

VANDECIA REJANE MONTEIRO FERNANDES

Uma metodologia para a criação de
arquiteturas de redes neurais convolucionais
otimizadas para a detecção e o diagnóstico
entre pneumonias

São Luís

2021

VANDECIA REJANE MONTEIRO FERNANDES

Uma metodologia para a criação de arquiteturas de
redes neurais convolucionais otimizadas para a
detecção e o diagnóstico entre pneumonias

Tese apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da UFMA como requisito parcial para a obtenção do grau de Doutora em Engenharia Elétrica.

Orientador: Prof. Dr. Anselmo Cardoso de Paiva

Coorientador: Prof. Dr. Aristófanés Corrêa Silva

São Luís

2021

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).
Diretoria Integrada de Bibliotecas/UFMA

Monteiro Fernandes, Vandecia Rejane.

Uma metodologia para a criação de arquiteturas de redes neurais convolucionais otimizadas para a detecção e o diagnóstico entre pneumonias / Vandecia Rejane Monteiro Fernandes. - 2021.

96 p.

Coorientador(a): Aristófanês Corrêa Silva.

Orientador(a): Anselmo Cardoso de Paiva.

Tese (Doutorado) - Programa de Pós-graduação em Engenharia Elétrica/ccet, Universidade Federal do Maranhão, São Luís, 2021.

1. Aprendizado profundo. 2. Estimação bayesiana. 3. Pneumonia. 4. Raio-x. I. Cardoso de Paiva, Anselmo. II. Corrêa Silva, Aristófanês. III. Título.

Tese de autoria de Vandecia Rejane Monteiro Fernandes, sob o título “**Uma metodologia para a criação de arquiteturas de redes neurais convolucionais otimizadas para a detecção e o diagnóstico entre pneumonias**”, apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Maranhão, como requisito parcial para obtenção do título de Doutora em Engenharia Elétrica, aprovada em 23 de julho de 2021 pela comissão examinadora constituída pelos doutores:

Prof. Dr. Anselmo Cardoso de Paiva
Orientador

Prof. Dr. Aristófanés Corrêa Silva
Coorientador

Prof. Dr. João Dallyson Sousa de Almeida
Examinador Interno

Prof. Dr. João Viana da Fonseca Neto
Examinador Interno

Profa. Dra. Aura Conci
Examinador Externo

Prof. Dr. Pedro de Alcantara dos Santos Neto
Examinador Externo

Agradecimentos

Agradeço a Deus, minha família, meus orientadores e a todos que contribuíram direta ou indiretamente no desenvolvimento desta Tese.

“O sucesso não é para aqueles que correm mais rápido, mas para os que sem mantêm em movimento”

(Bangambiki Habyarimana)

Resumo

A pneumonia é uma doença que afeta os pulmões, dificultando a respiração. De acordo com estudos publicados, a pneumonia é a doença que mais mata crianças com menos de cinco anos de idade no mundo, sendo que a maior parte do número de óbitos acontece em países em desenvolvimento. Com o surgimento da COVID-19, a pneumonia tornou-se novamente uma preocupação mundial. Um grande esforço de empresas, governos, comunidade médica e acadêmica tem sido feito em todo o mundo para conter a propagação do coronavírus. A rápida disseminação do vírus SARS-CoV-2 faz com que o diagnóstico precoce seja extremamente importante. Dessa forma, a automação ou aceleração no processo diagnóstico é algo desejável. O uso de métodos computacionais podem diminuir a carga de trabalho de especialistas e até oferecer uma segunda opinião, aumentando a quantidade de diagnósticos precisos. Esta tese propõe uma metodologia para a criação de arquiteturas de redes neurais convolucionais específicas para detecção de pneumonia e classificação entre os tipos bacteriano e virais (incluindo a COVID-19), através da otimização bayesiana de redes pré-treinadas. Os resultados obtidos são promissores, atingindo, entre outras métricas, 100% de acurácia para a detecção da COVID-19 e mais de 95% de acurácia no diagnóstico entre os tipos normal, bacteriano, viral e COVID-19.

Palavras-chave: pneumonia; raio-x; aprendizado profundo; estimação bayesiana

Abstract

Pneumonia is a disease that affects the lungs, making it difficult to breathe. According to published studies, pneumonia is the disease that kills most children under the age of five, with most deaths occurring in developing countries. With the emergence of COVID-19, pneumonia has again become a worldwide concern. A great effort by companies, governments, the medical and academic community has been made worldwide to contain the spread of the coronavirus. The rapid spread of the SARS-CoV-2 virus makes early diagnosis extremely important. Thus, automation or acceleration in the diagnostic process is desirable. The use of computational methods can decrease specialists' workload and even offer a second opinion, increasing accurate diagnoses. This thesis proposes a methodology for constructing convolutional neural network architectures specific for pneumonia detection and classification between bacterial and viral types (including COVID-19) through Bayesian optimization of pre-trained networks. The results obtained are promising, reaching, among other metrics, 100% accuracy for COVID-19 detection and more than 95% accuracy in the diagnosis between the normal, bacterial, viral, and COVID-19 types

Keywords: pneumonia; chest x-ray; deep learning; Bayesian optimization

Lista de figuras

Figura 1 – Aparelho Respiratório.	30
Figura 2 – Alvéolo normal x Alvéolo com Pneumonia.	31
Figura 3 – Radiografias: Normal x Pneumonia x COVID-19	34
Figura 4 – Tipos de incidência mais realizados em radiografias do tórax.	35
Figura 5 – Representação da estrutura de uma rede neural artificial.	36
Figura 6 – Gráfico comparativo de Deep Learning versus métodos anteriores de aprendizado.	37
Figura 7 – Fluxo de dados em uma RNP.	38
Figura 8 – Arquitetura simples de uma CNN.	40
Figura 9 – Esquerda: imagem de entrada em uma CNN. Direita: exemplo de filtro de ativação.	40
Figura 10 – Processo de convolução realizado pelo filtro (azul) sobre a imagem, resultando no mapa de características (laranja).	41
Figura 11 – Estágios de uma CNN típica.	42
Figura 12 – Demonstração da operação <i>max pooling</i>	42
Figura 13 – Arquitetura VGG-16	44
Figura 14 – Bloco Residual	45
Figura 15 – Arquitetura ResNet-34	46
Figura 16 – Bloco Denso	47
Figura 17 – Arquitetura DenseNet-169	47
Figura 18 – Visualização dos campos receptivos dos filtros.	49
Figura 19 – Metodologia Proposta	55
Figura 20 – Exemplos de imagens da base fornecida por Kermany et al. (2018) . . .	56
Figura 21 – Exemplos de imagens da base fornecida por Rahman (2020)	57
Figura 22 – Dois exemplos de <i>data augmentation</i> da base para o balanceamento da mesma.	58
Figura 23 – Arquiteturas de Redes de Segmentação Semântica	60
Figura 24 – Dois exemplos de marcação manual do pulmão.	61
Figura 25 – Diagrama que ilustra os elementos do processo de estimação de topologia e parâmetros da rede convolucional proposta por esta metodologia . . .	63

Figura 26 – Ilustra o modelo de estimação da rede especializada baseado em uma sequência de blocos convolucionais concatenados.	63
Figura 27 – Ilustra o modelo de estimação da rede especializada baseado em uma sequência de blocos Inception concatenados.	64
Figura 28 – Comparação do quantitativo de imagens por DSC para todas as redes .	69
Figura 29 – Caso de estudo para o melhor resultado com a U-Net. A esquerda, um caso de sucesso com DSC= 0,980 e a direita, um caso de erro com DSC=0,663.	70
Figura 30 – Caso de estudo para o melhor resultado com a LinkNet. A esquerda, um caso de sucesso com DSC= 0,978 e a direita, um caso de erro com DSC=0.725.	70
Figura 31 – Arquitetura otimizada (VGG-16) - Detecção de Pneumonia	73
Figura 32 – Arquitetura otimizada (VGG-16) - Diagnóstico de Pneumonia	76

Lista de tabelas

Tabela 1 – Síntese dos trabalhos sobre pneumonia	25
Tabela 2 – Síntese dos trabalhos sobre COVID-19	29
Tabela 3 – Parâmetros x Hiperparâmetros	48
Tabela 4 – Decisões arquiteturais para serem avaliadas pelo otimizador ao estimar o melhor modelo.	65
Tabela 5 – Matriz de confusão para duas classes.	67
Tabela 6 – Resultados de segmentação no conjunto de validação, fase de treinamento do método	68
Tabela 7 – Resultados de segmentação no conjunto de testes, fase de aprendizado de transferência do método	69
Tabela 8 – Resultados obtidos para a Detecção de Pneumonia - VGG-16	72
Tabela 9 – Resultados obtidos para a Detecção de Pneumonia - ResNet-50	72
Tabela 10 – Resultados obtidos para a Detecção de Pneumonia - DenseNet-169	72
Tabela 11 – Hiperparâmetros estimados do melhor modelo - Detecção de pneumonia	72
Tabela 12 – Comparação de Parâmetros - Detecção de Pneumonia	73
Tabela 13 – Resultados obtidos para o Diagnóstico de Pneumonia - VGG-16	74
Tabela 14 – Resultados obtidos para o Diagnóstico de Pneumonia - ResNet-50	74
Tabela 15 – Resultados obtidos para o Diagnóstico de Pneumonia - DenseNet-169	74
Tabela 16 – Hiperparâmetros estimados do melhor modelo - Diagnóstico de pneumonia	75
Tabela 17 – Comparação de Parâmetros - Diagnóstico de Pneumonia	75
Tabela 18 – Comparação com outros trabalhos	77
Tabela 19 – Resultados obtidos para a Detecção de COVID-19 - VGG-16	78
Tabela 20 – Resultados obtidos para a Detecção de COVID-19 - ResNet-50	78
Tabela 21 – Resultados obtidos para a Detecção de COVID-19 - DenseNet-169	78
Tabela 22 – Hiperparâmetros estimados pelos melhores modelos da VGG-16 - Detecção COVID-19.	79
Tabela 23 – Hiperparâmetros estimados pelo melhor modelo da ResNet-50 - Detecção COVID-19.	79
Tabela 24 – Comparação de Parâmetros - Detecção COVID-19 - VGG-16.	79
Tabela 25 – Comparação de Parâmetros - Detecção COVID-19 - ResNet-50.	80
Tabela 26 – Resultados obtidos para o Diagnóstico (COVID-19) - VGG-16	80

Tabela 27 – Resultados obtidos para o Diagnóstico (COVID-19)- ResNet-50	80
Tabela 28 – Resultados obtidos para o Diagnóstico (COVID-19) - DenseNet-169 . .	81
Tabela 29 – Hiperparâmetros estimados pelo melhor modelo - Diagnóstico COVID-19	81
Tabela 30 – Comparação de Parâmetros - Diagnóstico COVID-19	81
Tabela 31 – Comparação com outros trabalhos	82
Tabela 32 – Artigos produzidos diretamente relacionados a Tese	84

