

**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO SUDESTE  
DE MINAS GERAIS - CAMPUS RIO POMBA**

Fábio Júnior Barbosa

**CLASSIFICAÇÃO DE RADIOGRAFIAS  
TORÁCICAS COM COVID-19 UTILIZANDO  
REDES NEURAIS CONVOLUCIONAIS**

Rio Pomba

2021

Fábio Júnior Barbosa

# **CLASSIFICAÇÃO DE RADIOGRAFIAS TORÁCICAS COM COVID-19 UTILIZANDO REDES NEURAIIS CONVOLUCIONAIS**

Trabalho de Conclusão curso apresentado ao *campus* Rio Pomba, do Instituto Federal de Educação, Ciência e Tecnologia do Sudeste de Minas Gerais, como parte das exigências do curso de Bacharelado em Ciência da Computação para a obtenção do título de Bacharel em Ciência da Computação.

Orientadora: Profa. Dra. Alessandra Martins Coelho

Coorientador: Matheus de Freitas Oliveira Baffa

Rio Pomba

2021

CLASSIFICAÇÃO DE RADIOGRAFIAS TORÁCICAS COM COVID-19  
UTILIZANDO REDES NEURAS CONVOLUCIONAIS

**Ficha Catalográfica elaborada pela Biblioteca Jofre Moreira – Instituto Federal de Educação, Ciência e Tecnologia do Sudeste de Minas Gerais-Campus Rio Pomba.**  
**Bibliotecária Tatiana dos Reis Gonçalves Ferreira**  
**CRB 6 - 2711**

B238c

Barbosa, Fábio Júnior.

Classificação de radiografias torácicas com Covid-19 utilizando redes neurais convolucionais. / Fábio Júnior Barbosa. – Rio Pomba, 2021.  
35f. : il.

Orientador (a): Prof. Dsc Alessandra Martins Coelho.

Trabalho de Conclusão do Curso Bacharel em Ciência da Computação - Instituto Federal de Educação, Ciência e Tecnologia do Sudeste de Minas Gerais - Campus Rio Pomba.

1. Redes neurais. 2. Visão computacional. 3. Aprendizagem profunda. 4. Covid-19. I. COELHO, Alessandra Martins (Orient.). II.Título.

CDD: 006.3

Fábio Júnior Barbosa

# CLASSIFICAÇÃO DE RADIOGRAFIAS TORÁCICAS COM COVID-19 UTILIZANDO REDES NEURAIS CONVOLUCIONAIS

Trabalho de Conclusão curso apresentado ao *campus* Rio Pomba, do Instituto Federal de Educação, Ciência e Tecnologia do Sudeste de Minas Gerais, como parte das exigências do curso de Bacharelado em Ciência da Computação para a obtenção do título de Bacharel em Ciência da Computação.

Trabalho aprovado. Rio Pomba, 26 de agosto de 2021.

---

**Profa. Dra. Alessandra Martins Coelho**  
Orientadora, IF Sudeste MG - Rio Pomba

---

**Matheus de Freitas Oliveira Baffa**  
Coorientador, USP - Ribeirão Preto

---

**Prof. Dr. Lucas Grassano Lattari**  
IF Sudeste MG - Rio Pomba

---

**Prof. Me. Marjorye Polinati da Silva  
Vecchi**  
IF Sudeste MG - Rio Pomba

Rio Pomba  
2021

*Este trabalho é dedicado a todos aqueles que perderam  
seus entes queridos, amigos, pais, filhos.*

*"O primeiro passo para a cura é saber qual é a doença."  
(Provérbio Latino, 2021)*

# Resumo

Atualmente, não só o Brasil, mas o mundo todo está enfrentando uma crise sanitária gerada por uma doença conhecida como COVID-19. Causada pelo novo coronavírus ou SARS-COV-2, essa doença foi considerada uma pandemia pela Organização Mundial da Saúde (OMS). Por ser uma doença com alta taxa de propagação, é fundamental uma rápida detecção e controle. Para a sua detecção é possível usar testes RT-PCR e de sorologia de anticorpos IgM e IgG, porém esses testes tem um custo mais elevado e sua disponibilidade é limitada. Dessa forma, as imagens médicas têm sido utilizadas para uma rápida visualização e detecção da progressão da doença no pulmão. Assim, este trabalho apresenta uma ferramenta baseada em Visão Computacional, para auxiliar no diagnóstico da COVID-19 utilizando imagens de radiografias do tórax. Uma Rede Neural Convolucional (CNN) para a classificação de imagens foi modelada em três diferentes arquiteturas, com a finalidade de obter resultados de melhor qualidade e menor custo computacional. Usando o protocolo de experimentação de validação cruzada *k-fold* e uma base de dados de 1000 imagens, divididas igualmente entre pacientes saudáveis e de pacientes com COVID-19, obteve-se como resultado uma Acurácia de 99,31%, Sensibilidade de 99,44% e Especificidade de 99,90%, em um dos três modelos considerados para análise, reforçando o alto poder de detecção de padrões das CNNs. A partir dos resultados obtidos, espera-se que este sistema possa auxiliar os médicos na detecção eficaz da doença.

**Palavras-chaves:** Aprendizagem Profunda. COVID-19. Redes Neurais Convolucionais. Visão Computacional.

# Abstract

Recently, not only Brazil, but the whole world is facing a health crisis caused by a disease known as COVID-19 caused by the new coronavirus or SARS-COV-2, considered a pandemic by the World Health Organization (WHO). It is a disease with a high spread rate, rapid detection and control is essential. To detect, it is possible to use RTPCR tests and IgM and IgG antibody serology, but these tests are more expensive and their availability is limited. Thus, medical imaging has been used for rapid visualization and detection of disease progression in the lung. Thus, this work presents a tool based on Computer Vision to assist in the diagnosis of COVID-19 using X-Ray images. In this work, a Convolutional Neural Network was used to classify X-ray images, modeled in three different architectures, in order to obtain a higher quality and lower computational cost. Using the k-fold cross validation experiment protocol and an image base of 1000 images, 500 from healthy patients and 500 from patients with COVID-19, an accuracy of 99.31%, sensitivity of 99.44% and specificity of 99.90% was obtained in one of the three models considered, reinforcing the high pattern detection efficiency of CNNs. From the results obtained, it is expected that this system can help physicians in the effective detection of the disease.

**Key-words:** Deep Learning. COVID-19. Convolutional Neural Networks. Computer vision.



# Lista de ilustrações

Figura 1 – Arquitetura de um neurônio artificial. . . . .	5
Figura 2 – Modelo de Rede Neural Artificial com a introdução de uma camada oculta. . . . .	6
Figura 3 – Modelo de Rede Neural Profunda. . . . .	7
Figura 4 – Entrada tridimensional de uma CNN. . . . .	9
Figura 5 – Camada de convolução. . . . .	10
Figura 6 – Processo de <i>Pooling</i> . . . . .	11
Figura 7 – Exemplo de Max-Pooling. . . . .	11
Figura 8 – Exemplo de camadas de uma rede convolucional. . . . .	12
Figura 9 – Operação de achatamento ou <i>flatten</i> . . . . .	13
Figura 10 – Exemplo de Imagens de Radiografias do Tórax da Base de Dados. . . . .	19
Figura 11 – Representação da arquitetura do modelo I. . . . .	21
Figura 12 – Representação da arquitetura do modelo II. . . . .	22
Figura 13 – Representação da arquitetura do modelo III. . . . .	22
Figura 14 – Representação visual do <i>k-fold</i> para $k = 5$ . . . . .	23
Figura 15 – Acurácia média. . . . .	28

# Lista de tabelas

Tabela 1 – Resultados do <i>k-fold</i> , para o modelo I. . . . .	26
Tabela 2 – Resultados do <i>k-fold</i> , para o modelo II. . . . .	26
Tabela 3 – Resultados do <i>k-fold</i> , para o modelo III. . . . .	27
Tabela 4 – Resultados dos experimentos. . . . .	27

# Lista de abreviaturas e siglas

ACC	Accuracy
AUC	<i>Area Under the Curve</i>
CAD	<i>Computer-Aided Diagnosis</i>
CNN	<i>Convolutional Neural Network</i>
DeTraC	<i>Decompose, Transfer, and Compose</i>
FCNN	<i>Fully-Connected Neural Networks</i>
GPU	<i>Graphics Processing Unit</i>
SVM	<i>Support Vector Machine</i>
MLCP	Memória Longa a Curto Prazo
OMS	Organização Mundial da Saúde
ReLU	<i>Rectified Linear Unit</i>
RT-PCR	<i>Reverse Transcriptase reaction followed by Polymerase Chain Reaction</i>
SEN	Sensibilidade
SPE	Especificidade
SUS	Sistema Único de Saúde

# Sumário

<b>Introdução</b>	<b>1</b>
<b>2 Fundamentação Teórica</b>	<b>4</b>
2.1 Fundamentos de Aprendizado de Máquina	4
2.2 Redes Neurais Artificiais	4
2.3 <i>Deep Learning</i>	7
2.4 Redes Neurais Convolucionais	8
<b>3 Trabalhos Relacionados</b>	<b>14</b>
3.1 Classificação de COVID-19 com Rede Neural Totalmente Conectada	14
3.2 Classificação de COVID-19 com Rede Neural Convolucional	15
3.3 Classificação de COVID-19 com Outras Abordagens	17
3.4 Conclusão Sobre os Trabalhos Relacionados	17
<b>4 Materiais e métodos</b>	<b>18</b>
4.1 Pré-Processamento da Base de Dados	18
4.2 Arquitetura de Rede Neural Artificial	20
4.2.1 Modelo I	20
4.2.2 Modelo II	21
4.2.3 Modelo III	22
4.3 Protocolo de Experimentação	23
4.4 Métricas utilizadas	23
<b>5 Experimentos e Resultados</b>	<b>25</b>
5.1 Especificações de <i>Hardware</i> e <i>Software</i>	25
5.2 Resultados do Protocolo de Avaliação Utilizado	25
5.3 Resultados e Discussão	27
<b>6 Conclusão</b>	<b>29</b>
<b>Referências</b>	<b>30</b>
<b>Glossary</b>	<b>34</b>

# Introdução

Visão Computacional é a área de estudo que simula a capacidade de visão dos seres humanos. Quando falamos de visão, não estamos falando apenas da capacidade de captar uma imagem, mas sim de um processamento complexo, como a interpretação de uma cena. A partir de uma imagem ou cena é possível extrair informações, como objetos de interesse, cores, separação de regiões, forma e textura. Essa área de estudo tem gerado grandes contribuições em várias áreas do conhecimento, como na astronomia, medicina, entretenimento, análise de impressão digital, robótica, entre outras (RUDEK; COELHO; CANGIOLIERI JUNIOR, 2001).

Segundo Backes e Sá Junior (2019), um sistema de Visão Computacional pode ser dividido em aquisição, processamento, segmentação, extração de características e análise de imagens. Na fase de aquisição é feita a captura das imagens, podendo ser por equipamentos eletrônicos, como câmeras, filmadoras e *scanners*. Na fase de processamento são realizados procedimentos para melhorar a imagem retirando ruídos, suavizando a imagem e realçando as bordas, por exemplo. Na fase de segmentação ocorre a divisão da imagem em regiões de interesse. A fase de extração de características é responsável por encontrar uma representação numérica para as características de uma imagem. Já a fase de análise de imagens é responsável por reconhecer características e padrões com o intuito de agrupá-los.

A Visão Computacional pode auxiliar a área médica nos diagnósticos de doenças, através de Sistemas de Auxílio ao Diagnóstico (*Computer-Aided Diagnosis - CAD*). Esses sistemas ajudam na tomada de decisão, funcionando como uma opinião secundária, de modo a melhorar a precisão do diagnóstico (SALOMÃO; MARQUES, 2011).

Sistemas CAD podem ser modelados em várias áreas médicas, a partir de imagens de radiografia, ultrassonografia e ressonância magnética, por exemplo. Com o intuito de melhorar tais imagens e extrair informações úteis para o diagnóstico, são utilizadas técnicas de Processamento de Imagens, Inteligência Artificial e Reconhecimento de Padrões (NUNES, 2006).

