003.09871**, 2020.

WANG, X.; PENG, Y.; LU, L.; LU, Z.; BAGHERI, M.; SUMMERS, R. M. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S. l.: s. n.], 2017. p. 2097–2106.

WARE, L. B.; MATTHAY, M. A. Acute pulmonary edema. **New England Journal of Medicine**, Mass Medical Soc, v. 353, n. 26, p. 2788–2796, 2005.

XIE, X.; ZHONG, Z.; ZHAO, W.; ZHENG, C.; WANG, F.; LIU, J. Chest ct for typical 2019-ncov pneumonia: relationship to negative rt-pcr testing. **Radiology**, p. 200343–200343, 2020.

YARMUS, L.; FELLER-KOPMAN, D. Pneumothorax in the critically ill patient. **Chest**, Elsevier, v. 141, n. 4, p. 1098–1105, 2012.

YEYGE, A.; GHAG, P.; SOLANKI, C.; MISHRA, A. **Review Paper on Cloud Service Provider–AWS, Azure, GCP**. [S. l.], 2022.

YIN, Y.; WUNDERINK, R. G. Mers, sars and other coronaviruses as causes of pneumonia. **Respirology**, Wiley Online Library, v. 23, n. 2, p. 130–137, 2018.

ZHANG, G. P. Neural networks for classification: a survey. **IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)**, IEEE, v. 30, n. 4, p. 451–462, 2000.

ZHANG, J.; XIE, Y.; LI, Y.; SHEN, C.; XIA, Y. Covid-19 screening on chest x-ray images using deep learning based anomaly detection. **arXiv preprint arXiv:2003.12338**, 2020.

ZHANG, W.; ITOH, K.; TANIDA, J.; ICHIOKA, Y. Parallel distributed processing model with local space-invariant interconnections and its optical architecture. **Applied optics**, Optical Society of America, v. 29, n. 32, p. 4790–4797, 1990.

ZHENG, C.; DENG, X.; FU, Q.; ZHOU, Q.; FENG, J.; MA, H.; LIU, W.; WANG, X. Deep learning-based detection for covid-19 from chest ct using weak label. **medRxiv**, Cold Spring Harbor Laboratory Press, 2020.