Lista de abreviaturas e siglas

AUC	<i>Area Under the Curve</i>
CNN	<i>Convolutional Neural Network</i>
DBN	<i>Deep Belief Network</i>
DNN	<i>Deep Neural Network</i>
DRNN	<i>Deep Recurent Neural Network</i>
EI	<i>Expected Improvement</i>
FCN	<i>Fully Convolutional Network</i>
FN	<i>False Negative</i>
FP	<i>False Positive</i>
FrMEMs	<i>Fractional Multichannel Exponent Moments</i>
GPU	<i>Graphics Processing Unit</i>
IRA	Infecção Respiratória Aguda
IA	Inteligência Artificial
KNN	<i>K-Nearest Neighbor</i>
LSTM	<i>Long Short-Term Memory</i>
MERS	<i>Middle East respiratory syndrome</i>
MLP	<i>Multilayer Perceptron</i>
MRFO	<i>Manta-Ray Foraging Optimization and differential evolution</i>
mRMR	<i>Minimum Redundancy Maximum Relevance</i>
NIH	<i>National Institute of Health</i>
OMS	Organização Mundial de Saúde
ONG	Organização Não Governamental

PAC	Pneumonia Adquirida em Comunidade
RNA	Rede Neural Artificial
RNP	Rede Neural Profunda
RT-PCR	<i>Reverse-Transcriptase Polymerase Chain Reaction</i>
SARS	<i>Severe Acute Respiratory Syndrome</i>
SMBO	<i>Sequential Model Based Optimization</i>
SUS	Sistema Único de Saúde
SVM	<i>Support Vector Machine</i>
TN	<i>True Negative</i>
TP	<i>True Positive</i>
TPE	<i>Tree-structured Parzen Estimator</i>

Sumário

1	Introdução	16
1.1	Objetivos	19
1.2	Contribuições	20
1.3	Organização do trabalho	20
2	Trabalhos Relacionados	22
3	Fundamentação Teórica	30
3.1	Pneumonia	30
3.2	Radiografia do Tórax	33
3.3	Aprendizado Profundo	36
3.3.1	<i>Feedforward e Backpropagation</i>	37
3.3.2	Redes Neurais Convolucionais	39
3.3.2.1	Camada de Convolução	39
3.3.2.2	Camada de <i>Pooling</i>	41
3.3.2.3	Camada de Classificação	42
3.3.2.4	Rede Convolucional VGG	44
3.3.2.5	Rede Convolucional ResNet	45
3.3.2.6	Rede Convolucional DenseNet	46
3.3.3	Hiperparâmetros	47
3.3.4	<i>Transfer Learning</i>	49
3.4	Meta-aprendizagem	50
3.4.1	Otimização Bayesiana	51
4	Metodologia proposta	55
4.1	Aquisição das Imagens	55
4.1.1	Base de imagens de Pneumonia	56
4.1.2	Base de imagens de COVID-19	57
4.2	Pré-Processamento	58
4.2.1	Segmentação	59
4.3	Estimação e Avaliação	62

4.4	Otimização	64
5	Resultados e Discussão	66
5.1	Resultados da Segmentação do Pulmão	67
5.2	Experimentos para detecção e diagnóstico de Pneumonia	71
5.2.1	Resultados para a detecção de Pneumonia	71
5.2.2	Resultados para o diagnóstico entre Pneumonias	73
5.2.3	Comparação com outros trabalhos	76
5.3	Experimentos para detecção e diagnóstico de COVID- 19	77
5.3.1	Resultados para a detecção de COVID-19	77
5.3.2	Resultados para o diagnóstico de pneumonia (incluindo COVID-19)	79
5.3.3	Comparação com outros trabalhos	81
6	Conclusão	83
	Referências	85

1 Introdução

As infecções respiratórias agudas (IRA) constituem um grave problema de saúde pública, apresentando altas taxas de morbidade e mortalidade na população em geral (HALAJI et al., 2019). A pneumonia adquirida em comunidade (PAC) é a forma mais grave dessas infecções, sendo uma das principais causas de mortes no mundo (CORRÊA et al., 2018).

Pneumonia é uma inflamação que acomete o(s) pulmão(ões) causada por vírus, bactérias, fungos ou reações alérgicas. Essa inflamação enche o(s) pulmão(ões) de secreção, o que dificulta a respiração. Qualquer pessoa está sujeita a contrair pneumonia, mas crianças, idosos e pessoas imunocomprometidas estão mais vulneráveis ao contágio, que, em alguns casos, pode ser fatal (AMERICAN THORACIC SOCIETY, 2018). A pneumonia adquirida em comunidade é a pneumonia contraída no convívio social, fora do ambiente hospitalar, ou que se manifesta em até 48 horas após a internação.

De acordo com um estudo feito pela ONG *Save the Children* em conjunto com a Universidade John Hopkins, a pneumonia é a doença que mais mata crianças menores de 5 anos no mundo (SAVE THE CHILDREN, 2018). As mortes infantis causada pela pneumonia são maiores que as mortes infantis causadas por malária, diarreia e sarampo somadas. Somente em 2017, 808.694 crianças morreram devido a essa doença (WORLD HEALTH ORGANIZATION, 2019). Ainda de acordo com os dados de Save the children (2018), a pneumonia matará 11 milhões de crianças até 2030, se nenhuma medida for tomada. A Organização Mundial de Saúde (OMS) há mais de duas décadas tem promovido discussões acerca da pneumonia. Diversos estudos foram publicados na tentativa de diminuir a taxa de mortalidade (WORLD HEALTH ORGANIZATION, 2019).

No Brasil, considerando a taxa de mortalidade por pneumonia em crianças menores de 5 anos no início dos anos 2000, houve uma redução significativa no número de mortes, caindo de 101 mortes a cada 100 mil crianças, para 29 mortes a cada 100 mil (DADONAITE; ROSER, 2020). A melhora da situação socioeconômica, o maior acesso a cuidados de saúde, a disponibilidade nacional de antibióticos e as políticas de vacinação podem ser responsáveis por essa redução de casos (CORRÊA et al., 2018)

A pneumonia também é uma das maiores causadoras de internações hospitalares no mundo, acarretando altos custos aos serviços de saúde. No Brasil, de janeiro a novembro de 2019, 554.184 pessoas foram hospitalizadas devido a pneumonia, representando gastos

de um pouco mais de R\$ 490 milhões com serviços hospitalares (MINISTÉRIO DA SAÚDE, 2020).

A mortalidade infantil por pneumonia está fortemente ligada à desnutrição, pobreza e acesso inadequado a cuidados de saúde. Assim, a incidência costuma ser maior em locais onde falta infraestrutura de saúde e as pessoas carecem de recursos financeiros (UNICEF; ORGANIZATION, 2008). A maior parte dos casos de pneumonia encontra-se em países em desenvolvimento, sendo esses também responsáveis pela maior parte do número de óbitos. Apenas 5 países: Índia, Nigéria, Paquistão, República Democrática do Congo e Etiópia, foram responsáveis por mais da metade de todas as mortes por pneumonia infantil em 2017 (DADONAITE; ROSER, 2020).

Contudo, um novo tipo de infecção respiratória tornou a pneumonia uma preocupação mundial. Um novo vírus, da família dos coronavírus, foi identificado pela primeira vez em Wuhan, na China, em dezembro de 2019. O vírus SARS-CoV-2 mostrou-se altamente contagioso e rapidamente se espalhou pelo mundo. Em fevereiro de 2020 a doença causada pelo vírus SARS-CoV-2 foi nomeada de COVID-19 e, no dia 11 de março de 2020, a Organização Mundial da Saúde declarou pandemia.

De acordo com dados do Sistema Único de Saúde (SUS), em 2020 observou-se uma queda de aproximadamente 25% no quantitativo de internações no país em relação a 2019. Acredita-se que essa queda esteja diretamente relacionada a a COVID-19. Devido às recomendações de distanciamento social e muitas vezes receando ser contagiada, a população reduziu intensamente a busca por serviços de saúde, incluindo as emergências dos hospitais, além da suspensão por parte do SUS das consultas ambulatoriais e cirurgias eletivas. Considerando somente as internações por doenças virais, houve um aumento expressivo no número de internações, alcançando o segundo lugar dentre as principais causas de morbidade nas internações no ano de 2020 (ALBUQUERQUE, 2020).

O espectro clínico da COVID-19 varia de infecções assintomáticas a quadros graves, nos quais ocorre uma pneumonia severa. O diagnóstico da doença é feito por um médico através de exame clínico. Para confirmar o diagnóstico geralmente são solicitados alguns exames adicionais, como exames laboratoriais ou de imagem (MINISTÉRIO DA SAÚDE, 2021).

O exame laboratorial *Reverse-Transcriptase Polymerase Chain Reaction (RT-PCR)* permanece como padrão de referência para o diagnóstico definitivo de infecção por COVID-19. Esse exame detecta também a Influenza ou a presença de Vírus Sincicial Respiratório

(VSR) normalmente até o oitavo dia de início de sintomas. As amostras são colhidas por *swab* nasal ou orofaríngeo e por material de escarro ou lavado broncoalveolar (MINISTÉRIO DA SAÚDE, 2021).

Os exames de imagem que tem sido utilizado amplamente no combate a pandemia são a radiografia do tórax e a tomografia computadorizada. A radiografia é o exame de imagem mais antigo do mundo (SOCIEDADE BRASILEIRA DE REUMATOLOGIA, 2011). O surgimento de outros exames de imagem, como ultrassonografia, tomografia computadorizada e ressonância nuclear magnética, não foi capaz de diminuir a importância da radiografia simples no atendimento clínico e emergencial. Sua ampla disponibilidade, baixo custo e rapidez, podendo ser realizada “no leito” em paciente com instabilidade clínica, fazem da radiografia simples um dos primeiros exames complementares a serem solicitados na maioria dos casos atendidos (BRENTANO et al., 2014). Além da contribuição ao diagnóstico, a radiografia de tórax permite ainda avaliar a extensão das lesões, detectar complicações e auxiliar no diagnóstico diferencial (CORRÊA et al., 2018).

A radiografia do tórax, embora tenha baixa sensibilidade especialmente nas formas leves e precoces da COVID-19, pode ser útil no diagnóstico diferencial e avaliação inicial em paciente ambulatorial ou no serviço de emergência, principalmente no acompanhamento evolutivo de pacientes hospitalizados com suspeita da doença. (CAPONE et al., 2020)

Contudo, por ser um exame de imagem mais simples, interpretar corretamente uma radiografia do tórax requer uma observação cuidadosa (RAOOF et al., 2012). Entender o resultado de uma radiografia requer experiência e, algumas vezes, o mesmo exame pode resultar em diagnósticos diferentes dependendo do especialista (DELRUE et al., 2011). Devido às complexas patologias e mudanças sutis de textura, os radiologistas podem cometer erros mesmo quando bem treinados. As principais formas de minimizar os erros de interpretação são o conhecimento profundo da anatomia e da fisiologia básica das doenças do tórax. A interpretação correta das radiografias do tórax além de acelerar o início do tratamento, muitas vezes evita a necessidade de exames mais caros e sofisticados (DELRUE et al., 2011).

A Inteligência Artificial (IA) envolve o desenvolvimento de algoritmos para executar tarefas tipicamente associadas à inteligência humana (HE et al., 2019). Aprendizado Profundo, do inglês *Deep Learning*, é uma subárea da inteligência artificial, inspirada na estrutura do cérebro humano. Os algoritmos de aprendizado profundo tentam tirar conclusões semelhantes às dos humanos, analisando continuamente os dados com uma

determinada estrutura lógica. Para conseguir isso, o aprendizado profundo usa uma estrutura de várias camadas denominada redes neurais. A aprendizagem profunda imita a atividade em camadas do neurônios no neocórtex, simulando o arranjo de neurônios em uma rede neural artificial (LECUN; BENGIO, 1995). Esses algoritmos tem mostrado avanços notáveis na capacidade das máquinas de compreender e manipular dados, incluindo imagens, linguagem e fala (ESTEVA et al., 2019).