O Coronavírus é um ácido ribonucleico viral da família *Coronaviridae*. O vírus responsável pela síndrome respiratória aguda grave (SARS-CoV), o vírus responsável pela síndrome respiratória do oriente médio (MERS-CoV) e o coronavírus 2, causador da síndrome respiratória aguda grave (SARS-CoV-2 ou COVID-19) são da mesma subfamília, infectando apenas mamíferos. Esses vírus são altamente patogênicos e podem causar síndrome respiratória e gastrointestinal (LIMA, 2020).

A COVID-19 foi encontrada pela primeira vez em *Wuhan*, uma província da China, em meados de dezembro de 2019. Essa doença passou a ser declarada uma Emergência de Saúde Pública de Importância Internacional pela Organização Mundial da Saúde (OMS), em 30 de janeiro de 2020 (GARCIA; DUARTE, 2020). No Brasil, isso ocorreu em três de fevereiro de 2020. Já em 11 de março foi considerada pandemia pela OMS. Dados disponibilizados pelo (BRASIL, 2021), em 15 de julho de 2021, informavam 19.262.518 casos confirmados da doença no Brasil, com 17.917.189 pacientes recuperados e 538.942 óbitos (cerca de 2,8% do total de casos confirmados).

A COVID-19 tem uma alta transmissibilidade, sendo necessárias medidas para detecção e diagnóstico precoce. Os exames mais usados atualmente são a reação de transcriptase reversa, seguida de reação em cadeia da polimerase (*Reverse Transcriptase reaction followed by Polymerase Chain Reaction* RT-PCR) e a sorologia. O RT-PCR verifica a existência do vírus através de uma análise do material genético coletado da garganta e do nariz, e a sorologia testa os níveis de anticorpos imunoglobulina M (IgM) e imunoglobulina G (IgG) em amostra de sangue, verificando se o paciente tem imunidade contra o vírus (DASA, 2020).

Os sintomas mais comuns da COVID-19 são tosse, febre, coriza e dor de garganta. No entanto, uma parcela possui o seu quadro agravado, acometendo os pulmões (BRASIL, 2020). Devido a este fato, exames de imagens, como as radiografias do tórax, podem ser usados para a verificação e acompanhamento dos pacientes mais graves, para a indicação de um tratamento mais adequado.

Com o grande número de casos de COVID-19 no Brasil e com um baixo número de testes de diagnóstico inicialmente disponibilizados para atendimento de todos os pacientes, bem como a dificuldade de acompanhamento da evolução da doença nos pulmões, viu-se nas radiografias do tórax uma alternativa para auxiliar os médicos no diagnóstico e acompanhamento da doença. As imagens geradas nesses exames mostram o estado do pulmão, indicando se existe uma pneumonia causada por bactéria ou por vírus (FORATO, 2020).

As imagens radiográficas, em geral, têm um baixo contraste e, por isso, acabam sendo de difícil leitura pelos radiologistas, podendo gerar em muitos casos falsos negativos, pela dificuldade de diferenciação entre uma pneumonia causada por COVID-19 ou outra pneumonia. Por esse motivo, um sistema CAD pode ser muito útil para auxiliar e reforçar o diagnóstico (PATROCÍNIO, 2020).

Assim, o objetivo deste trabalho consiste no desenvolvimento de um método de classificação de imagens torácicas, com a finalidade de verificar se o exame é de um paciente saudável ou com COVID-19. Como objetivos específicos pretende-se (I) apresentar o atual estado da arte sobre classificação da COVID-19 em radiografias do tórax; (II) desenvolver um método de classificação baseado em Redes Neurais Convolucio-

nais; e (III) avaliar a performance da classificação utilizando arquiteturas simplificadas, de modo a manter a eficiência da classificação e reduzir os custos computacionais.

Este trabalho está organizado como se segue. O Capítulo 2 contém toda a base teórica necessária para o entendimento deste trabalho. No Capítulo 3 tem-se os trabalhos relacionados encontrados na literatura. No Capítulo 4 são apresentados os materiais utilizados e o método de classificação desenvolvido. No capítulo 5 são apresentados os experimentos realizados e quais foram os seus resultados. E, por fim, no Capítulo 6, tem-se a conclusão deste trabalho.

## 2 Fundamentação Teórica

Neste capítulo são abordados temas fundamentais para o entendimento do trabalho, como, a Aprendizado de Máquina, que é amplamente utilizada para resolver problemas em diversas áreas, conteúdos sobre Redes Neurais Artificiais, entre outros assuntos relevantes para o entendimento do trabalho.

### 2.1 Fundamentos de Aprendizado de Máquina

A ideia de se criar máquinas que podem pensar (máquinas pensantes) é um tema que reúne cientistas do mundo inteiro e das mais variadas áreas do conhecimento, como Linguística, Psicologia e Ciência da Computação. O objetivo é estudar e criar *softwares* que buscam reproduzir comportamentos inteligentes (TEIXEIRA, 2019).

Por mais que a tecnologia tenha avançado, tentar representar o pensamento humano ainda é um grande desafio. Até o momento não se conseguiu criar uma máquina com a capacidade de reproduzir as funções cognitivas do cérebro humano perfeitamente. Entretanto, do mesmo modo que são criadas máquinas que imitam a capacidade de andar, são desenvolvidas formas de imitar os pensamentos (TEIXEIRA, 2019).

Com o aumento na quantidade de dados, a Inteligência Artificial vem se tornando uma ferramenta cada vez mais útil para a análise desses dados e na tomada de decisões. Pode-se destacar o uso de Aprendizado de Máquina (*Machine Learning*) nas últimas décadas, com grande importância para o desenvolvimento tecnológico. (GOMES, 2019).

O Aprendizado de Máquina tem como objetivo o estudo e construção de regras lógicas para a melhor execução de uma tarefa ou tomada de decisão, conforme o contexto, a partir do reconhecimento de padrões dos dados analisados (ALECRIM, 2018). Suas aplicações são variadas, por exemplo, em Processamento de Linguagem Natural, filtragem de *spam*, reconhecimento de fala, Visão Computacional, diagnóstico médico, sistemas de busca (*internet*), entre muitas outras (GOMES, 2019).

### 2.2 Redes Neurais Artificiais

Redes Neurais Artificiais são sistemas computacionais inspirados no processamento de informações do cérebro humano, sendo formadas por um conjunto de neurônios artificiais que tentam reproduzir, de forma artificial, o comportamento de um



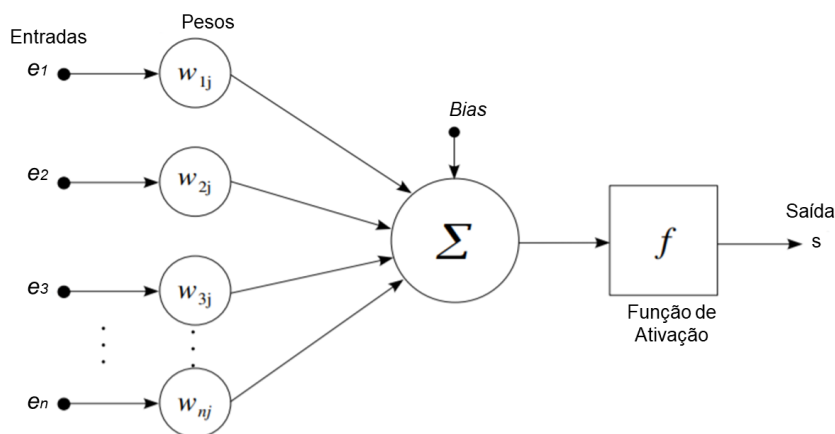
neurônio biológico (MATSUNAGA, 2012).

Na Figura 1 tem-se a representação da arquitetura básica de um neurônio artificial. O neurônio artificial recebe diversos valores de entrada ( $e_1, e_2, e_3, \dots, e_n$ ), que são multiplicados por um valor conhecido como peso ( $w_i$ ). Assim, a primeira entrada é multiplicada por um peso ( $w_1$ ), a segunda entrada por outro peso ( $w_2$ ), e assim sucessivamente. No corpo do neurônio é feito um somatório dos valores, representado na Expressão 2.1.

$$\sum = \sum_{i=1}^n e_i * w_i \quad (2.1)$$

Em seguida, o resultado da Expressão 2.1 é parametrizado para uma função de ativação  $f$  (que será abordada em outros capítulos), gerando uma saída  $S$  (GRÜBLER, 2018).

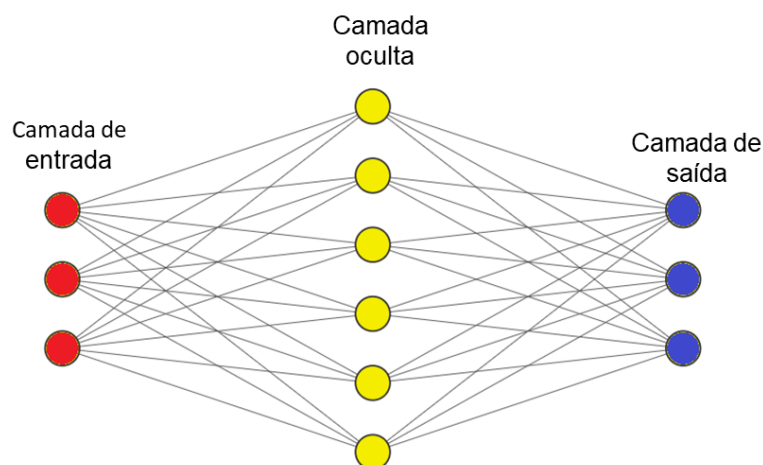
Figura 1 – Arquitetura de um neurônio artificial.



Fonte: Adaptada de Matsunaga (2012)

Uma Rede Neural Artificial pode ser formada por um ou vários neurônios, que são organizados em camadas. A primeira camada é chamada de entrada e a última, de saída. Já as camadas que estão entre a camada de entrada e a camada de saída são chamadas camadas ocultas. Na Figura 2 tem-se a representação de um modelo de Rede Neural Artificial, sendo a camada camada de entrada representada por neurônios em vermelho, a camada oculta em amarelo e a camada de saída em azul (DATA SCIENCE ACADEMY, 2021).

Figura 2 – Modelo de Rede Neural Artificial com a introdução de uma camada oculta.



Fonte: Elaborada pelo autor.

Uma Rede Neural Artificial precisa aprender a reconhecer padrões, e para que isso seja possível é necessário fazer o treinamento. No treinamento são utilizados diferentes conjuntos de dados, compostos por amostras que, no contexto deste trabalho, são imagens. Para cada imagem processada pela Rede Neural Artificial, verifica-se o resultado considerando as taxas de erros e, em seguida, os pesos da rede neural são ajustados com o intuito de gerar um resultado melhor. Comparando o treinamento de uma Rede Neural Artificial com o processo de aprendizado humano, pode-se dizer que ele é semelhante à memorização, ou seja, a rede memoriza os padrões encontrados nas imagens.

O processo de treinamento também é conhecido como aprendizagem e pode ser dividido em Aprendizagem Supervisionada e Aprendizagem Não Supervisionada. Na Aprendizagem Supervisionada, os resultados da rede são comparados ao que se espera como resultado, e os neurônios têm os pesos ajustados de modo a diminuir a diferença entre a saída e o resultado desejado. Já na Aprendizagem Não Supervisionada, são criados grupos com padrões semelhantes, baseado em conceitos de vizinhança e agrupamento (MATSUNAGA, 2012).

Conforme Monard e Baranauskas (2003), alguns conceitos são fundamentais para um bom entendimento do assunto em questão, por exemplo, atributo, classe, conjunto de exemplos, ruído, Viés (*Bias*), Acurácia, distribuição de classes, prevalência de classe.

## 2.3 Deep Learning

*Deep Learning* ou Aprendizagem Profunda é uma subárea do Aprendizado de Máquina. Esta pode ser definida como o campo preocupado com o estudo de Redes Neurais Artificiais que possuem um número maior de camadas ocultas, tendo como destaque aplicações em análise de dados, classificação, reconhecimento de objetos visuais, dentre outras (DATA SCIENCE ACADEMY, 2021).