A capacidade desses algoritmos de extrair abstrações complexas e de alto nível de grandes volumes de dados, especialmente dados não estruturados, torna-os uma ferramenta valiosa na área da saúde. Com o grande volume de dados gerado devido a crescente proliferação de dispositivos médicos e sistemas de registro digital (ESTEVA et al., 2019), a aprendizagem profunda tem obtido resultados promissores na medicina, como por exemplo, na detecção de glaucoma (HAVAELI, 2017), tumores cerebrais (CERENTINI. et al., 2017) e câncer de pele (ESTEVA et al., 2017). O objetivo geral do uso de aprendizagem profunda na medicina é usar algoritmos para descobrir informações relevantes dos dados, auxiliando na tomada de decisão clínica através da interpretação rápida e precisa.

Após mais de um ano da declaração de pandemia, a COVID-19 continua a assolar os países do mundo. Dados da Organização Mundial da Saúde mostram que, em março de 2021, já são 124.215.843 casos confirmados, incluindo 2.734.374 mortes. Estados Unidos, Brasil, Índia, Rússia e Reino Unido são os países com maior números de casos e Estados Unidos e Brasil lideram os números de mortes (WORLD HEALTH ORGANIZATION, 2020). Com a urgente necessidade de controlar a pandemia, novas estratégias de diagnóstico se fazem necessárias, já que os testes laboratoriais atualmente disponíveis muitas vezes não estão acessíveis a população infectada. Sabendo que o diagnóstico rápido e preciso é o principal fator para reduzir a taxa de transmissão da COVID-19 e a redução da mortalidade por pneumonia, a automatização ou aceleração no processo de diagnóstico é algo desejável. O uso de métodos computacionais podem diminuir a carga de trabalho dos especialistas e ainda oferecer uma segunda opinião, aumentando a rapidez do tratamento.

1.1 Objetivos

Considerando que a região suspeita de pneumonia em uma imagem de raio-x torácico possui informações de difícil percepção, o principal objetivo deste trabalho consiste em propôr uma metodologia para estimação de uma rede neural convolucional específica para

detecção de pneumonia e classificação quanto aos tipos viral, bacteriano e COVID-19. Para tanto, a construção dessa nova arquitetura é baseada nos princípios de *transfer learning* (YOSINSKI et al., 2014) e na estimação da topologia da rede, assim como de seus hiperparâmetros, utilizando um modelo de otimização Bayesiano (BERGSTRA et al., 2011).

Destacam-se como objetivos específicos deste trabalho:

- Propor e validar uma metodologia para a estimação de arquiteturas de redes neurais convolucionais com aproveitamento de pré-treinamento em outros problemas;
- Desenvolver uma estratégia para realizar a segmentação dos pulmões;
- Desenvolver um modelo otimizado capaz de detectar a pneumonia e fazer a classificação entre os diferentes tipos de pneumonia;
- Avaliar o funcionamento e acurácia dos modelos encontrados com bases de imagens de radiografias do tórax.

1.2 Contribuições

Esta tese apresenta como principais contribuições:

- Uma nova abordagem para extração e combinação otimizada de redes neurais pré-treinadas para a detecção e diagnóstico da pneumonia;
- Uma nova abordagem para segmentação de pulmão em base de imagens que não possuam marcação de um especialista;
- Uma estratégia para viabilizar uma arquitetura otimizada de redes neurais convolucionais através da adaptação da otimização Bayesiana.

1.3 Organização do trabalho

Esta tese está organizada em seis capítulos, de forma a apresentar o conteúdo mais claramente, conforme os parágrafos a seguir.

O Capítulo 2, Trabalhos Relacionados, apresenta como alguns trabalhos trataram a detecção automatizada de pneumonia.

O Capítulo 3, Fundamentação Teórica, apresenta o embasamento teórico necessário para o entendimento e desenvolvimento da metodologia proposta.

O Capítulo 4, Metodologia Proposta, apresenta as etapas necessárias no desenvolvimento do modelo proposto para a detecção e diagnóstico da pneumonia.

O Capítulo 5, Resultados e Discussão, apresenta e discute os resultados obtidos pela metodologia proposta, comparando-os com os resultados obtidos pelos trabalhos relacionados.

O Capítulo 6, Conclusão, apresenta as considerações finais acerca do trabalho realizado e as propostas de trabalhos futuros.

2 Trabalhos Relacionados

Neste capítulo são apresentados trabalhos que utilizaram técnicas de inteligência artificial para determinar presença de pneumonia através de imagens.

Inteligência artificial (IA) é a tecnologia capaz de simular o processo humano de aprendizado e tomada de decisões. Na IA os algoritmos funcionam de forma semelhante ao nosso cérebro: processam informações baseadas em aprendizado de experiências anteriores. A inteligência artificial está continuamente à nossa volta, de carros e drones autônomos a assistentes virtuais e softwares automáticos de tradução ou investimento (KLAUS SCHWAB, 2016). Nos últimos anos, a IA foi impulsionada por aumentos exponenciais no poder da computação e pela disponibilidade de grandes quantidades de dados.

Com o grande volume de dados que está sendo gerado e a crescente proliferação de dispositivos médicos e sistemas de registro digital (ESTEVA et al., 2019), a área da medicina e de cuidados da saúde tem sido alvo constante de aplicação de sistemas que utilizam inteligência artificial.

Na área médica, a utilização da IA começou a impactar os especialistas, predominantemente através da interpretação rápida e precisa da imagem; os sistemas de saúde, melhorando o fluxo de trabalho e potencialmente reduzindo erros médicos; e os pacientes, permitindo que eles processem seus próprios dados para promover a saúde (TOPOL, 2019).

Como veremos nos trabalhos a seguir, já há algum tempo a pneumonia vem sendo objeto de estudo dos métodos automáticos de detecção.

Em 2017, o *National Institute of Health (NIH)* disponibilizou publicamente uma das maiores bases de radiografias de tórax, consistindo de 108.948 imagens radiográficas de 32.797 pacientes, abrangendo oito tipos de doenças torácicas: atelectasia, cardiomegalia, efusão, infiltração, massa, nódulo, pneumonia e pneumotórax (WANG et al., 2017). Posteriormente o NIH expandiu seu conjunto de dados para 112.120 imagens, incluindo mais seis doenças torácicas: consolidação, edema, enfisema, fibrose, derrame pleural e hérnia. Essa base de imagens foi chamada de ChestX-ray14.

Alguns trabalhos utilizaram a base ChestX-ray14: Kumar, Grewal e Srivastava (2018) usaram uma DenseNet para classificar as 14 doenças, obtendo acurácia de 0,70 para a pneumonia. Yao et al. (2017) utilizaram uma DenseNet junto com uma LSTM (*Long Short-Term Memory*), obtendo acurácia de 0,71 na detecção da pneumonia. Rajpurkar et

al. (2017) propuseram uma DenseNet para realizar a extração automática de características e a classificação. O trabalho atingiu acurácia de 0,76 para a pneumonia. Antin, Kravitz e Martayan (2017) usaram uma DenseNet-121 para classificar se um dado raio-X contém pneumonia, alcançando AUC de 0,60. Além de usar um modelo de aprendizado profundo, eles também usaram um modelo de regressão logística para fazer a detecção de pneumonia. O modelo de regressão logística atingiu AUC de 0,60. Yao et al. (2018) apresentam um *ensemble* das redes ResNetV2 e Fully Convolutional Network (FCN) com uma função de *pooling* customizada para identificar e localizar as várias doenças da base ChestX-ray14. Esse modelo atingiu AUC de 0,69 na detecção de pneumonia. Usando um subconjunto de imagens do ChestX-ray14, Sirazitdinov et al. (2019) propõem um *ensemble* de duas Redes Neurais Convolucionais: RetinaNet e Mask R-CNN para detecção e localização de pneumonia. O modelo atingiu precisão de 0,75.

Utilizando a base de imagens MIMIC-CXR (JOHNSON et al., 2019), Rubin et al. (2018) propõem a DualNet, uma versão modificada da DenseNet-121 que aceita entradas de raio-X de tórax frontais e laterais emparelhadas. Essa solução atingiu AUC de 0,62 para detecção de pneumonia. Já Bertrand, Hashir e Cohen (2019) utilizam a base de imagens PadChest(BUSTOS et al., 2020) para comparar o uso de incidências radiológicas pósterio-anterior (PA) lateral e frontal usando um modelo DenseNet simples. Esse trabalho mostra que o uso da vista lateral aumentou a AUC de 8 das 56 doenças e atingiu o mesmo desempenho que a exibição PA para 21 doenças. Uma dessas doenças é a pneumonia.

Embora as bases de imagens ChestX-ray14(WANG et al., 2017), CheXpert(IRVIN et al., 2019), MIMIC-CXR, (JOHNSON et al., 2019) e PadChest (BUSTOS et al., 2020) sejam compostas de milhares de imagens de raios-X, tais bases contém poucos exemplos de pneumonia, sendo muito desbalanceadas. Em 2018, Kermany et al. (2018) publicou um conjunto de imagens de raios-X composto por 5.232 imagens apenas de casos de pneumonia. Essa base de imagens é formada por radiografias pediátricas, apresentando pacientes saudáveis e pacientes com pneumonia viral ou bacteriana. Nessa base, Kermany et al. (2018) alcançaram resultados promissores utilizando *transfer learning* com o modelo Inception, alcançando uma acurácia de 0,96 na detecção de pneumonia e 0,90 no diagnóstico entre os tipos de pneumonia bacteriana e viral.

Os trabalhos a seguir utilizaram a base proposta por Kermany et al. (2018). Utilizando uma VGG-19 otimizada através da otimização bayesiana, Sousa, Fernandes e Paiva (2019) obtiveram acurácia de 0,95 para detecção de pneumonia e 0,83 para o

diagnóstico entre pneumonias. Utilizando segmentação do pulmão e as características obtidas com uma VGG-16 submetidas a uma *Multilayer Perceptron (MLP)* otimizada, o trabalho de Rajaraman et al. (2018) atinge acurácia de 0,96 para detecção de pneumonia e 0,93 no diagnóstico entre os tipos bacteriana e viral. Ainda utilizando redes neurais convolucionais para a detecção de pneumonia, Saraiva et al. (2019) atingiram acurácia de 0,95 e Stephen et al. (2019) obtiveram acurácia de 0,93. Reforçando a ideia de *transfer learning*, Liang e Zheng (2019) utilizaram uma DenseNet pré-treinada na base apresentada em Wang et al. (2017) para tratar o problema de detecção de pneumonia. A rede conseguiu adaptar-se bem ao problema, obtendo uma acurácia de 0,90. Ayan e Ünver (2019) usaram dois modelos de redes neurais convolucionais, Xception e VGG-16, para detecção de pneumonia. Os resultados obtidos mostraram que a VGG-16 obteve melhor acurácia, atingindo 0,87. O trabalho de Rahman et al. (2020) analisa o desempenho de quatro diferentes arquiteturas de redes pré-treinadas: AlexNet, ResNet-18, DenseNet-201 e SqueezeNet para detectar pneumonia e classificar entre pneumonia bacteriana e viral. Nesse trabalho, a DenseNet-201 superou os outros modelos, atingindo uma acurácia de 0,98 na detecção de pneumonia e acurácia de 0,95 no diagnóstico entre as pneumonias. Hashmi et al. (2020), propõem um modelo que combina as previsões ponderadas de quatro modelos de aprendizagem profunda: ResNet18, Xception, InceptionV3, DenseNet-121 e MobileNetV3. As previsões desses modelos foram combinadas para calcular a previsão final, atingindo uma acurácia de 0,98 na detecção de pneumonia. Em Toğaçar et al. (2020) três modelos de CNN pré-treinados, AlexNet, VGG-16 e VGG-19, são usados para extrair características. As características extraídas foram combinadas usando o algoritmo *minimum redundancy maximum relevance (mRMR)*. O melhor desempenho atingiu uma acurácia de 0,99. Moujahid et al. (2020) comparam cinco modelos de CNN: VGG-16, VGG-19, NasNetMobile, ResNet152V2 e InceptionResNetV2 para detectar a pneumonia. O modelo que obteve a melhor acurácia foi o VGG-16, atingindo 0,96. Chouhan et al. (2020) também combinam cinco arquiteturas de redes pré-treinadas para detectar a pneumonia. As saídas de cada rede pré-treinada foram combinadas em um vetor de previsão e a classe majoritária foi usada para chegar a uma previsão final de detecção, obtendo acurácia de 0,93.

A Tabela 1 sumariza os trabalhos citados sobre detecção e diagnóstico de pneumonia.

Tabela 1 – Síntese dos trabalhos sobre pneumonia

Trabalho	Tipo		Base			Resultados	
	Det.	Diag.	A	B	C	Acurácia	AUC
Kumar, Grewal e Srivastava (2018)	x		x			0.70	
Yao et al. (2017)	x		x			0.71	
Rajpurkar et al. (2017)	x		x			0.76	
Antin, Kravitz e Martayan (2017)	x		x			0.69	
Yao et al. (2018)	x		x				0.69
Rubin et al. (2018)	x			x			0.62
Kermany et al. (2018)	x				x	0.96	
Kermany et al. (2018)		x			x		0.90
Sousa, Fernandes e Paiva (2019)	x				x	0.95	
Sousa, Fernandes e Paiva (2019)		x			x		0.83
Rajaraman et al. (2018)	x				x	0.96	
Rajaraman et al. (2018)		x			x		0.93
Saraiva et al. (2019)	x				x	0.95	
Stephen et al. (2019)	x				x	0.93	
Liang e Zheng (2019)	x				x	0.90	
Ayan e Ünver (2019)	x				x	0.87	
Rahman et al. (2020)	x				x	0.98	
Rahman et al. (2020)		x			x		0.95
Hashmi et al. (2020)	x				x	0.98	
Toğaçar et al. (2020)	x				x	0.99	
Moujahid et al. (2020)	x				x	0.96	
Chouhan et al. (2020)	x				x	0.93	

Bases - (A):Wang et al. (2017), (B):Johnson et al. (2019), (C):Kermany et al. (2018)

Um novo tipo de pneumonia viral tem concentrado a atenção da humanidade. A pandemia de COVID-19 causada pelo vírus SARS-CoV-2 se espalhou rapidamente em todo o mundo, levando a um surto global. Um grande esforço de empresas, governos, comunidade médica e acadêmica tem sido feito em todo o mundo para conter a propagação do coronavírus. A rápida disseminação do vírus SARS-CoV-2 faz com que o diagnóstico precoce seja extremamente importante. O sucesso do uso de inteligência artificial em vários problemas na área da saúde na última década, como detecção de tuberculose (ROHILLA; HOODA; MITTAL, 2017), (LAKHANI; SUNDARAM, 2017), (MURPHY et al., 2019), detecção de nódulos e massas no pulmão (TANG; KIM; XIE, 2018), (XIE, 2018), (MA et al., 2018), (YASAKA et al., 2018), (TANG; LIU; XIE, 2019), detecção de câncer (ARDILA et al., 2019), (YALA et al., 2019), (KATHER et al., 2019), (LAKSHMANAPRABU et al., 2019), (XU et al., 2019), (COCCIA, 2020), retinopatia diabética (TING et al., 2019), (SAHLSTEN et al., 2019),

glaucoma (ABBAS, 2017), (LI et al., 2018), (ASAOKA et al., 2019), entre outros, reforçam o potencial que a IA tem para contribuir no controle e diagnóstico da COVID-19.

A primeira base de imagens pública contendo casos de COVID-19 (COHEN; MORRISON; DAO, 2020) foi disponibilizada em Fevereiro de 2020. Com o avanço da pandemia, outros conjuntos de imagens de raio-X e tomografia computadorizada surgiram, como Chung (2020), Rahman (2020), Wang, Lin e Wong (2020), Vayá et al. (2020), entre outros. Com a disponibilização pública de imagens, vários trabalhos de detecção automática de COVID-19 começaram a ser publicados.

Em Wang, Lin e Wong (2020) é proposta a COVID-Net, uma rede neural convolucional projetada para a detecção da COVID-19 a partir de imagens de raio-X. A COVID-Net foi pre-treinada na base ImageNet e depois treinada na base COVIDx, também proposta pelos autores. A base COVIDx é composta por outras cinco bases públicas de imagens de COVID-19. A COVID-Net atingiu acurácia geral de 0,93. Asnaoui e Chawki (2020) utilizam imagens de raio-X e tomografia computadorizada para fazer um comparativo de sete modelos de redes neurais: VGG-16, VGG-19, DenseNet-201, InceptionResNetV2, InceptionV3, Resnet-50 e MobileNetV2, para classificação de imagens em normal, pneumonia bacteriana e pneumonia por coronavírus. O modelo que obteve melhor acurácia foi o InceptionResnetV2, atingindo 0,92. Misra et al. (2020) propuseram um *ensemble* de ResNets para detectar a COVID-19. Eles treinaram três modelos de ResNet18: um modelo foi treinado com imagens de pulmão normal e imagens de pulmão com alguma outra doença. O segundo modelo foi treinado com imagens de casos de pneumonia e casos de não pneumonia e o terceiro modelo foi treinado com imagens de casos de COVID-19 e casos de não COVID-19. As camadas de classificação dos modelos foram retiradas e foi feito o *ensemble* dos três modelos treinados para classificar as imagens em normal, pneumonia e COVID-19. Essa abordagem atingiu acurácia de 0,95. Ozturk et al. (2020) propõem a DarkCovidNet: um modelo de CNN para a detecção automática de COVID-19. O modelo proposto possui 17 camadas convolucionais e é formado por camadas chamadas de DarkNet. Cada camada DarkNet é composta por uma camada convolucional seguida pelas operações *batch normalization* e *LeakyRelu*. Esse modelo atingiu acurácia de 0,98 na detecção da COVID-19 e acurácia de 0,87 no diagnóstico entre outras pneumonias.

Pereira et al. (2020) exploram diferentes tipos de métodos de extração de características para identificar a COVID-19 dentre outras pneumonias. Essa abordagem atingiu F1-Score de 0,89. Khan, Shah e Bhat (2020) apresentam a Coronet, uma Rede

Neural Convolutacional para a detecção automática da COVID-19. O modelo utilizado é baseado no modelo Xception e contém 71 camadas treinadas na base ImageNet. Esse modelo atingiu acurácia média de 0,87 na identificação da COVID-19. Em Ouchicha e Ammor (2020) é proposta a CVDNet, uma ResNet customizada para detecção e diagnóstico da COVID-19. O modelo apresentado usa diferentes tamanhos de *kernel* para extrair características locais e globais das imagens de entrada. Esse modelo atinge acurácia média de 0,97 na detecção da COVID-19 e acurácia média de 0,96 no diagnóstico entre imagens de pulmão normal, com pneumonia viral ou com COVID-19. Brunese et al. (2020) usam dois modelos de VGG-16 para analisar imagens de raio-X. O primeiro modelo visa detectar se um determinado raio-X pertence a um paciente saudável ou a um paciente com alguma doença pulmonar. Caso o raio-X seja classificado com alguma doença pulmonar, o segundo modelo irá determinar se a doença é COVID-19. O modelo alcançou acurácia de 0,98 na classificação da COVID-19. Em Nour, Cömert e Polat (2020) uma CNN é utilizada somente para extrair características das imagens. A etapa de classificação das imagens de COVID é feita com algoritmos de aprendizado de máquina. Os algoritmos avaliados foram o *K-Nearest Neighbor (KNN)*, *Support Vector Machine (SVM)* e árvore de decisão. Os hiperparâmetros dos modelos de aprendizado de máquina foram otimizados utilizando otimização bayesiana. O melhor resultado alcançado foi utilizando o SVM, atingindo acurácia de 0,98.

Chowdhury et al. (2020) investigam a utilidade da inteligência artificial para a detecção da COVID-19. Para isso, os autores avaliam o desempenho de oito modelos de CNN: SqueezeNet, MobileNetv2, ResNet-18, InceptionV3, ResNet-101, CheXNet, DenseNet-201 e VGG-19. Cada modelo foi treinado em uma base com *data augmentation* e sem *data augmentation*. A DenseNet-201 atingiu a maior acurácia tanto na classificação binária, alcançando 0,99, quanto na multiclasse, atingindo 0,97. Para a detecção da COVID-19 em imagens de raio-X, Elaziz et al. (2020) propõem um método que utiliza *Fractional Multichannel Exponent Moments (FrMEMs)* para extrair características das imagens de raios-X e em seguida, aplicam uma versão modificada do algoritmo *Manta-Ray Foraging Optimization and differential evolution (MRFO)* para selecionar as características mais relevantes extraídas. A classificação das imagens de COVID-19 é feita utilizando o algoritmo KNN, atingindo acurácia de 0,98. Ucar e Korkmaz (2020) introduzem a COVIDiagnosis-Net, um modelo de aprendizado profundo baseado na SqueezeNet e otimizado através da otimização bayesiana. O modelo proposto alcançou acurácia de 0,98 no diagnóstico entre

imagem normal, COVID-19 e pneumonia. Das et al. (2020) utilizam o modelo Xception para classificar as imagens de raio-X em normal, pneumonia e COVID-19. O modelo proposto atingiu acurácia de 0,97. Em Rahimzadeh e Attar (2020) é apresentado um modelo de rede de aprendizado profundo que é composto pela concatenação dos modelos Xception e ResNet50V2. A acurácia média do modelo para classificação das imagens de raio-X em normal, pneumonia e COVID-19 é 0,91.