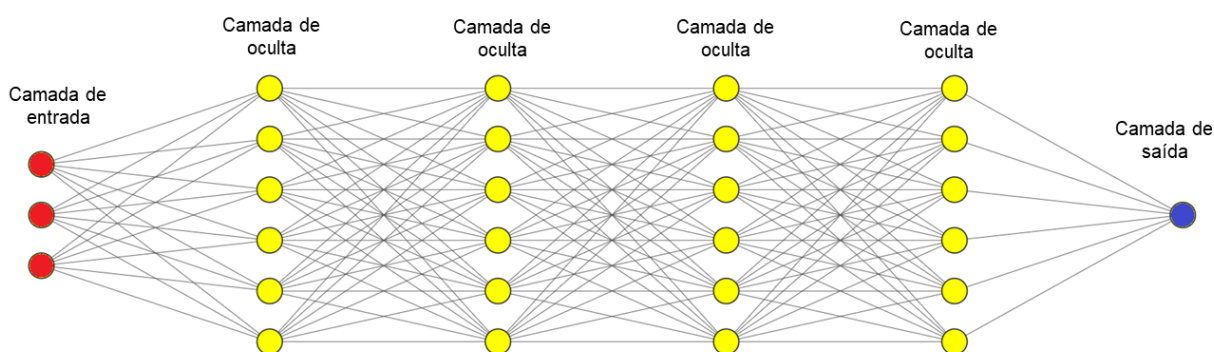
O *Deep Learning* obtém resultados consideráveis em problemas relacionados a Inteligência Computacional, pelo fato de ter uma alta capacidade de análise de padrões, podendo ser aplicado em diferentes áreas, por exemplo, ciência, esporte e finanças (LECUN; BENGIO; HINTON, 2015).

O aumento do volume de dados inviabiliza a utilização da programação com regras pré-definidas. Em situações desse tipo, *Deep Learning* pode ser mais eficaz. Grandes empresas têm necessidade de executar tarefas, como análise, organização e tratamento de suas bases de dados e *Deep Learning* pode auxiliar nessas atividades (KAUFMAN, 2019).

Deep Learning adota uma arquitetura com várias camadas intermediárias de processamento e abstração. Cada camada pode aprender a reconhecer um tipo de padrão. Por exemplo, se estiver realizando o reconhecimento de padrões visuais, a primeira camada pode reconhecer bordas, a segunda reconhecer formas mais complexas, como quadrados e retângulos, a terceira abstrair outros tipos de formas ainda mais complexas, e assim por diante (DATA SCIENCE ACADEMY, 2021).

Na Figura 3 é apresentado um exemplo de uma Rede Neural Profunda com seis camadas. De vermelho está representada a camada de entrada, que possui três neurônios. De amarelo estão descritas quatro camadas ocultas, com seis neurônios cada uma e, de azul, a camada de saída com um neurônio.

Figura 3 – Modelo de Rede Neural Profunda.



Fonte: Elaborada pelo autor.

## 2.4 Redes Neurais Convolucionais

As Redes Neurais Convolucionais (*Convolutional Neural Networks* - CNN) são Redes Neurais Profundas que podem classificar imagens, agrupá-las conforme padrões recorrentes e realizar reconhecimento de objetos. Essa arquitetura de Rede Neural Artificial pode identificar rostos, indivíduos, objetos e padrões visuais de maneira melhor adaptada (DATA SCIENCE ACADEMY, 2021).

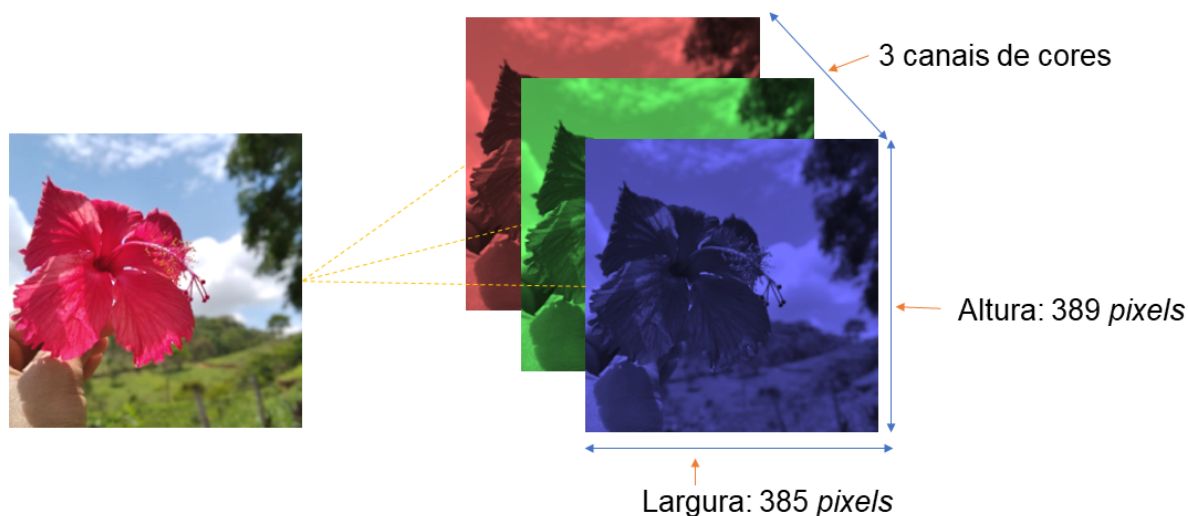
A primeira CNN foi criada em 1998 com sete camadas convolucionais e uma camada totalmente conectada (*fully connected*). Durante esse tempo até os dias atuais elas evoluíram muito, se tornando mais profundas (mais camadas) e complexas (ALVES, 2018).

Um problema estudado recorrentemente na Visão Computacional é o Reconhecimento de Objetos em imagens, e as CNNs têm mostrado bom desempenho nessas aplicações. A principal ideia de uma CNN é filtrar linhas, curvas e bordas de uma imagem, repetindo esses processos várias vezes, com o intuito de extrair informações mais complexas (ALVES, 2018).

Uma inovação da CNN é a habilidade de criar filtros capazes de extrair os atributos mais relevantes na diferenciação entre as classes, durante o treinamento. As CNNs se destacam por gerarem ótimos resultados na detecção de recursos, como cores, bordas e linhas em uma imagem, tendo alta capacidade de reconhecimento de padrões (ALIGER, 2019).

Em se tratando de classificação de imagens, a entrada de uma CNN é uma matriz tridimensional com altura, largura e profundidade (determinada pelo número de canais de cores) contendo os *pixels* que compõe a imagem. Na Figura 4 é apresentada uma entrada de uma CNN com altura de 389 *pixels*, largura de 385 *pixels* e com três canais (vermelho, verde e azul) (ALVES, 2018).

Figura 4 – Entrada tridimensional de uma CNN.

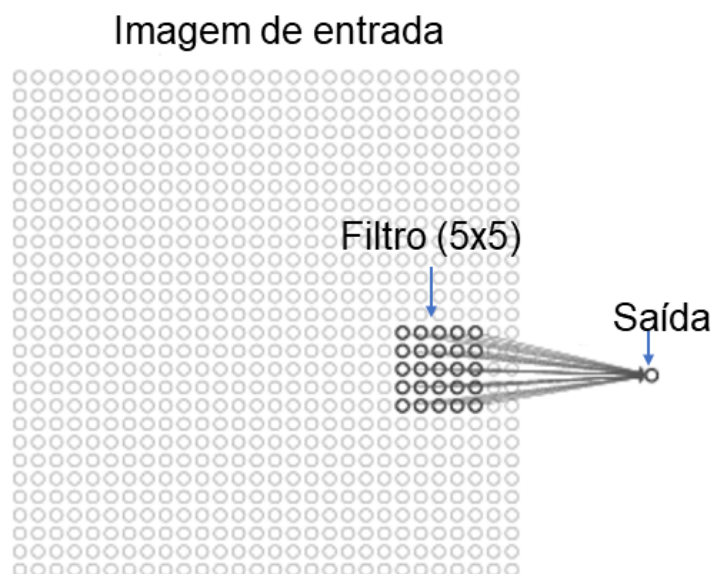


Fonte: Elaborada pelo autor.

Uma CNN é formada por vários tipos de camadas, de maneira que cada uma tem uma função. Na camada convolucional realiza-se as operações de convolução, definidas como a aplicação de um filtro capaz de produzir uma nova imagem com apenas seus detalhes realçados, de modo a permitir a classificação entre diferentes objetos. Pode-se pensar nas convoluções como filtros que percorrem toda a imagem, realçando traços, bordas e linhas horizontais, por exemplo, produzindo imagens que abordam pontos ou características específicas da imagem original (ALVES, 2018).

Na Figura 5, tem-se a representação do processo de convolução em uma imagem de 28x28 *pixels*, utilizando um filtro de 5x5 *pixels*. Esse filtro percorre toda a imagem, gerando como saída um *pixel* de uma nova imagem.

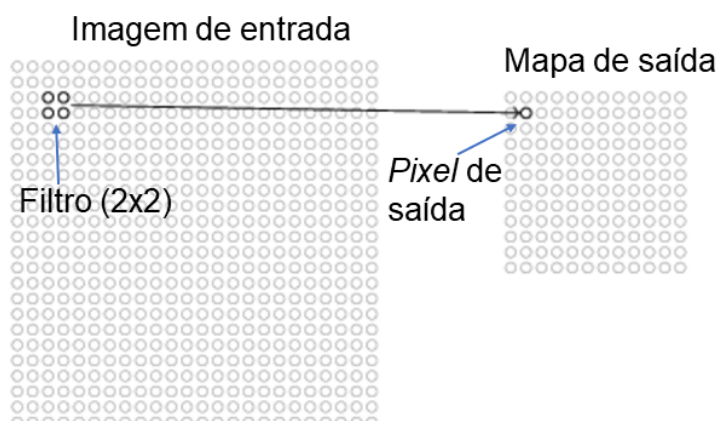
Figura 5 – Camada de convolução.



Fonte: Adaptado de *Data Science Academy*, 2019.

Nas camadas convolucionais são gerados os mapas de recursos (também conhecidos como mapa de características). Esses são criados ao se aplicar um filtro em uma imagem. Nesse momento, os filtros absorvem características dessa imagem e as guardam em uma matriz, gerando assim um mapa de recursos da imagem. Podem ser criados vários mapas de recursos para uma mesma imagem, conforme a situação.

As CNNs também possuem as camadas de agrupamento ou de *pooling*, que têm como função condensar ou diminuir o tamanho das imagens. O tamanho da saída de uma camada de *pooling* é influenciado pelo filtro utilizado. Na Figura 6 tem-se uma imagem na qual se aplica um filtro de tamanho 2x2 *pixels*. Esse filtro percorre toda a imagem, gerando como saída um conjunto de dados de menor dimensão espacial. O tamanho da saída é influenciado pelo tamanho do filtro (DATA SCIENCE ACADEMY, 2021).

Figura 6 – Processo de *Pooling*.

Fonte: Adaptado de *Data Science Academy*, 2019.

Apesar de existirem diferentes estratégias para definir o resultado gerado pelo filtro de *pooling*, a mais utilizada é a *Max-pooling*. Essa técnica consiste em selecionar o maior valor dentro de um conjunto de *pixels* ou simplesmente gerar a ativação máxima da região (DATA SCIENCE ACADEMY, 2021).

Na Figura 7 tem-se a representação de uma imagem de tamanho 4x4 *pixels*, na qual se aplica a operação de *Max-pooling*. Na primeira área de atuação do *Max-pooling* tem-se os valores 4, 2, 8, 6, sendo 8 o maior valor e, portanto, este será o valor selecionado. O procedimento é realizado em todo o conjunto de dados, de forma que *Max-pooling* percorra toda a imagem, gerando no final uma imagem de tamanho 2x2 *pixels*.

Figura 7 – Exemplo de Max-Pooling.



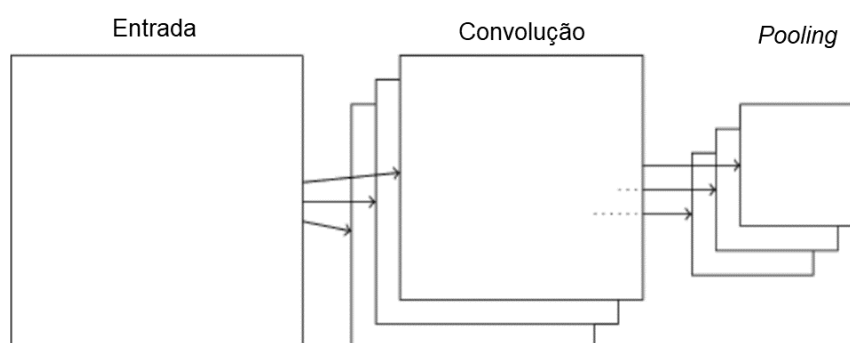
Fonte: PELTARION, 2021.

Na camada totalmente conectada, todos os neurônios são conectados entre si, ou seja, nessa camada todos os neurônios da própria camada são conectados à camada anterior e posterior, desempenhando a função de classificação dos dados. Devido a esse fato, possui a desvantagem de necessitar de maior processamento (OLIVEIRA JUNIOR, 2018).

Para que a rede consiga aprender adequadamente, existem as funções de ativação que são aplicadas em neurônios de diferentes camadas, conforme a situação. Apesar das funções *sigmoide*, *tanh* e *softmax* serem exemplos de funções de ativação, para as redes convolucionais, a mais indicada, segundo (ALVES, 2018), é a função de ativação *ReLU* por ser mais eficiente computacionalmente, sem grandes diferenças de Acurácia.

Para reconhecer uma imagem, são necessários alguns mapas de imagem para reconhecimento de padrões. Na Figura 8 é apresentada uma imagem de entrada, gerando três mapas de imagens na camada convolucional. Na camada seguinte (camada de *pooling*), as saídas da camada anterior foram comprimidas, gerando saídas menores (DATA SCIENCE ACADEMY, 2021).

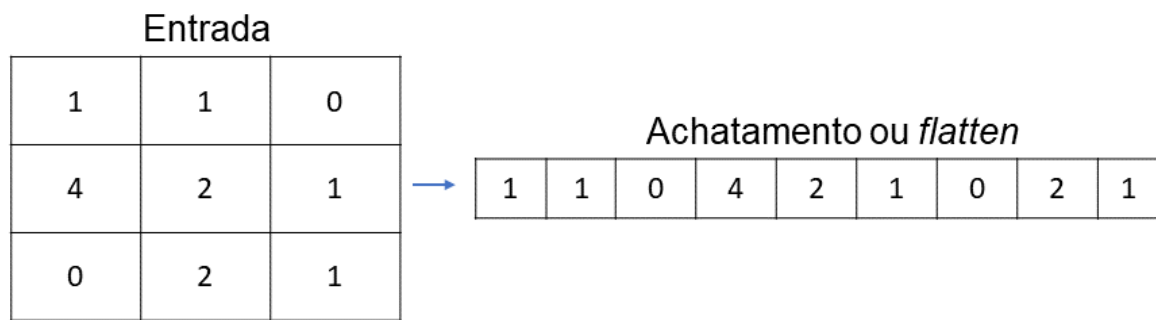
Figura 8 – Exemplo de camadas de uma rede convolucional.



Fonte: Adaptado de *Data Science Academy*, 2019.

Além dos tipos de camadas já abordados, também existem algumas operações para auxiliar na formatação dos dados de saída de uma camada para outra. Neste trabalho utilizou-se a operação de achatamento conhecida como *flattening*. A operação de achatamento tem a finalidade de transformar a saída matricial em um saída unidimensional, como é demonstrado na Figura 9.



Figura 9 – Operação de achatamento ou *flatten*.

Fonte: Elaborado pelo autor.

## 3 Trabalhos Relacionados

Um passo fundamental em qualquer trabalho é a leitura e análise de outros trabalhos, pois dessa forma é possível verificar o que já foi feito, como foi desenvolvido e qual a relação desses conteúdos com o seu trabalho. Neste capítulo serão abordados trabalhos que têm como foco principal o diagnóstico da COVID-19 por meio da classificação de imagens de radiografias do tórax, com destaque para a utilização de Redes Neurais Profundas para executar tal tarefa.

### 3.1 Classificação de COVID-19 com Rede Neural Totalmente Conectada

No trabalho de Carvalho et al. (2020), foi utilizada uma Rede Neural Totalmente Conectada (*Fully-Connected Neural Networks FCNN*), possuindo uma camada de entrada com 20 neurônios, seis camadas internas totalmente conectadas com 16 neurônios cada, e uma camada de saída com um neurônio, produzindo uma classificação binária. Para extrair características da imagem, utilizou-se o descritor de textura de *Haralick* e *Threshold Adjacency Statistics* (TAS) e, ao todo, foram extraídas 20 características. Na etapa de treinamento foram usadas 100 épocas. Utilizou-se uma base com 219 imagens de pacientes com COVID-19, sendo 1.345 de pneumonia viral e 1.341 saudáveis. Para os experimentos foram usadas apenas imagens de pacientes com COVID-19 e saudáveis. Os resultados obtidos foram avaliados mediante métricas de Acurácia de 98,39%, Sensibilidade (classificação para casos de verdadeiros positivos) de 99,18% e Especificidade (classificação para casos de verdadeiros negativos) de 93,47%.

Em Viveiros et al. (2020) também foi utilizada uma FCNN, diferenciando na camada de entrada e também nas camadas totalmente conectadas, que possuíam 268 neurônios e seis camadas totalmente conectadas com 180 neurônios cada, respectivamente. Como em Carvalho et al. (2020), a camada de saída possuía um neurônio logístico com saída binária. A FCNN foi treinada usando 5.000 épocas. As imagens foram redimensionadas para 512x512 *pixels*. Foi utilizada a segmentação de imagens, bem como um pré-processamento das imagens através da Equalização do Histograma, aplicando a U-Net desenvolvida em Ronneberger, Fischer e Brox (2015). Utilizando as bases de dados de Cohen et al. (2020) e Rahman, Chowdhury e Khandakar (2020), foram empregadas 525 imagens de radiografias de tórax, sendo 263 de pacientes com COVID-19 e 262 saudáveis. Obtiveram como resultado uma Acurácia de 95,48%, Sen-

sibilidade de 95,34% e Especificidade de 95,79%.

Já Martins et al. (2021) utilizaram uma FCNN com uma camada de entrada de 13 neurônios, seis camadas ocultas com 16 neurônios cada, junto a seis camadas de *Dropout* (processo de exclusão aleatória e temporária de alguns dos neurônios da camada oculta de uma rede) e uma camada de saída contendo um único neurônio. Para a extração de características das imagens, utilizou-se o Padrão Binário Local (*Local Binary Patterns* - LBP), desenvolvido por He e Wang (1990). Foram usadas as imagens de pacientes saudáveis e de pacientes com COVID-19, disponibilizadas em Rahman, Chowdhury e Khandakar (2020), composta por 219 imagens de pacientes com COVID-19, 1341 de pacientes saudáveis e 1345 de pacientes com pneumonia viral. Para evitar a prevalência de classe, a base foi balanceada. Os experimentos geraram valores médios de Acurácia de 96,52%, Sensibilidade de 94,00% e Especificidade de 98,64%.

## 3.2 Classificação de COVID-19 com Rede Neural Convolutacional

Abbas, Abdelsamea e Gaber (2021) usaram um tipo de CNN denominada Decompor, Transferir e Compor (*Decompose, Transfer, and Compose* - *DeTraC*), que é derivada de uma VGG19. A partir de um conjunto de imagens coletadas de Cohen et al. (2020), Jaeger et al. (2014) e Candemir et al. (2014), foram aplicadas técnicas de Aumento de Dados (*Data Augmentation*), criando múltiplas versões das figuras da base de dados, após a aplicação de transformações geométricas, como Translação e Rotação. Ao final, foram produzidas 1.674 amostras. Como resultados obtiveram uma Acurácia de 97,35%, Sensibilidade de 98,23% e Especificidade de 96,34%.

Narin, Kaya e Pamuk (2020) em sua pesquisa foram utilizados cinco modelos baseados em CNNs pré-treinadas (ResNet50, ResNet101, ResNet152, InceptionV3 e Inception-ResNetV2). Para os experimentos foram usadas 341 imagens de radiografias do tórax de um repositório do GitHub de código aberto (COHEN et al., 2020), 2.800 imagens de radiografias do tórax normais (saudáveis) (WANG et al., 2017), 2.772 imagens de radiografias do tórax de pneumonia viral e 1.493 de pneumonia bacteriana (MOONEY, 2018). Os modelos ResNet50 e ResNet101 obtiveram 96,10% de Acurácia, gerando os melhores resultados. Por ter uma taxa menor de erro, os autores consideraram o ResNet50 um pouco melhor.

Em seus experimentos, Ozturk et al. (2020) utilizaram um modelo de CNN denominado DarkNet, para desenvolver uma nova arquitetura chamada de DarkCovidNet. Seu modelo continha um número menor de camadas e filtros, em comparação com o modelo original. Para o desenvolvimento do respectivo trabalho, foram usadas 127 imagens de pacientes com COVID-19 de Cohen et al. (2020), 500 imagens de pacientes com pneumonia e 500 imagens de pacientes saudáveis de Wang et al. (2017). Os

testes realizados com os três tipos de imagens obtiveram uma Acurácia de 87,02%, Sensibilidade de 85,35% e Especificidade de 92,18%. Já os testes realizados com as imagens de pacientes com COVID-19 e pacientes saudáveis atingiram uma Acurácia de 98,08%, Sensibilidade de 95,13% e Especificidade de 95,30%.

Apostolopoulos e Mpesiana (2020) abordaram um método de aprendizagem por transferência. Para tal tarefa foram utilizadas CNNs pré-treinadas, adaptadas para uma nova tarefa a partir de um retreinamento do modelo. Os modelos utilizados foram VGG19, Mobile Net, Inception, Xception e Inception ResNet v2. Para a realização dos experimentos foram usadas 1.428 imagens de radiografias do tórax, sendo 224 imagens de pacientes com COVID-19, 700 imagens de pacientes com pneumonia comum e 504 imagens de pacientes em condições normais, extraídas de diversas fontes. Após o treinamento, o melhor resultado foi alcançado pela Mobile Net, com uma Acurácia de 96,78%, Sensibilidade de 98,66% e Especificidade de 96,46%.

Zhang et al. (2021) propuseram a utilização de um modelo de CNN convencional. Esse modelo utiliza agrupamento estocástico (amostragem de uma distribuição multinomial gerada por uma região de ativação) nas camadas de *pooling*, combinação das camadas convolucionais com camadas de normalização e união, camadas *dropout* e, finalmente, camadas *Fully-Connected*, gerando melhora no desempenho do modelo. Utilizando uma base com 142 imagens de pacientes com COVID-19 e 142 imagens de pacientes saudáveis, obtidas de hospitais locais, o modelo atingiu uma Acurácia de 93,64% (com variação 1,42% positiva ou negativa), uma Especificidade de 94,00% (com variação 1,56% positiva ou negativa) e Sensibilidade de 93,28% (com variação 1,50% positiva ou negativa) no diagnóstico da COVID-19.

Islam, Islam e Asraf (2020) utilizaram uma CNN e uma arquitetura de Rede Neural Artificial recorrente que se recorda de valores em intervalos arbitrários (*Long Short-Term Memory* LSTM). No modelo proposto, a CNN foi usada para extração de recursos e a LSTM foi usada para classificar os dados com base nos recursos extraídos. Seu modelo é formado por 12 camadas convolucionais, cinco camadas de *pooling*, uma camada totalmente conectada, uma camada LSTM e, por fim, uma camada de saída com a função *softmax*, totalizando 20 camadas. Usando uma base de dados de 4.575 imagens de radiografias do tórax dividida em 1.525 imagens de COVID-19, 1.525 imagens de pneumonia e 1.525 imagens de pacientes normais, composta por um compilado de outras bases. Em seus experimentos foram produzidos resultados médios com Acurácia de 99,4%, AUC de 99,9%, Especificidade de 99,2% e Sensibilidade de 99,3%.