Gupta et al. (2021) apresentam a InstaCovNet-19, um modelo de aprendizado profundo composto por cinco outros modelos para detectar a COVID-19. Os modelos utilizados são: InceptionV3, MobileNetV2, ResNet-101, NASNet e Xception. Cada modelo recebe uma cópia da imagem de entrada para fazer a extração de características de forma paralela. A saída de cada modelo é combinada em uma nova camada e uma última camada faz a classificação das imagens. Essa proposta alcançou acurácia de 0,99 na detecção da COVID-19 e também acurácia de 0,99 na classificação entre pneumonia, COVID-19 e normal. Jia, Lam e Xu (2021) propõem duas CNNs para fazer a detecção da COVID-19 através de imagens de raio-X e tomografia computadorizada (CT). Uma MobileNet modificada é utilizada para fazer a classificação em imagens de raio-X, atingindo a acurácia de 0,99 na classificação em 5 classes (COVID-19, pneumonia bacteriana, pneumonia viral(exceto COVID-19), tuberculose e saudável), e uma ResNet modificada é utilizada para fazer a detecção da COVID-19 em imagens de CT, atingindo acurácia de 0,99. Em Quan et al. (2021) é proposta a DenseCapsNet, uma rede formada pela fusão de uma DenseNet-121 e uma *Capsule Network* (CapsNet). Com a utilização de imagens de raio-X segmentadas, a DenseCapsNet atinge acurácia de 0,90 na detecção da COVID-19. A CovidXrayNet, um modelo de *deep learning* baseado na EfficientNet-B0, é proposto em Monshi et al. (2021). A classificação de imagens de raios-X em normal, COVID-19 ou pneumonia alcançou acurácia de 0,95 nesse trabalho. Aslan et al. (2021) propõem uma arquitetura formada por uma AlexNet com uma camada adicional composta pela rede BiLSTM (*Bidirectional LSTM*) para fazer a detecção da COVID-19, além de um método para fazer a segmentação do tórax nas imagens de raio-X. A arquitetura proposta atingiu acurácia de 0,98 na classificação entre normal, pneumonia viral e COVID-19.

Para melhor compreensão, a Tabela 2 sumariza os trabalhos citados sobre COVID-19.

A metodologia proposta nesta Tese utiliza as bases de imagens fornecida por Kermany et al. (2018) e Rahman (2020) para treinamento e validação de suas técnicas.

Tabela 2 – Síntese dos trabalhos sobre COVID-19

Trabalho	Tipo		Base								Acurácia
	Detec.	Diag.	A	B	C	D	E	F	G	H	
Wang, Lin e Wong (2020)	x		x								0.93
Asnaoui e Chawki (2020)	x			x	x						0.92
Misra et al. (2020)	x			x		x				x	0.95
Ozturk et al. (2020)	x			x			x				0.98
Ozturk et al. (2020)		x		x			x				0.87
Ouchicha e Ammor (2020)		x						x			0.96
Brunese et al. (2020)	x			x			x			x	0.98
Nour, Cömert e Polat (2020)		x						x			0.98
Chowdhury et al. (2020)	x							x			0.99
Chowdhury et al. (2020)		x						x			0.97
Elaziz et al. (2020)	x			x							0.98
Ucar e Korkmaz (2020)		x	x								0.98
Das et al. (2020)		x		x			x				0.97
Rahimzadeh e Attar (2020)		x		x		x					0.91
Gupta et al. (2021)	x			x				x			0.99
Gupta et al. (2021)		x		x				x			0.99
Jia, Lam e Xu (2021)		x	x								0.99
Quan et al. (2021)	x			x				x	x		0.98
Quan et al. (2021)		x		x				x	x		0.90
Monshi et al. (2021)		x	x								0.95
Aslan et al. (2021)		x						x			0.98

Bases - (A):Wang, Lin e Wong (2020), (B):Cohen, Morrison e Dao (2020), (C):Kermany et al. (2018), (D):RSNA (2019), (E):Wang et al. (2017), (F):Rahman (2020), (G):Praveen (2020), (H):Privado

Embora vários dos trabalhos relacionados aqui tenham explorado *transfer learning* como mecanismo de aproveitamento de treinamento para o reconhecimento de pneumonia, a metodologia proposta neste trabalho diferencia-se das demais por supor que existe a possibilidade da criação de topologias completas de redes neurais convolucionais que possam combinar pré-treinamentos juntamente com especializações, a fim de gerar um método global único e otimizado para o problema de detecção e diagnóstico de pneumonia. Essa suposição se sustenta na teoria de *transfer learning*, onde a especialização de treinamento prévio pode levar a determinação de características complexas específicas do problema tratado. E ainda que diferentes topologias de rede possuem capacidades de caracterização diferentes, assim supõe-se que estejam adquirindo características complementares que, quando utilizadas conjuntamente, podem se tornar o modelo mais eficiente da resolução do problema.

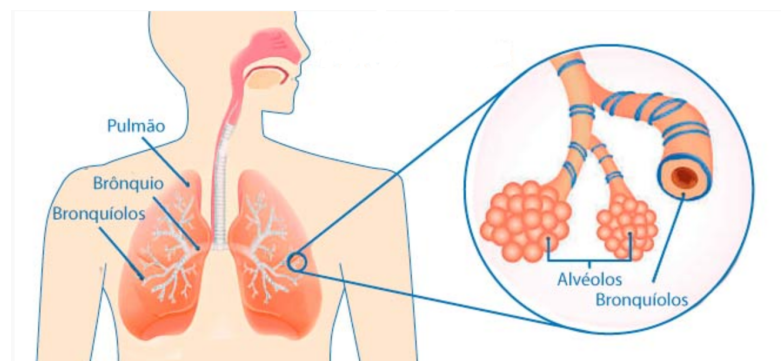
3 Fundamentação Teórica

Este capítulo apresenta a fundamentação teórica utilizada no desenvolvimento deste trabalho e necessária para compreensão das técnicas utilizadas na metodologia proposta.

3.1 Pneumonia

A Pneumonia é uma infecção que acomete os pulmões e que pode ser causada por bactérias, vírus, fungos ou reações alérgicas (AMERICAN LUNG ASSOCIATION, 2019). Essa infecção faz com que os alvéolos dos pulmões (Figura 1) fiquem inflamados e se encham com secreção.

Figura 1 – Aparelho Respiratório.

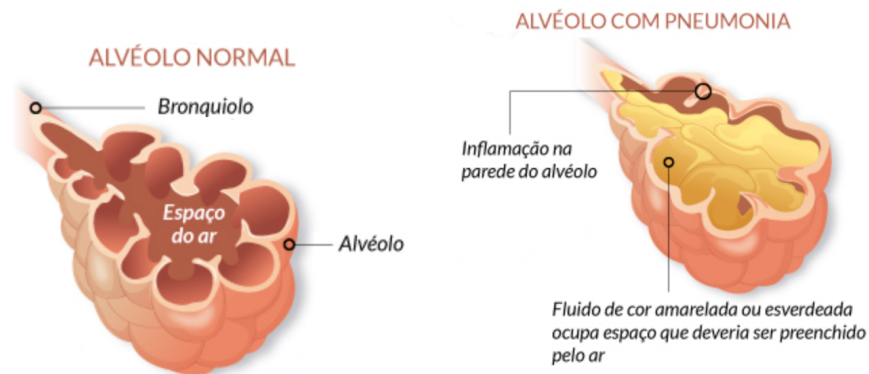


Fonte: (COSTA, 2019)

Os alvéolos são pequenas bolsas através das quais são feitas as trocas dos gases respirados (oxigênio e gás carbônico). O alvéolo entrega o oxigênio respirado para as hemácias (glóbulos vermelhos) e recebe delas o gás carbônico. Na pneumonia, os alvéolos ficam cheios de secreções (líquido ou pus), impedindo a entrada e saída dos gases (Figura 2). Nesses alvéolos infeccionados não há troca de oxigênio por gás carbônico, dificultando e até impedindo a respiração. Quanto mais alvéolos acometidos pela pneumonia, mais grave é o quadro (COSTA, 2019).

Considerando o local onde foi adquirida, a pneumonia pode ser classificada como comunitária ou hospitalar. A pneumonia adquirida na comunidade ou comunitária, é aquela adquirida fora do hospital ou nas primeiras 48 horas de internamento hospitalar. Já a pneumonia hospitalar ou nosocomial é aquela que se adquire entre as 48 horas após o

Figura 2 – Alvéolo normal x Alvéolo com Pneumonia.



Fonte: Adaptada de Costa (2019)

internamento hospitalar e as 72 horas após a alta, mas que não estava em incubação no momento da entrada (SCHWARTZMANN et al., 2010).

Os sintomas da pneumonia podem variar de leves a graves e independente de onde ela foi adquirida, geralmente incluem (AMERICAN LUNG ASSOCIATION, 2019):

- Tosse, que pode produzir muco esverdeado, amarelo ou até mesmo com sangue
- Febre, sudorese e calafrios
- Falta de ar
- Respiração rápida e superficial
- Dor no peito aguda ou penetrante que piora quando você respira profundamente ou tosse
- Perda de apetite, baixa energia e fadiga
- Náuseas e vômitos, especialmente em crianças pequenas
- Confusão, especialmente em pessoas idosas

Considerando o micro-organismo causador da infecção, a pneumonia pode ser classificada em pneumonia bacteriana, pneumonia viral ou pneumonia fúngica.

A pneumonia bacteriana é aquela causada por bactérias e geralmente não é transmissível. Seu tratamento é feito através de antibióticos. A pneumonia viral ou vírica é aquela provocada por vírus e ocorre predominantemente nos meses de inverno. Geralmente é contagiosa e o tratamento inclui apenas antitérmicos e analgésicos para aliviar os sintomas, mas, nas formas graves da doença, podem ser necessários medicamentos antivirais. Já a pneumonia fúngica é aquela causada por fungos, sendo o tipo mais raro da

doença. Esse tipo ocorre essencialmente em indivíduos imunodeficientes. O tratamento deve ser feito com medicamentos antifúngicos e dependendo do fungo causador da doença, esse tipo de pneumonia pode ser contagioso. (SETHI, 2019)

COVID-19 é o nome da doença causada pelo vírus SARS-CoV-2, conhecido por novo coronavírus. Existem sete tipos de coronavírus que causam doenças em humanos, sendo todos provenientes de animais. Três desses tipos de coronavírus causam infecção respiratórias graves em humanos, sendo responsáveis por grandes surtos de pneumonia fatal enquanto que os outros quatro são causas comuns de resfriados (NISHIOKA, 2020). O primeiro coronavírus patogênico foi o Sars-CoV, identificado em 2002 como causador da epidemia de Síndrome Respiratória Aguda Grave, do inglês *Severe Acute Respiratory Syndrome (SARS)*. Em 2012 foi identificado o MERS-CoV, o coronavírus de maior letalidade (até 34,4% dos infectados morriam), causador da Síndrome Respiratória do Oriente Médio - *Middle East respiratory syndrome (MERS)*. Em dezembro de 2019 o coronavírus SARS-CoV-2 foi identificado em Wuhan, na China, e embora seja menos letal até o momento, tem uma alta taxa de transmissibilidade, se espalhando rapidamente pelo mundo (TESINI, 2020).

A transmissão acontece de uma pessoa doente para outra ou por contato próximo por meio de gotículas de saliva, espirro, tosse, aperto de mãos e toque em objetos ou superfícies contaminadas, como celulares, mesas, talheres, maçanetas etc. Principalmente em ambientes fechados, o coronavírus é capaz de ficar suspenso no ar por algum tempo e, com isso, entrar em contato com a boca ou os olhos de qualquer pessoa (MINISTÉRIO DA SAÚDE, 2021).

Os sintomas da COVID-19 podem variar de um resfriado até uma pneumonia severa. Os sintomas mais comuns, mas não obrigatórios, são (MINISTÉRIO DA SAÚDE, 2021):

- Tosse
- Febre
- Coriza
- Dor de garganta
- Dificuldade para respirar
- Perda de olfato (anosmia)
- Alteração do paladar (ageusia)
- Distúrbios gastrintestinais (náuseas/vômitos/diarreia)

- Cansaço (astenia)
- Diminuição do apetite (hiporexia)
- Dispnéia (falta de ar)

Muitos fatores afetam a gravidade de um caso de pneumonia, como o tipo de agente causador da infecção pulmonar, a idade da pessoa e sua saúde geral. Diferente dos outros tipos de infecções respiratórias, a grande maioria das crianças com COVID-19 dificilmente evoluem para a forma grave da doença (WEISBERG et al., 2021). Apesar de ainda estar em debate, as principais questões que aumentam a probabilidade de agravamento do quadro de COVID-19 são: idade avançada, obesidade, diabetes, hipertensão, imunidade baixa, condições cardíacas graves, doenças pulmonares e doenças renais graves (VEJA SAÚDE, 2021).