### 3.3 Classificação de COVID-19 com Outras Abordagens

Özkaya, Öztürk e Barstugan (2020) propuseram a utilização de uma Máquina de Vetor de Suporte (*Support Vector Machine* - SVM). Foram obtidos vetores de características de três CNNs pré-treinadas (VGG-16, GoogleNet e ResNet-50), com o intuito de treinar a SVM. Em seus experimentos foram usadas 53 imagens de tomografia computadorizada, nas quais foram extraídos fragmentos de tamanho 16x16 e 32x32 *pixels*, a partir da base de dados da *Societa Italiana di Radiologia Medica e Interventistica* (2020). A SVM foi comparada à VGG-16, GoogleNet e ResNet-50, obtendo resultados superiores nos testes, alcançando resultados de 98,27% de Acurácia, 98,93% de Sensibilidade e 97,60% de Especificidade.

No trabalho desenvolvido por Heidarian et al. (2020) foi proposta a *COVID-CAPS* - um tipo de Rede Neural Artificial denominada Rede Encapsulada (*Capsule Networks*). Essas redes possuem capacidade de capturar informações espaciais (informações de um objeto e suas propriedades, como localização no espaço, sua relação com outros objetos, cor, tamanho, contraste etc.). Utilizando um conjunto de imagens de tomografias volumétricas de tórax (exame de radiografias utilizadas para a visualização interna no corpo de forma tridimensional), com 171 imagens de pacientes com COVID-19, 60 pacientes com pneumonia e 76 saudáveis de Afshar et al. (2021), atingiram nos experimentos uma Acurácia de 90,82%, Sensibilidade de 94,55% e Especificidade de 86,04%.

### 3.4 Conclusão Sobre os Trabalhos Relacionados

Ao analisar os trabalhos relacionados, percebeu-se um grande potencial de conhecimentos de padrões das CNNs, gerando resultados de Acurácia superiores a 90%.

Dentre os trabalhos que usaram FNCCs, considerando que mesmos adotaram parcialmente ou totalmente a mesma base de dados, possibilitando uma comparação de resultados, em Carvalho et al. (2020) foi obtida a melhor Acurácia, com 98,39%.

Se destacou a utilização das CNNs em conjunto com outros métodos, como a união de uma CNN a uma LSTM, em Islam, Islam e Asraf (2020), aplicação de métodos de translação e rotação de imagem a fim de gerar uma base maior, em Abbas, Abdelsamea e Gaber (2021), e aprendizagem por transferência, em Apostolopoulos e Mpesiana (2020).

Outro fator relevante nos trabalhos foi a utilização de modelos de CNNs já populares na tarefa de classificação de imagens, por exemplo, VGG-16, VGG-19 e ResNet-50, como a comparação de eficiência apresentada em Narin, Kaya e Pamuk (2020).

## 4 Materiais e métodos

Neste capítulo são abordados os materiais e métodos utilizados neste trabalho. Ele está dividido nos seguintes tópicos: Pré-processamento da Base de Dados, Arquitetura de Rede Neural Artificial, Protocolo de Experimentação e Métricas Utilizadas.

Utilizou-se três modelos de rede neurais diferentes, cada um com números diferentes de camadas, com a finalidade de analisar o desempenho desses modelos para a classificação de imagens de radiografia do tórax.

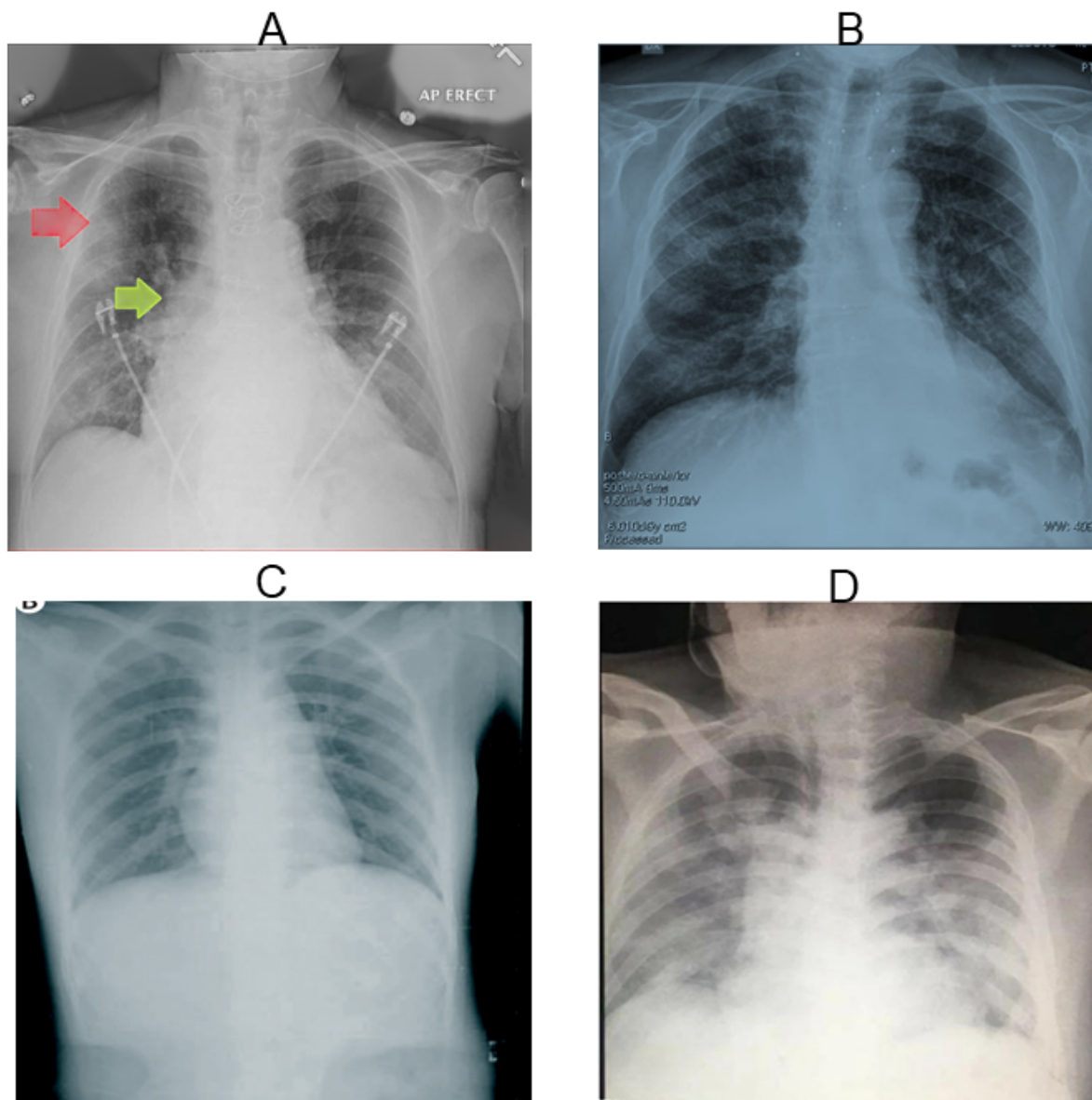
### 4.1 Pré-Processamento da Base de Dados

Dada a natureza deste trabalho, faz-se necessária a utilização de uma base de dados para fundamentar o aprendizado da Rede Neural Convolutacional proposta. Para tal, foi usada a base disponibilizada em Patel (2021), que é uma junção de outras bases de dados, totalizando 6.432 imagens de tamanhos variados, com 1.583 imagens de pacientes saudáveis, 576 imagens de pacientes diagnosticados com COVID-19 e 4.273 imagens de pacientes com pneumonia.

Para a formação do conjunto de dados, Patel (2021) coletou as imagens de três repositórios distintos: Cohen et al. (2020), com licença para pesquisas científicas, Mooney (2018), também licenciada e, por fim, ambos do mesmo autor, Chung (2020a) e Chung (2020b), que apesar de não terem encontrado licenças vinculadas a tais bases, o autor menciona que elas devem ser usadas para pesquisas científicas.

Na Figura 10, tem-se quatro imagens de exemplo de pacientes com COVID-19 retiradas da base de dados, escolhidas para ilustrar algumas nuances entre as imagens que a compõe. Em Figura 10(A), pode-se observar setas coloridas na imagem, bem como linhas na região dos pulmões. Em Figura 10(B) é mostrado um texto na parte inferior da imagem, além da imagem ter uma tonalidade mais azulada. Em Figura 10(C) destaca-se uma região mais alongada abaixo dos pulmões e, por fim, em Figura 10(D) é exibida uma imagem com coloração amarelada. Apesar das CNNs terem grande capacidade de classificação, as observações salientadas nas imagens podem prejudicar a sua classificação e, conseqüentemente, diminuir a Acurácia. Assim, para se evitar o reconhecimento de padrões irrelevantes para a classificação optou-se por realizar um pré-processamento da base de dados.

Figura 10 – Exemplo de Imagens de Radiografias do Tórax da Base de Dados.



Fonte: Adaptado de Patel, 2021.

No processo de pré-processamento, considerando que as cores das imagens não seriam uma característica adequada para a classificação das imagens, tanto pelo fato de aumentar a complexidade computacional quanto pela falta de homogeneidade entre as cores, todas as imagens da base foram convertidas para escala de cinza. Apesar da base ser composta em sua maioria por imagens nesse padrão, algumas não estavam representadas dessa maneira.

Utilizou-se a biblioteca OpenCV para a conversão das imagens em tons de cinza. Essa biblioteca é amplamente utilizada para processamento de imagens e vídeos, usufruindo de grande destaque na detecção de objetos, aplicação de filtros, detecção de bordas, filtragem de canais de cores etc.

As imagens também foram redimensionadas para o tamanho de 1024x1024 *pixels*, para uma padronização do seu tamanho. No entanto, verificou-se que esse processo não era necessário, já que a CNN redimensiona as imagens em sua camada de entrada.

Como mencionado por Monard e Baranauskas (2003), o balanceamento da base de dados é fundamental para gerar um melhor desempenho no treinamento da base, além de ajudar na distribuição das classes de forma mais homogênea.

Com a finalidade de balancear o conjunto de dados usado, foram selecionadas 500 imagens de pacientes saudáveis (400 para treino e 100 para teste) e 500 imagens de pacientes com COVID-19 (400 para treino e 100 para teste). O processo de seleção foi feito randomicamente, gerando 1000 imagens para o treinamento do modelo de CNN.

## 4.2 Arquitetura de Rede Neural Artificial

Neste trabalho foi utilizada uma CNN para a classificação da base de imagens, pois as CNNs possuem um bom desempenho na tarefa de classificação de imagens, podendo detectar linhas, bordas, objetos etc. Além disso, CNNs vêm sendo amplamente usadas nos últimos anos para a resolução de problemas de Visão Computacional, se destacando na análise de padrões e classificação de imagens.

Foram propostos três modelos de Rede Neural Artificial, que foram intitulados modelo I, modelo II e modelo III. Os modelos são compostos por camadas de convolução, *Max-pooling* e *Dense*. A seguir cada modelo será abordado detalhadamente.

### 4.2.1 Modelo I

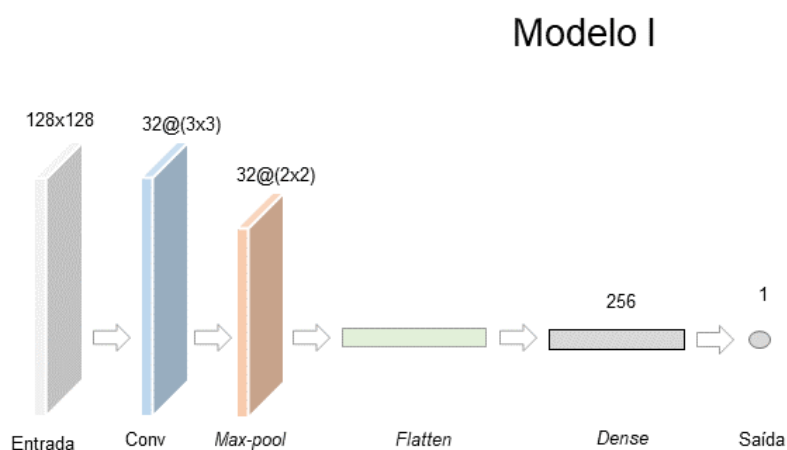
O modelo I tem um número menor de camadas, sendo composto por quatro camadas. Possui como entrada uma imagem, que é redimensionada para o de tamanho 128x128 *pixels*, sendo cada *pixel* interpretado por um neurônio. É seguida por uma camada convolucional com 32 filtros de tamanho 3x3 *pixels* e função de ativação ReLU. A seguir, inclui-se uma camada *Max-pooling* de tamanho 2x2 *pixels*. Posteriormente, aplica-se a operação de achatamento (*flatten*) e utiliza-se uma camada totalmente conectada (*dense*), com 256 neurônios e a inclui-se uma taxa de 20% *dropout*, junto à função de ativação ReLU. Finalmente, introduz-se um neurônio de saída para a classificação binária, usando uma função de ativação *Sigmoid*.

Na Figura 11 encontra-se a representação gráfica do modelo I. A entrada da CNN é representada por um plano retangular de cor cinza, sendo o tamanho da imagem de entrada, 128x128 *pixels* descrito na parte superior da figura e na parte inferior



encontra-se a legenda figura. As camadas convolucionais são representadas por uma figura retangular de cor azul claro, sendo em sua posição inferior incluída uma legenda Conv e na parte superior incluído o número de filtros junto a sua dimensão (representado por: *filtro@dimensões*). Já as camadas *Max-pooling* são representadas por retângulos da cor laranja, em sua posição inferior tem-se uma legenda com o conteúdo Max-pool e na parte superior esta incluído o número de filtros junto a sua dimensão (representado por: *filtro@dimensões*). A operação de achatamento ou *flatten* é representada por um retângulo bidimensional de cor verde claro, com a legenda *flatten*. A camada totalmente conectada ou *dense* é representada por um retângulo bidimensional cinza com a legenda *Dense* abaixo e na parte superior o número 256 que indica o número de neurônios. A saída é representada por um círculo, também de cor cinza, na parte inferior, com a palavra *Saída*, na parte superior, o número 1 que indica o número de neurônios.

Figura 11 – Representação da arquitetura do modelo I.



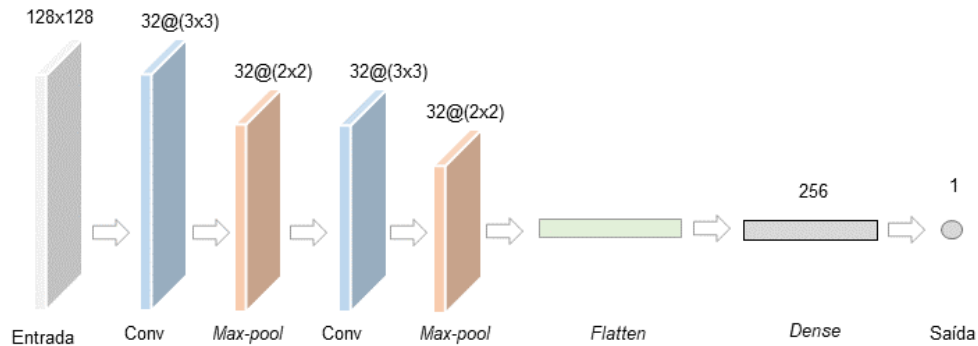
Fonte: Elaborada pelo autor.

#### 4.2.2 Modelo II

O modelo II tem a mesma arquitetura do modelo I, porém foram adicionadas uma camada de convolução com 32 filtros de tamanho 3x3 *pixels*, aplicando uma função de ativação ReLU, seguida por uma camada *Max-pooling* de tamanho 2x2 *pixels*, da operação de *flattening*. Nesse caso, o modelo II tem seis camadas em sua arquitetura. Na Figura 12 é exibida a representação gráfica do modelo II, no qual as camadas são ilustradas da mesma forma que na Figura 11.

Figura 12 – Representação da arquitetura do modelo II.

## Modelo II



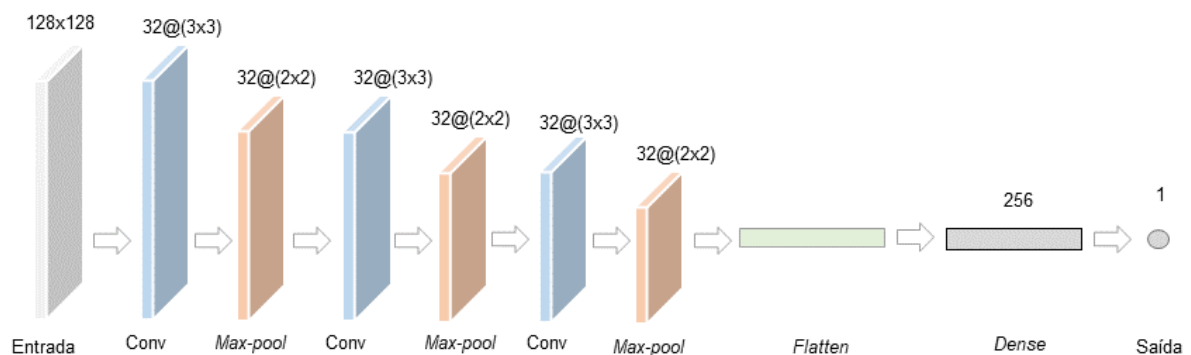
Fonte: Elaborada pelo autor.

## 4.2.3 Modelo III

O modelo III, possui o maior número de camadas, totalizando oito. Esse modelo possui uma arquitetura muito parecida com os modelos anteriores. Foram introduzidas duas camadas a mais em relação ao modelo II, sendo uma camada de convolução, com 32 filtros de tamanho 3x3 *pixels*, função de ativação ReLU e uma camada *Max-pooling* de tamanho 2x2 *pixels*, antes de uma camada totalmente conectada de 256 neurônios. Na Figura 13 exibe-se a representação gráfica do modelo III. Pode-se observar que as camadas são representadas similarmente às Figuras 11 e 12.

Figura 13 – Representação da arquitetura do modelo III.

## Modelo III



Fonte: Elaborada pelo autor.

A metodologia apresentada foi adotada com o objetivo de analisar quais os impactos de se utilizar modelos de Redes Neurais Profundas com um número de ca-

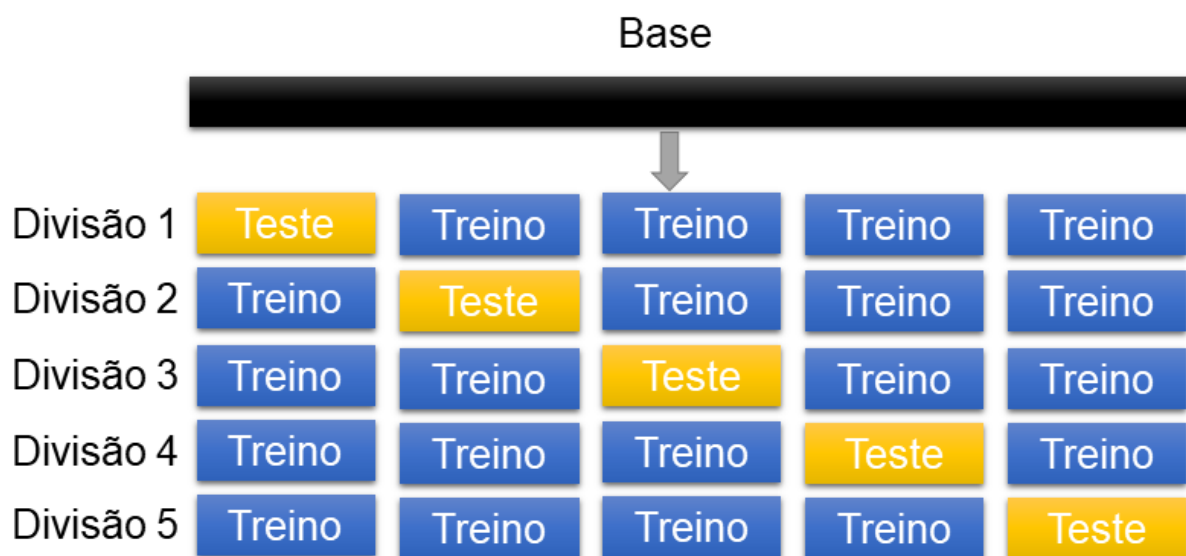
madras menor do que as arquiteturas convencionais, que são usadas para a resolução de trabalhos semelhantes.

### 4.3 Protocolo de Experimentação

Uma parte essencial em qualquer trabalho é o processo de validação dos seus experimentos. O protocolo de validação cruzada *k-fold* é amplamente utilizado para problemas de classificação, tendo a finalidade de realizar avaliações de maneira a observar a capacidade de generalização do modelo.

A validação cruzada *k-fold* divide a base de dados em  $k$  partes. São definidas  $k - 1$  partes para treino e uma parte para validação, contemplando  $k$  números de interação, de maneira que seja uma parte diferente para teste em cada interação. Na Figura 14 tem-se uma exemplificação do *k-fold*.

Figura 14 – Representação visual do *k-fold* para  $k = 5$ .



Fonte: Elaborado pelo autor.

### 4.4 Métricas utilizadas

É fundamental avaliar os resultados de qualquer experimento de modo, a validar uma hipótese e, para isso, são utilizadas as métricas de avaliação. Entretanto, antes de abordar essas métricas deve-se saber alguns conceitos. Em todo experimento binário são gerados quatro valores: o total de verdadeiros positivos (VP), o valor de verdadeiros negativos (VN), o total de resultados positivos rotulados de forma errônea à classe positiva (paciente com COVID-19) e, por fim, o número de resultados negativos rotula-

dos de forma errônea ao conjunto negativo (paciente sem COVID-19). Essas variáveis são fundamentais para o cálculo das métricas de avaliação.

Nos experimentos foram consideradas quatro métricas. A Acurácia (ACC), que é a probabilidade do teste fornecer resultados corretos, ou seja, das classificações positivas e negativas estarem de fato corretas. Esse cálculo é realizado conforme previsto na Equação 4.1.

$$ACC = \frac{(VP + VN)}{(VP + VN + FP + FN)} \quad (4.1)$$

A Sensibilidade (SEN), por sua vez, é a probabilidade do teste fornecer resultados verdadeiros corretos, considerando apenas os resultados positivos. Essa métrica pode ser calculada considerando a Equação 4.2.

$$SEN = \frac{(VP)}{(VP + FN)} \quad (4.2)$$

A Especificidade (ESP) descreve o grau de acerto do método apenas para a classe negativa e ela é calculada conforme a Equação 4.3.

$$ESP = \frac{(VN)}{(VN + FP)} \quad (4.3)$$

A Curva Característica de Operação do Receptor (*Area under the Curve* - AUC) é uma representação gráfica que mede o quão bem o modelo separa as classes, considerando diferentes limiares. O cálculo da AUC, calculada conforme a Equação 4.4, é baseado no cálculo da área da curva de acerto e, por isso, se utiliza de uma integral, tendo como base a Taxa de Verdadeiros Positivos (TPR) e a Taxa de Falsos Positivos (FPR) (PAULINO, 2018).

$$AUC = \int_0^1 TPR(FPR) dFPR \quad (4.4)$$

A escolha dessas métricas foi influenciada pelo fato de serem amplamente utilizadas em trabalhos de classificação de imagens, sendo bons indicativos de como avaliar devidamente a evolução do treinamento do modelo em questão.

## 5 Experimentos e Resultados

Neste capítulo são abordados os protocolos de experimentação empregados neste trabalho, bem como o equipamento e aplicativos computacionais utilizados para realizar os experimentos, além dos resultados gerados pelos mesmos.

### 5.1 Especificações de *Hardware* e *Software*

Os testes foram realizados em um *notebook* Acer Nitro AN515-52, processador Intel(R) Core(TM) i5-8300H de 2.30GHz, com 16GB de memória RAM e uma placa de video NVIDIA GeForce GTX 1050 de 4GBs dedicados, com um HD SSD tipo NVMe de 512GB. O sistema operacional usado foi o Windows 10 *Home Single Language*, baseado em uma arquitetura de 64 *bits*.

Para a implementação utilizou-se a linguagem de programação *Python* 3.8.10, por ser uma tecnologia composta por diferentes bibliotecas amplamente utilizadas para resolução de problemas na área de Visão Computacional. Junto à linguagem supracitada, também foram utilizadas as seguintes bibliotecas: *Keras* 2.4.3, ferramenta que faz a interface *front-end* com bibliotecas mais robustas de *Deep Learning*; *TensorFlow* 2.5.0, muito empregada para a construção de arquiteturas de Redes Neurais Profundas com suporte para Unidade de Processamento Gráfico (*Graphics Processing Unit* - GPU); *OpenCV* 4.5.2.54, usada para a aplicação de algoritmos de Processamento de Imagens e Visão Computacional, e também *Matplotlib* 3.4.2, solução robusta para a construção de gráficos para análise.