A maioria dos casos de pneumonia pode ser tratada com sucesso, embora possa levar semanas para se recuperar completamente. Para a COVID-19, até o momento, não existe nenhum medicamento que evite a infecção ou que impeça o agravamento dos casos leves.

Diagnosticar a pneumonia é um trabalho realizado por médicos através de um exame clínico e geralmente requer o auxílio de exames laboratoriais e de imagem, junto com o histórico médico do paciente. Por ser um exame amplamente acessível, a radiografia do tórax é normalmente solicitada para confirmação do diagnóstico de pneumonia.

3.2 Radiografia do Tórax

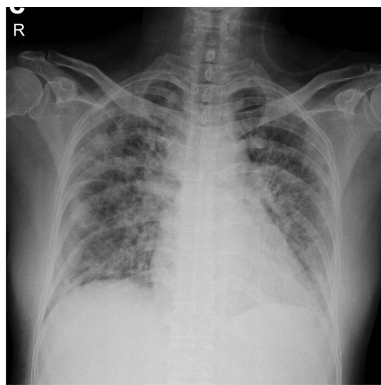
A radiografia (ou raio-X) é um exame de imagem que pode ser usado em diversas partes do corpo com funções diagnósticas distintas. Esse exame é feito usando radiação ionizante, que utiliza os raios X para gerar uma imagem médica de diferentes partes do corpo humano (SAUDE BEM ESTAR, 2011).

Devido a tecnologia simples, a radiografia é o método de imagem mais barato que existe, estando amplamente disponível nos centros de saúde. O exame é não invasivo, sendo realizado por um técnico treinado. O tempo de realização é curto e diversas partes do corpo podem ser radiografadas em uma única ocasião (SOCIEDADE BRASILEIRA DE REUMATOLOGIA, 2011). Em uma radiografia, quanto mais denso for o tecido, mais ele

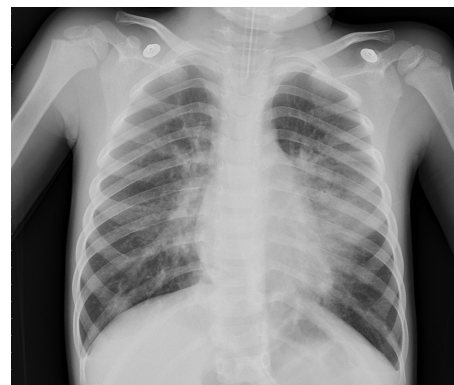
absorve os raios X. Os ossos, portanto, aparecem brancos nas imagens, pois são muito densos. Já partes moles, como órgãos, são mostradas mais escuras.

Em uma radiografia do tórax, os alvéolos cheios de secreção aparecem como uma mancha branca em lugares que deveriam ser mais escuros. Isso acontece devido ao fato dos raios X não conseguirem atravessar o pulmão por completo, sendo em parte absorvido pelos líquidos acumulados no pulmão (PARVEENA, 2011). Em casos de COVID-19, os principais achados na radiografia são opacidades em vidro fosco, consolidações e opacidades lineares multifocais e bilaterais, em graus variados de extensão de acometimento do parênquima dependendo da gravidade do paciente e tempo de evolução da doença (CAPONE et al., 2020). Exemplos de radiografias de pacientes afetados pela pneumonia podem ser vistos na Figura 3.

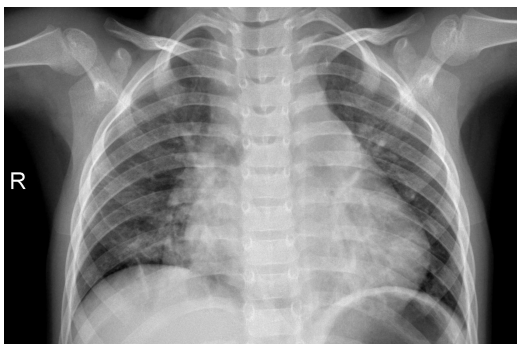
Figura 3 – Radiografias: Normal x Pneumonia x COVID-19



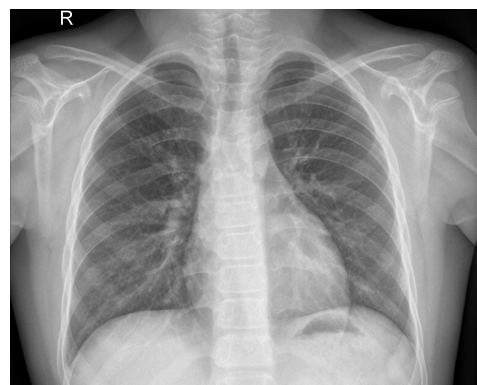
(a) COVID-19



(b) Pneumonia viral



(c) Pneumonia bacteriana

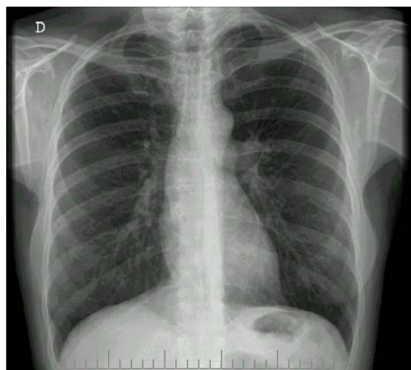


(d) Normal

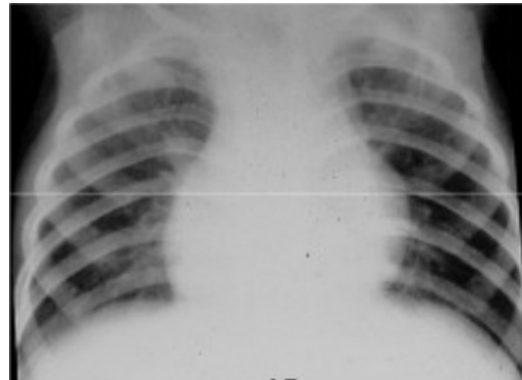
Na radiologia, incidência é o nome dado à trajetória percorrida pelo feixe de raios X durante o exame, indicando por onde a radiação penetra no corpo e por onde ela sai. Assim, médicos podem recomendar duas ou mais incidências, de forma que os resultados mostrem perspectivas diferentes dos mesmos tecidos (MORSCH, 2019). As incidências

básicas quando se trata de raio-X do tórax são: incidência póstero-anterior (PA), quando os raios entram pela região posterior das costas e saem pela anterior; incidência antero-posterior (AP), quando a radiação penetra o organismo pela parte anterior e sai pela posterior; e a incidência de perfil, sendo que rotineiramente realiza-se o perfil esquerdo, deixando o lado esquerdo em contato com o filme e com entrada do feixe pela direita, para não magnificar o coração (WADA; RODRIGUES; SANTOS, 2019). Esses tipos de incidência podem ser observados na Figura 4. Além dessas incidências citadas, pode-se utilizar outras incidências especiais, apesar de menos frequentes: PA em decúbito dorsal, decúbito lateral, lordótica, oblíqua anterior e oblíqua posterior.

Figura 4 – Tipos de incidência mais realizados em radiografias do tórax.



(a) Póstero-anterior (PA)



(b) Antero-posterior (AP)



(c) Perfil

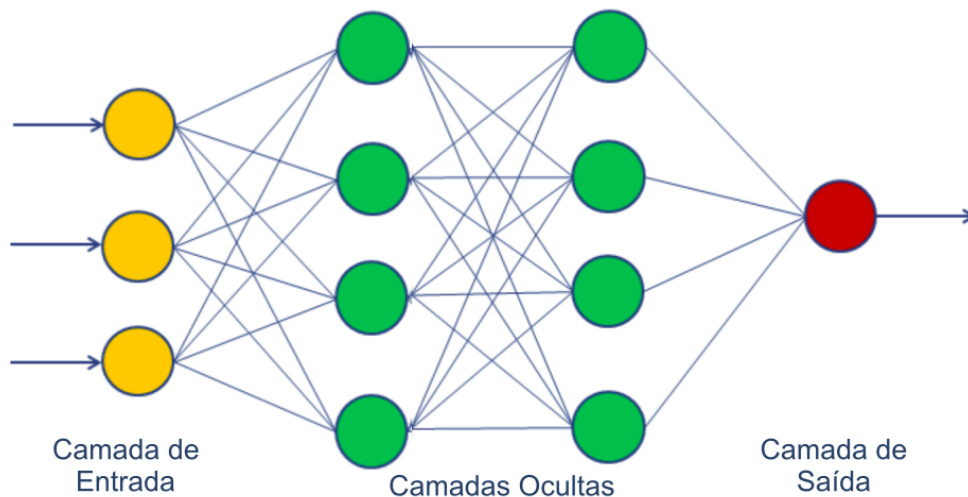
A radiação emitida pelos aparelhos – principalmente os mais modernos – é segura, desde que a pessoa não se submeta ao raio-X a todo momento. Em radiografias do tórax, a quantidade de exposição à radiação com uma única imagem é semelhante à quantidade que a maioria das pessoas recebe do ambiente em 2,4 dias (KOCAK, 2019).

A análise precisa de uma radiografia requer bastante observação e experiência do especialista e esse fato pode resultar em diagnósticos diferentes para a mesma imagem.

3.3 Aprendizado Profundo

O Aprendizado Profundo é um subcampo de aprendizado de máquina que consiste principalmente em algoritmos de redes neurais artificiais (RNA), pois são inspirados na estrutura e funcionamento do cérebro. O termo profundo diz respeito à estrutura da rede neural artificial, que é construída em camadas (NIELSEN, 2018). Normalmente essas camadas são separadas entre camada de entrada, camada de saída e camada(s) escondida(s) (que ficam entre a entrada e a saída). Quanto mais camadas uma RNA possuir, mais profunda ela é. Essa estrutura pode ser visualizada na representação da Figura 5. Cada círculo representa um neurônio, e um conjunto de neurônios compõe uma camada.

Figura 5 – Representação da estrutura de uma rede neural artificial.

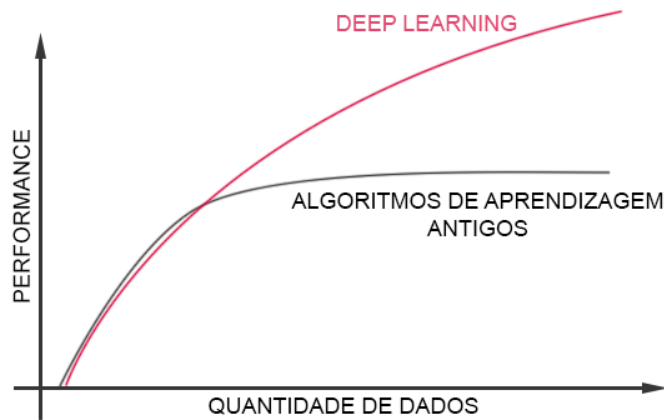


Fonte: Adaptada de Nielsen (2018).