Para auxiliar o treinamento via GPU, utilizou-se a tecnologia CUDA (*Compute Unified Device Architecture*) versão 11.3, desenvolvida pela Nvidia para possibilitar a utilização de suas GPUs na aceleração das aplicações.

### 5.2 Resultados do Protocolo de Avaliação Utilizado

Os experimentos foram conduzidos seguindo o protocolo de Validação Cruzada *K-fold*, pois é amplamente utilizado em problemas de natureza classificatória. Utilizou-se 80% da base para treinamento e 20% da base para teste, em cinco iterações, ou seja,  $k = 5$ .

A seguir, apresenta-se a Tabela 1, em que estão contidos os resultados do treinamento utilizando o método *K-fold*. Na primeira coluna ( $K$ ) são descritos cada ciclo (*fold*) durante a etapa de treinamento. Conjuntamente, são apresentados os valores

de Acurácia, Sensibilidade, Especificidade e AUC para cada iteração. Esses valores são apresentados em porcentagem. Na última linha são descritos os valores médios de cada métrica. O valor da Média (em porcentagem) foi calculado somando-se os valores produzidos após a aplicação das métricas durante cada ciclo  $k$ , dividido pelo total de ciclos (no caso 5).

Tabela 1 – Resultados do  $k$ -fold, para o modelo I.

K	Acurácia (%)	Sensibilidade (%)	Especificidade (%)	AUC (%)
1	98,89	98,61	99,17	99,81
2	98,89	98,89	98,89	99,96
3	99,03	98,89	99,17	99,81
4	99,17	99,44	98,89	99,95
5	98,47	98,33	98,61	99,91
<b>Média</b>	<b>98,89</b>	<b>98,83</b>	<b>98,94</b>	<b>99,89</b>

Fonte: Elaborado pelo autor.

Na Tabela 2 são descritos os resultados do treinamento do modelo II, utilizando o método de Validação Cruzada  $K$ -fold. A Tabela 2 segue a mesma estrutura da Tabela 1.

Tabela 2 – Resultados do  $k$ -fold, para o modelo II.

K	Acurácia (%)	Sensibilidade (%)	Especificidade (%)	AUC (%)
1	99,58	99,44	99,72	100,00
2	99,44	99,44	99,44	99,99
3	99,58	99,72	99,44	99,98
4	99,03	98,89	99,17	99,99
5	98,33	98,33	98,33	99,81
<b>Média</b>	<b>99,19</b>	<b>99,17</b>	<b>99,22</b>	<b>99,95</b>

Fonte: Elaborado pelo autor.

Finalmente, na Tabela 3 são apresentados os resultados do treinamento do modelo III, empregando a mesma metodologia de experimentação usada nos modelos anteriores. A tabela segue o mesmo formato das anteriores.

Tabela 3 – Resultados do *k-fold*, para o modelo III.

K	Acurácia (%)	Sensibilidade (%)	Especificidade (%)	AUC (%)
1	99,17	99,17	99,17	99,82
2	99,31	98,89	99,72	99,98
3	99,72	99,72	99,72	99,84
4	99,58	99,44	99,72	100,0
5	98,75	98,61	98,89	99,86
<b>Média</b>	<b>99,31</b>	<b>99,17</b>	<b>99,44</b>	<b>99,90</b>

Fonte: Elaborado pelo autor.

### 5.3 Resultados e Discussão

Para avaliar a performance dos modelos apresentados nos experimentos, foram analisadas as métricas de Acurácia, Sensibilidade, Especificidade e AUC. Na Tabela 4, são apresentadas as médias dos resultados dos experimentos.

Tabela 4 – Resultados dos experimentos.

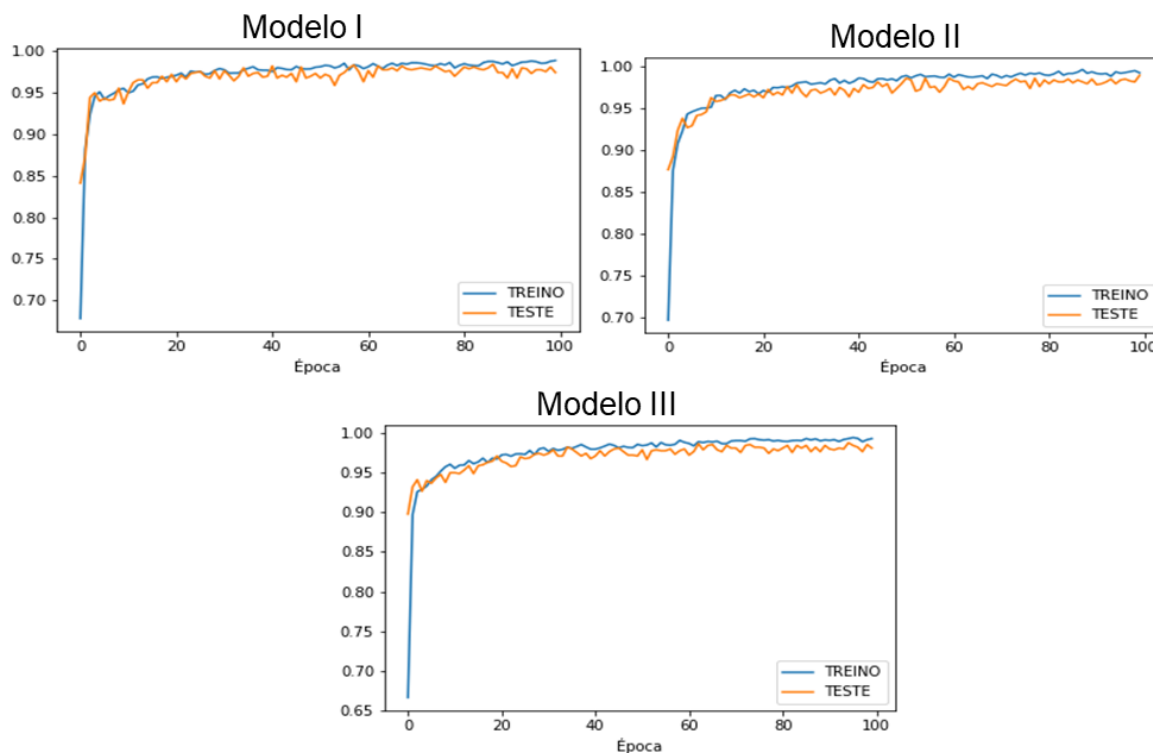
Modelo	Acurácia (%)	Sensibilidade (%)	Especificidade (%)	AUC (%)
Modelo I	98,89	98,83	98,94	99,89
Modelo II	99,19	99,17	99,22	99,95
Modelo III	<b>99,31</b>	99,17	99,44	99,90

Fonte: Elaborado pelo autor.

Analisando os resultados, pode-se perceber que o modelo I teve o menor desempenho, com uma Acurácia média de 98,89%. O modelo II gerou resultados melhores que o modelo I, tendo uma Acurácia média de 99,19%. Já o modelo III obteve o melhor desempenho, com uma Acurácia média de 99,31%. Conforme pode-se observar, modelos que possuem um número maior de camadas produzem resultados melhores.

Na Figura 15 tem-se três gráficos, gerados pelos valores médios de Acurácia obtidos pelo *k-fold*, representados em uma mesma imagem para melhor visualização e comparação. No gráfico intitulado Modelo I encontra-se o gráfico da Acurácia média do modelo I. Os gráficos Modelo II e Modelo III seguem a mesma linha. Os números no eixo *y* representam o nível de Acurácia de maneira normalizada (0 representando 0% e 1 representando 100%). Os valores do eixo *x* representam o total de épocas de treinamento, variando de 0 a 100 épocas. As linhas azuis representam a Acurácia média no treinamento de seu modelo, já as linhas laranjas descrevem a mesma métrica no que se refere ao teste do modelo.

Figura 15 – Acurácia média.



Fonte: Elaborado pelo autor.

É perceptível uma semelhança entre os gráficos, já que ambos os modelos iniciam o treinamento com um rápido aumento na Acurácia e, após algumas épocas, começam a se estabilizar. Analisando o Modelo I, após essa estabilização, observa-se a ocorrência de um movimento instável em seu treinamento, entre as épocas 40 e 60. No modelo II também é possível ver esse movimento.



## 6 Conclusão

Atualmente vive-se em um momento histórico conturbado, ocasionado por uma doença que se tornou uma pandemia com taxas de letalidade elevadas, bem como causadora de um alto número de sequelas. Por esse motivo, é fundamental a produção de conhecimento científico, com a finalidade de gerar soluções para auxiliar e minimizar essa situação.

Neste trabalho produziu-se uma arquitetura de Rede Neural Convolucional para o reconhecimento de padrões em imagens de pacientes que possuíam a COVID-19 e pacientes saudáveis. Para tal, foram analisados diferentes modelos de CNN, com objetivo de verificar o desempenho com diferentes números de camadas.

A partir dos resultados gerados, percebeu-se que modelos com maior número de camadas têm uma propensão a gerar resultados melhores. Em teoria, um maior número de camadas gera uma maior absorção de padrões e, conseqüentemente, de identificação. No entanto, modelos com um número inferior de camadas já podem ser suficientes para se obter bons resultados. Tudo isso exemplifica o poder de identificação tipicamente conseguido pelas CNNs.

A partir dos resultados produzidos, concluiu-se que o modelo III apresentou o melhor resultado entre os modelos analisados, podendo ser uma alternativa para aprimorar a classificação de imagens de radiografias torácicas, permitindo detectar lesões que os radiologistas podem ter dificuldades para perceber, auxiliando em um diagnóstico mais preciso.

Como trabalho futuro, pode-se desenvolver um sistema que utilize a CNN já treinada para a classificação de imagens, com o objetivo de aplicar, de forma prática, o conhecimento adquirido neste trabalho.

## Referências

- ABBAS, A.; ABDELSAMEA, M. M.; GABER, M. M. **Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network**. *Applied Intelligence*, Springer, v. 51, n. 2, p. 854–864, 2021.
- AFSHAR, P. et al. **COVID-CT-MD, COVID-19 computed tomography scan dataset applicable in machine learning and deep learning**. *Scientific Data*, Nature Publishing Group, v. 8, n. 1, p. 1–8, 2021.
- ALECRIM, E. **Machine learning: o que é e por que é tão importante**. 2018. Disponível em: <<https://tecnoblog.net/247820/machine-learning-ia-o-que-e>>. Acesso em: 21 fev. 2021.
- ALIGER. **As Redes Neurais Convolucionais no Deep Learning**. 2019. Disponível em: <<https://www.aliger.com.br/blog/as-redes-neuronais-convolutivas-no-deep-learning>>. Acesso em: 17 mar. 2021.
- ALVES, G. **Entendendo Redes Convolucionais (CNNs)**. 2018. Disponível em: <<https://medium.com/neuronio-br/entendendo-redes-convolucionais-cnns-d10359f21184>>. Acesso em: 17 mar. 2021.
- 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, v. 43, n. 2, p. 635–640, 2020.
- BACKES, A. R.; Sá JUNIOR, J. J. de M. Book. **Introdução à visão computacional usando Matlab**. [S.l.]: Alta Books Editora, 2019. ISBN 8550806285.
- BRASIL. Ministério da Saúde. **Sobre a doença**. 2020. Disponível em: <<https://coronavirus.saude.gov.br/sobre-a-doenca>>. Acesso em: 17 jun. 2020.
- BRASIL. Ministério da Saúde. **Painel de casos de doença pelo coronavírus 2019 (COVID-19) no Brasil pelo Ministério da Saúde**. 2021. Disponível em: <<https://covid.saude.gov.br>>. Acesso em: 15 de jul. 2021.
- CANDEMIR, S. et al. **Lung Segmentation in Chest Radiographs Using Anatomical Atlases With Nonrigid Registration**. *IEEE Transactions on Medical Imaging*, v. 33, n. 2, p. 577–590, 2014.
- CARVALHO, É. de et al. **Automatic Detection of COVID-19 in X-Ray Images Using Fully-Connected Neural Networks**. In: SBC. *Anais do XVI Workshop de Visão Computacional*. [S.l.], 2020. p. 41–45.
- CHUNG, A. **Figure1-COVID-chestxray-dataset**. 2020a. Disponível em: <<https://github.com/agchung/Figure1-COVID-chestxray-dataset>>.
- CHUNG, A. **Actualmed-COVID-chestxray-dataset**. 2020b. Disponível em: <<https://github.com/agchung/Actualmed-COVID-chestxray-dataset>>.