O uso de RNAs para realizar tarefas de Aprendizado de Máquina se popularizou com o crescimento expressivo de dados disponíveis. Os modelos anteriores às RNAs lidavam com conjuntos pequenos de dados, logo, com o crescimento ocorrido nas últimas décadas, novos modelos eram necessários para processar o grande volume de dados existente. Outro fator que contribuiu para o crescimento de uso das RNAs foi o aumento do poder computacional, principalmente pelo uso de placas gráficas (*GPUs*) para realizar o processamento dos dados de entrada. Essa mudança de paradigma é descrita no gráfico da Figura 6.

Outro grande benefício das RNAs é a sua habilidade de realizar a extração automática de características dos dados (*Feature Learning*) (BENGIO, 2012). Em um outro trabalho de Bengio et al. (2009), a extração de características é explicada: métodos de

Figura 6 – Gráfico comparativo de Deep Learning versus métodos anteriores de aprendizado.



Fonte: Adaptada de Ng (2015).

aprendizagem profunda aprendem as características de forma hierárquica; as características são extraídas dos níveis mais altos até os níveis mais baixos. Dessa forma, as RNAs aprendem tais características em diferentes níveis de abstração, permitindo ao sistema aprender funções complexas sem depender de características extraídas de forma manual.

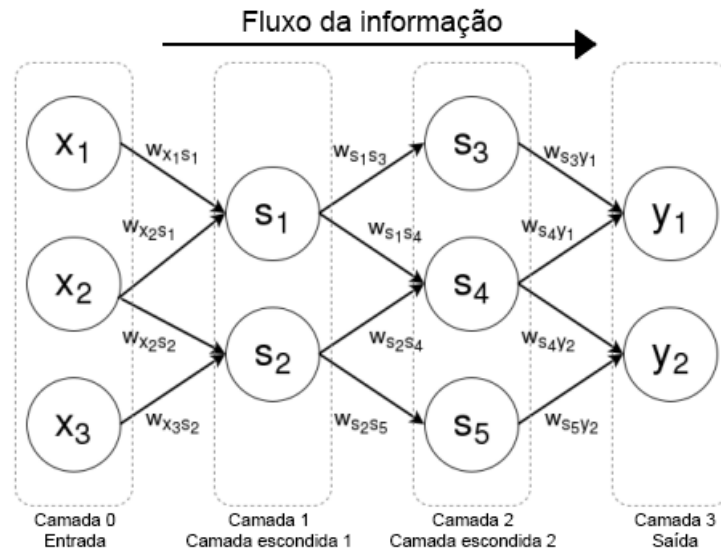
Várias arquiteturas de aprendizagem profunda, tais como Redes Neurais Profundas (DNNs - do inglês *Deep Neural Networks*), Redes Neurais Convolucionais (CNNs - do inglês *Convolutional Neural Networks*), Redes de Crenças Profundas (DBN - do inglês *Deep Belief Network*) e Redes Neurais Recorrentes Profundas (DRNNs - do inglês *Deep Recurrent Neural Networks*) têm sido aplicadas em áreas como visão computacional, reconhecimento automático de fala, processamento de linguagem natural, reconhecimento de áudio e bioinformática, onde elas têm se mostrado capazes de produzir resultados do estado-da-arte em várias tarefas (AHMAD; FARMAN; JAN, 2019).

3.3.1 *Feedforward e Backpropagation*

Feedforward e *Backpropagation* são dois conceitos importantes para melhor compreensão de como as Redes Neurais Profundas (RNP) realizam o aprendizado de máquina. Alguns autores chamam as RNPs de *Feedforward Neural Networks*, porque os dados seguem na mesma direção por todas as camadas da rede, nas quais diversas computações são feitas sobre esses dados até chegar na última camada retornando uma

saída. A Figura 7 exemplifica o fluxo de dados em uma RNP: Os dados fluem da entrada $[x_1, x_2, x_3]$ para a saída $[y_1, y_2]$, passando pelos neurônios $[s_1, s_2, s_3, s_4, s_5]$ das camadas centrais.

Figura 7 – Fluxo de dados em uma RNP.



Fonte: Adaptada de Brilliant.org (2018).

Cada ligação entre neurônios possui um peso associado. Cada neurônio recebe a informação dos neurônios que possui ligação na camada anterior, e processa esses dados junto ao peso da ligação. O resultado do cálculo obtido por esse neurônio é redirecionado aos neurônios que possui ligação na camada seguinte. Esse procedimento é feito por cada neurônio a partir da primeira camada escondida até a camada de saída, na qual o resultado é comparado à saída esperada por meio do cálculo do erro. O objetivo de uma RNP é diminuir esse erro, de forma a obter na saída o resultado mais próximo da classificação desejada. Esse procedimento é chamado de *Feedforward*.

Após a propagação do sinal camada a camada, as respostas produzidas pelas saídas da rede são comparadas com as respectivas respostas desejadas que estejam disponíveis. Os erros entre as respostas desejadas e aquelas produzidas pelos neurônios de saída são então calculados e, subsequentemente utilizados para ajustar os pesos de todos os neurônios (NUNES, 2016). Esse processo de retorno do erro para correção dos pesos é chamado de *Backpropagation*. O *Feedforward* é o processo de transferência dos dados à frente, enquanto o *Backpropagation* faz o caminho inverso, realizando a correção dos pesos.

Os dados de entrada são processados pela RNP por um número finito de vezes (quantidade de épocas), até se alcançar uma acurácia desejada ou a taxa de erro ser igual ou menor a uma taxa mínima definida.

3.3.2 Redes Neurais Convolucionais

Redes Neurais Convolucionais (LECUN et al., 1989) são um tipo especializado de rede neural projetadas para processar dados que estão na forma de múltiplos *arrays*. Alguns exemplos desses dados são: sinais e sequências (uma dimensão), imagem ou áudio (duas dimensões), vídeos ou imagens volumétricas (três dimensões). A imagem é o tipo de dado mais comum utilizado como entrada em uma CNN.

O termo “Rede Neural Convolutiva” indica que a rede faz uso de uma operação matemática chamada convolução. A convolução é um tipo específico de operação linear, utilizada no lugar das multiplicações de matrizes que são realizadas em uma Rede Neural comum.

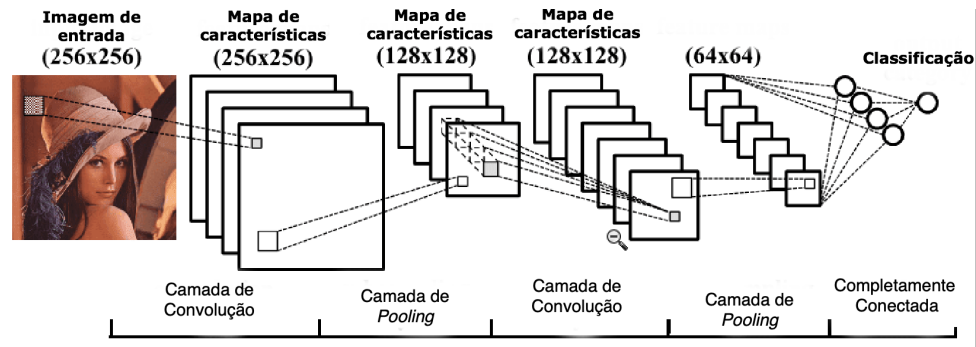
A arquitetura de uma CNN é estruturada como uma série de estágios, assim como a distribuição em camadas de uma RNP. Os primeiros estágios geralmente são compostos de dois tipos de camadas: camadas convolucionais e camadas *pooling*. O objetivo da camada convolutiva é detectar conjuntos de características dos dados recebidos da camada imediatamente anterior a ela, enquanto que a camada *pooling* une as características detectadas que são semanticamente similares, realizando uma sub-amostragem dessas características.

As camadas são empilhadas de forma que a saída processada por ela serve como entrada para a outra. A Figura 8 ilustra uma arquitetura geral de uma rede neural convolutiva.

3.3.2.1 Camada de Convolução

Uma rede neural aprende características de cada classe a partir da correção dos pesos associados às ligações entre neurônios. No caso das CNNs, os pesos da rede estão associados nas camadas convolucionais por meio de conjuntos de filtros. Cada filtro é passado sobre a imagem recebida da camada anterior, resultando em um mapa de ativação de características daquele filtro. Intuitivamente, a rede irá aprender filtros que são ativados

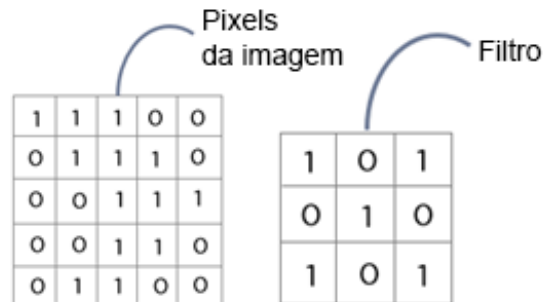
Figura 8 – Arquitetura simples de uma CNN.



Fonte: Adaptada de LeCun et al. (1998).

quando percebem algum tipo de característica visual na imagem como bordas, cores e padrões. Na Figura 9 é possível visualizar uma imagem de entrada e um filtro de ativação em forma de “X” da camada convolucional.

Figura 9 – Esquerda: imagem de entrada em uma CNN. Direita: exemplo de filtro de ativação.



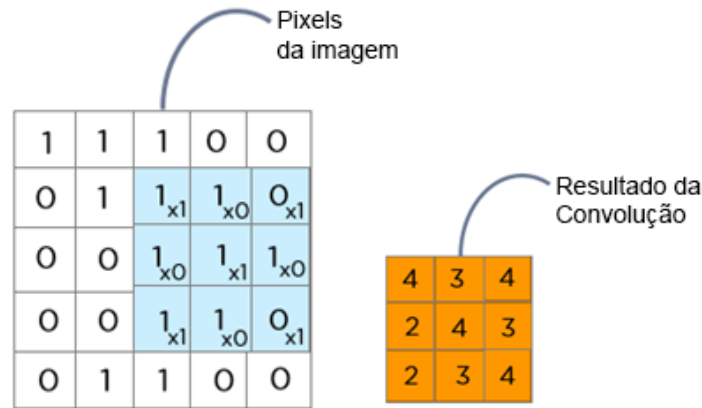
Fonte: Adaptada de Sharma (2018).

Como pode ser observado na Figura 10, a convolução é o resultado de uma soma ponderada local entre o filtro e a imagem, após esta ter sido percorrida totalmente pelo filtro. Definir o tamanho do filtro (*kernel*) menor que o tamanho da imagem permite que haja economia computacional, dado que apenas computações locais limitadas à área do filtro são feitas. A convolução pode ser definida pela Equação 1:

$$Y_k = f(W_k * x), \quad (1)$$

onde a imagem de entrada é dada por x ; o *kernel* de convolução relacionado ao k -ésimo filtro da camada é denotado por W_k ; o sinal $*$, neste contexto, significa o operador convolucional 2D usado para calcular o produto interno do modelo de filtro em cada

Figura 10 – Processo de convolução realizado pelo filtro (azul) sobre a imagem, resultando no mapa de características (laranja).



Fonte: Adaptada de Sharma (2018).

localização da matriz de entrada; e $f(x)$ representa a função de ativação não-linear aplicada ao resultado do produto (LECUN et al., 1998) e (RAWAT; WANG, 2017).