COHEN, J. P. et al. **Covid-19 image data collection: Prospective predictions are the future**. *arXiv preprint arXiv:2006.11988*, 2020. Disponível em: <<https://github.com/ieee8023/covid-chestxray-dataset>>.

DASA. **TUDO SOBRE A COVID-19**. 2020. Disponível em: <<https://dasa.com.br/coronavirus>>. Acesso em: 18 jun. 2020.

DATA SCIENCE ACADEMY. **Deep Learning Book**. 2021. Disponível em: <<https://www.deeplearningbook.com.br>>. Acesso em: 25 fev. 2021.

FORATO, F. **Não tem teste? Análise online de tomografias auxilia no diagnóstico de COVID-19**. 2020. Disponível em: <<https://canaltech.com.br/saude/nao-tem-teste-analise-online-de-tomografias-auxilia-no-diagnostico-de-covid-19-166080>>. Acesso em: 24 jun. 2020.

GARCIA, L. P.; DUARTE, E. **Intervenções não farmacológicas para o enfrentamento à epidemia da COVID-19 no Brasil**. *Epidemiologia e Serviços de Saúde*, scielo, v. 29, 00 2020. ISSN 2237-9622. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttextpid=S2237-96222020000200100nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttextpid=S2237-96222020000200100nrm=iso)>.

GOMES, P. C. T. **Introdução ao aprendizado de máquina**. 2019. Disponível em: <<https://www.datageeks.com.br/machine-learning>>. Acesso em: 21 fev. 2021.

GRÜBLER, M. **Entendendo o funcionamento de uma Rede Neural Artificial**. 2018. Disponível em: <<https://medium.com/brasil-ai/entendendo-o-funcionamento-de-uma-rede-neural-artificial-4463fcf44dd0>>. Acesso em: 25 jul. 2021.

HE, D.-C.; WANG, L. **Texture unit, texture spectrum, and texture analysis**. *IEEE transactions on Geoscience and Remote Sensing*, IEEE, v. 28, n. 4, p. 509–512, 1990.

HEIDARIAN, S. et al. **Covid-fact: A fully-automated capsule network-based framework for identification of covid-19 cases from chest ct scans**. *arXiv preprint arXiv:2010.16041*, 2020.

ISLAM, M. Z.; ISLAM, M. M.; ASRAF, A. **A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images**. *Informatics in Medicine Unlocked*, v. 20, p. 100412, 2020. ISSN 2352-9148. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2352914820305621>>.

JAEGER, S. et al. **Automatic Tuberculosis Screening Using Chest Radiographs**. *IEEE Transactions on Medical Imaging*, v. 33, n. 2, p. 233–245, 2014.

KAUFMAN, D. Book. **A inteligência artificial irá suplantará a inteligência humana?** [S.l.]: ESTAÇÃO DAS LETRAS E CORES EDI, 2019. ISBN 9788568552902.

LECUN, Y.; BENGIO, Y.; HINTON, G. **Deep learning**. *nature*, Nature Publishing Group, v. 521, n. 7553, p. 436–444, 2015.

LIMA, C. M. A. d. O. **Information about the new coronavirus disease (COVID-19)**. *Radiologia Brasileira*, scielo, v. 53, p. V – VI, 04 2020. ISSN 0100-3984. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttextpid=S0100-39842020000200001nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttextpid=S0100-39842020000200001nrm=iso)>.

MARTINS, J. V. G. et al. **Classificação da COVID-19 em Radiografias do Tórax Utilizando Redes Neurais Profundas e Padrões Binários Locais**. *Journal of Health Informatics*, v. 12, p. 387–392, 2021.

MATSUNAGA, V. Y. **Curso de Redes Neurais utilizando o MATLAB**. 2012.

Disponível em: <<https://www.muriloleal.com.br/visao/repositorio/centec/eai/ia/REDES%20NEURASIS%20-%20APOSTILA.pdf>>. Acesso em: 21 fev. 2021.

MONARD, M. C.; BARANAUSKAS, J. A. **Conceitos sobre aprendizado de máquina**. *Sistemas inteligentes-Fundamentos e aplicações*, Manole Ltda, v. 1, n. 1, p. 32, 2003.

MOONEY, P. **Chest X-Ray Images (Pneumonia)**. 2018. Disponível em: <<https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>>. Acesso em: 31 ago. 2021.

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.

NUNES, F. L. **Introdução ao processamento de imagens médicas para auxílio a diagnóstico—uma visão prática**. *Livro das Jornadas de Atualizações em Informática*, p. 73–126, 2006.

OLIVEIRA JUNIOR, W. R. D. **O USO DE REDES NEURASIS CONVOLUCIONAIS E TOTALMENTE CONECTADAS PARA A CATEGORIZAÇÃO DE TEXTOS EM IDIOMAS PORTUGUÊS E INGLÊS**. 129 f. Dissertação (Mestrado em Informática) — Pontifícia Universidade Católica do Paraná, Curitiba, 2018.

ÖZKAYA, U.; ÖZTÜRK, Ş.; BARSTUGAN, M. **Coronavirus (COVID-19) classification using deep features fusion and ranking technique**. In: *Big Data Analytics and Artificial Intelligence Against COVID-19: Innovation Vision and Approach*. [S.l.]: Springer, 2020. p. 281–295.

OZTURK, T. et al. **Automated detection of COVID-19 cases using deep neural networks with X-ray images**. *Computers in biology and medicine*, Elsevier, v. 121, p. 103792, 2020.

PATEL, P. **Chest X-ray (Covid-19 & Pneumonia)**. 2021. Disponível em: <<https://www.kaggle.com/prashant268/chest-xray-covid19-pneumonia>>. Acesso em: 01 jan. 2021.

PATROCÍNIO, A. C. Journal Article, **Pesquisadores da UFU analisam imagens radiológicas de pulmões para diagnóstico de Covid-19**. 2020. Disponível em: <<http://www.comunica.ufu.br/noticia/2020/04/pesquisadores-da-ufu-analisam-imagens-radiologicas-de-pulmoes-para-diagnostico-de>>.

PAULINO, A. **Área Abaixo da Curva ROC**. 2018. Disponível em: <<https://medium.com/ensina-ai/area-abaixo-da-curva-roc-15d2ae93a577>>. Acesso em: 7 jun. 2021.

RAHMAN, T.; CHOWDHURY, D. M.; KHANDAKAR, A. **Covid-19 radiography database**. 2020. Disponível em: <<https://www.kaggle.com/tawsifurrahman/covid19-radiography-database>>. Acesso em: 25 jun. 2020.

- RONNEBERGER, O.; FISCHER, P.; BROX, T. **U-net: Convolutional networks for biomedical image segmentation**. In: SPRINGER. *International Conference on Medical image computing and computer-assisted intervention*. [S.I.], 2015. p. 234–241.
- RUDEK, M.; COELHO, L. d. S.; CANGIOLIERI JUNIOR, O. **Visão Computacional Aplicada a Sistemas Produtivos: Fundamentos e Estudo de Caso**. *XXI Encontro Nacional de Engenharia de Produção-2001*, Salvador, 2001.
- SALOMÃO, S. C.; MARQUES, P. M. d. A. **Integrando ferramentas de auxílio ao diagnóstico no sistema de arquivamento e comunicação de imagens**. *Radiologia Brasileira*, v. 44, p. 374–380, 2011. ISSN 0100-3984. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttextpid=S0100-39842011000600009nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttextpid=S0100-39842011000600009nrm=iso)>.
- SIMONYAN, K.; ZISSERMAN, A. **Very deep convolutional networks for large-scale image recognition**. *arXiv preprint arXiv:1409.1556*, 2014.
- TEIXEIRA, J. Book. **O que é inteligência artificial**. e-galáxia, 2019. ISBN 9788584742615. Disponível em: <<http://books.google.com.br/books?id=oDSZDwAAQBAJ>>.
- Societa Italiana di Radiologia Medica e Interventistica. **COVID-19 Database**. 2020. Disponível em: <<https://sirm.org/covid-19/>>. Acesso em: 31 ago. 2021.
- VIVEIROS, V. H. et al. **Fully-Connected Neural Network for COVID-19 Chest X-Ray Imaging Classification Using Hybrid Features**. In: SBC. *Anais do XVI Workshop de Visão Computacional*. [S.I.], 2020. p. 30–35.
- WANG, X. et al. **Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases**. 2017. 2097–2106 p.
- ZHANG, Y.-D. et al. **A five-layer deep convolutional neural network with stochastic pooling for chest CT-based COVID-19 diagnosis**. *Machine Vision and Applications*, Springer, v. 32, n. 1, p. 1–13, 2021.

# Glossário

**Acurácia** É uma medida da quantidade de acertos.

**Atributo** Descreve uma característica de um exemplo. Geralmente podemos classificá-los em dois tipos: nominal, quando não se pode ordenar pelo valor e contínuo, quando há uma ordem de grandeza entre os valores.

**Classe** Pode ser considerado um atributo especial ou rótulo para um exemplo que descreve a meta atingida após a classificação de um exemplo.

**Conjunto de Exemplos** É composto por uma quantidade determinada de exemplos formado por seus valores de atributos e classe.

**Distribuição de Classes** É o valor que representa o quanto cada classe compõe do conjunto completo. Por exemplo, para um conjunto T com 100 registros, contendo as classes C1 com 60 registros, C2 com 15 registros e C3 com 25 registros. A distribuição de C1, C2 e C3 seria respectivamente 60%, 15% e 25%.

**Erro** Conhecido como taxa de erro ou taxa de classificação incorreta, é uma medida de desempenho que verifica o número de erros ocorridos em um conjunto de dados.

**Exemplo** Conhecido como caso, registro ou dado é uma sequência finita de atributos ou um vetor de atributos. Como exemplo, podemos citar dados médicos de um paciente doente, no qual o conjunto dos dados desse paciente seria um exemplo.

**Mapa de imagem** Conjunto de *pixels* gerada após aplicação de filtro.

**Prevalência de Classe** é o desbalanceamento de classes em um conjunto de exemplos. Por exemplo, se considerarmos um conjunto de exemplos T, contendo as classes C1, C2 e C3 compondo respectivamente 99%, 0,25% e 0,75%. Nesse caso, C1 tem prevalência de classe em relação a C2 e C3. Isso poderia dificultar o treinamento já que temos muitos exemplos de C1, porém em C2 e C3 temos uma quantidade muito pequena.

**ReLu** Unidade Linear Retificada (*Rectified Linear Unit*) é a função de ativação não linear, amplamente utilizada em projetar redes neurais .

**Ruído** São dados imperfeitos, que geram inconsistências no momento do treinamento. Podem ser gerados na aquisição dos dados, transformação ou na atribuição de valores para os atributos ou classe.

**Sigmoide** Função normalmente usadas para se referir à função logística.

**Softmax** Função matemática que converte um vetor de números em um vetor de probabilidades.

**Tanh** Também conhecida como tangente hiperbólica, é semelhante à função sigmoide sendo uma versão escalonada da função sigmoide.

**VGG16** Rede neural criada por (SIMONYAN; ZISSERMAN, 2014), composta por oito camadas convolucionais com filtros 3x3 *pixels*, três camadas *fully-connected* e cinco camadas de *max-pooling* com filtros 2x2 *pixels*, totalizando 16 camadas.

**VGG19** Rede neural criada por (SIMONYAN; ZISSERMAN, 2014), composta por 16 camadas convolucionais com filtros 3x3 *pixels* e três camadas *fully-connected*, totalizando 19 camadas.

**Viés (Bias)** Matematicamente dizendo, o *bias* é uma variável criada para aumentar a liberdade dos ajustes feitos nos pesos.