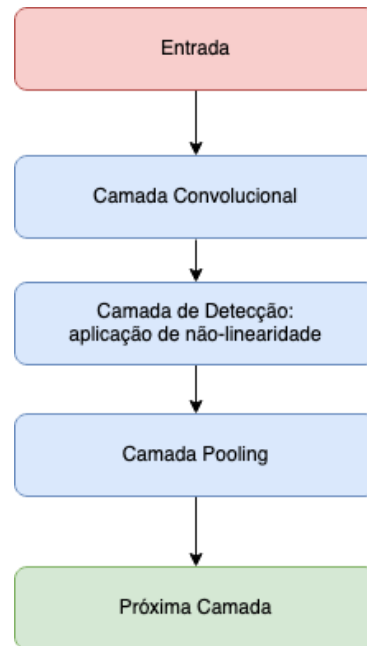
3.3.2.2 Camada de *Pooling*

Uma CNN é tipicamente composta de três estágios (Figura 11). No primeiro estágio diversas convoluções são realizadas para produzir um conjunto de ativações lineares. Em seguida, cada ativação linear é passada por uma função de ativação não-linear, como por exemplo a ativação *Rectified Linear Unit (ReLU)*. Por último, a rede faz uso de uma função *pooling* para modificar a saída da camada.

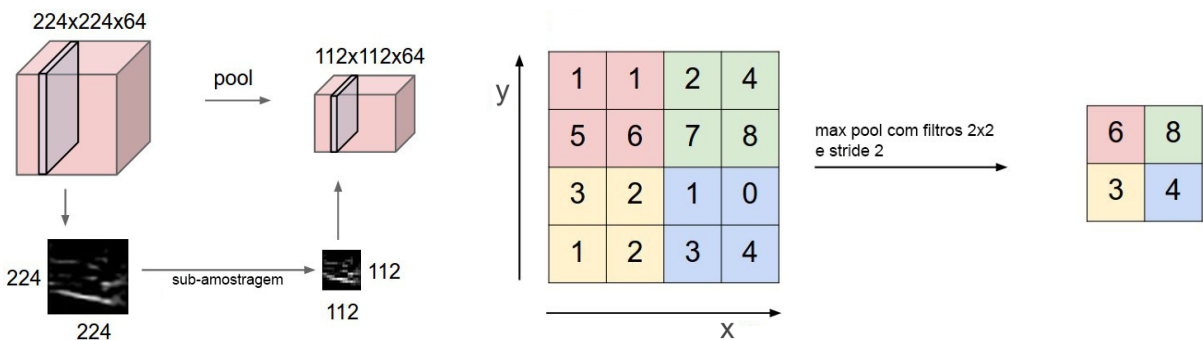
Uma função *pooling* substitui a saída da rede em certo local por uma representação estatística. Por exemplo, a função *max pooling* (ZHOU et al., 1988) retorna uma saída máxima em uma vizinhança retangular. Outra função *pooling* conhecida é *average pooling* que retorna a média de uma vizinhança retangular.

A função *max pooling* é a mais comum de ser utilizada em arquiteturas de CNN, e pode ser vista na Figura 12. É possível notar que ao final, uma sub-amostragem da imagem dada como entrada é retornada pela operação *pooling*. É importante mencionar que a quantidade de dados devolvida após a operação é significativamente reduzida sem grandes perdas, tornando viável a utilização dessa operação para realização de sub-amostragem dos mapas de ativação.

Figura 11 – Estágios de uma CNN típica.



Fonte: Adaptada de Goodfellow, Bengio e Courville (2016).

Figura 12 – Demonstração da operação *max pooling*.

Fonte: Adaptada de Vision e Lab (2018).

3.3.2.3 Camada de Classificação

Por último, geralmente, no final das redes neurais convolucionais, são acopladas camadas completamente conectadas para efetuar a classificação das características extraídas pelas camadas anteriores. Neste momento, definem-se as funções de custo a serem utilizadas, que tem como objetivo representar o erro entre o valor esperado e a predição dada pela rede neural. Esse erro é usado no processo de retro-propagação para ajuste de pesos.

As camadas empilhadas anteriormente a essa camada têm como função extrair características significativas da imagem de entrada (LECUN et al., 1998). No final, essas

características são vetorizadas e servem de entrada para uma rede neural artificial completamente conectada, cuja tarefa é categorizar os descritores informados em sua entrada.

Com a classificação dada pela camada completamente conectada a partir de uma imagem informada, necessita-se calcular o erro entre a categorização e o valor verdadeiro do rótulo da entrada. Geralmente, usa-se o operador *softmax* devido à sua simplicidade e interpretação probabilística, que serve para mapear a saída não normalizada da rede para uma distribuição de probabilidade sobre as classes de saída previstas (NASRABADI, 2007). Formalmente, para o i -ésimo vetor de características x_i , que tem classe rotulada por y_i , a função de custo *softmax* (L) pode ser escrita como:

$$L = \frac{1}{N} \sum_i -\log \left(\frac{e^{f_{y_i}}}{\sum_j e^{f_j}} \right), \quad (2)$$

onde o elemento j -ésimo ($j \in [1, K]$, K é o número de classes) do vetor de probabilidades de classe f é representado por f_j e N é o tamanho da base de treinamento. Para o custo, f representa as ativações de uma camada W totalmente conectada; assim, f_{y_i} , pode ser denotado como $f_{y_i} = W_{y_i}^T x_i$ em que W_{y_i} é a y_i -ésima coluna de W .

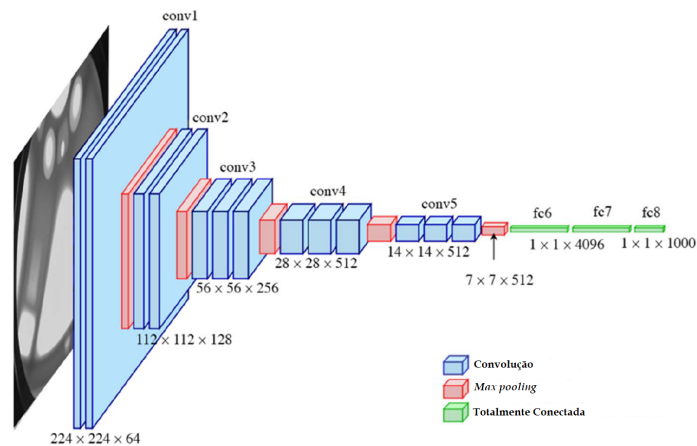
Uma técnica importante presente em algumas redes neurais é o *dropout* (SRIVASTAVA et al., 2014). Esse tipo de técnica é aplicada somente na fase de treino da rede, objetivando prevenir o *overfitting* de neurônios nas redes neurais artificiais. A ideia é simples e eficaz. A cada época de treino ou *batch* treinado, um sorteio sobre os neurônios envolvidos na RNA é feito e alguns são selecionados com probabilidade p e excluídos da próxima interação .

De acordo com a literatura, a CNN tem se mostrado o modelo ideal para todos os problemas relacionados à imagem. Devido a sua ampla popularidade, nos últimos anos surgiram várias arquiteturas de CNN : LeNet5 (LECUN et al., 1998), AlexNet (KRIZHEVSKY; SUTSKEVER; HINTON, 2012), Inception (SZEGEDY et al., 2014), ResNet (HE et al., 2015), Xception (CHOLLET, 2017), ResNeXt (XIE et al., 2017), entre outras. Por serem utilizadas como rede base nesta proposta, as redes VGG, ResNet e DenseNet serão mais detalhadas na próxima seção.

3.3.2.4 Rede Convolutacional VGG

A rede convolutacional VGG é composta por camadas convolucionais e camadas completamente conectadas. Entre cada bloco de camada convolutacional existem camadas pooling (SIMONYAN; ZISSERMAN, 2014). O nome VGG é devido a seus idealizadores, os pesquisadores do Visual Geometry Group, da Universidade de Oxford. A VGG possui duas variações: a VGG-16, mais popular, e a VGG-19. A diferença entre as variações é o número de camadas. Por exemplo, a VGG-16 é uma rede que possui 16 camadas de convolução. A Figura 13 ilustra a arquitetura da VGG16.

Figura 13 – Arquitetura VGG-16



Fonte: Adaptada de Zhou et al. (2014).

Para fazer o reconhecimento da imagem, a VGG-16 utiliza as camadas de convolução para retirar as características das imagens e as outras camadas totalmente conectadas para fazer as classificações. Devido a sua arquitetura profunda e uniforme, a VGG-16 tem sido muito utilizada como extrator de características-base (WANGENHEIM, 2018).

A VGG-16 tem tido resultados promissores em problemas médicos por ter características como: modelo sequencial de blocos de convolução, tamanhos de filtros uniformes em 3×3 e quantidade pequena de camadas. Essas características a tornam adequada para ser reutilizada para abordagem médica no que tange a capacidade do método em aprender corretamente os padrões necessários, mesmo que a base de dados não seja suficientemente grande.

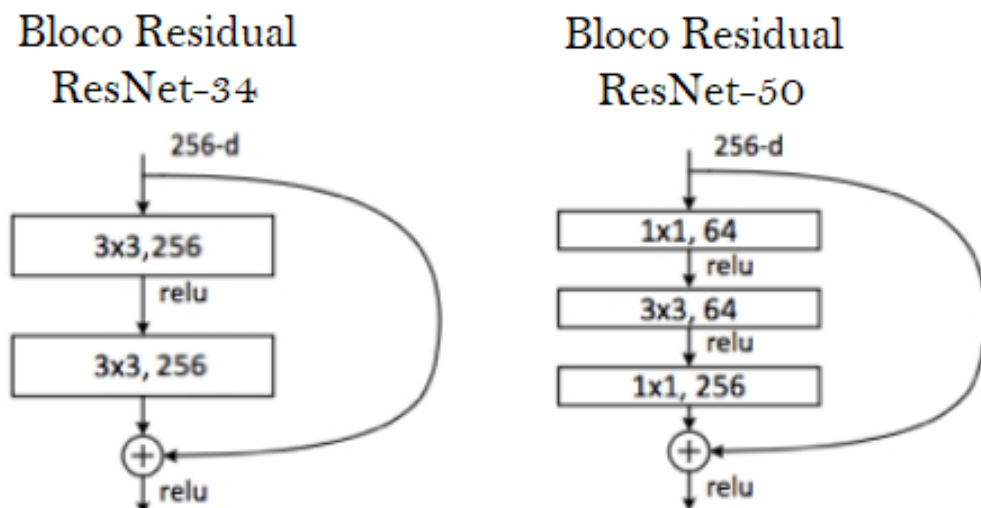
3.3.2.5 Rede Convolutacional ResNet

A rede Convolutacional Residual (ResNet) (HE et al., 2015) introduziu a aprendizagem residual: uma estrutura de aprendizagem pensada para facilitar o treinamento das redes que são substancialmente mais profundas.

Em teoria, à medida que o número de camadas em uma rede neural aumenta, ela deve ficar cada vez melhor no reconhecimento de funções e recursos complexos, resultando em melhor precisão e aprendizado. Porém, observou-se que a adição de camadas foi ineficiente em fornecer o resultado esperado, apresentando uma queda na precisão do treinamento a partir de certo ponto (ICHL.PRO, 2020).

Para contornar esse problema, as redes residuais utilizam os blocos residuais, que são conexões de salto (*skip connections* ou *shortcut connections*) que pulam pares de grupos de camadas convolucionais. A Figura 14 ilustra exemplos de bloco residual. Em um determinado ponto, a ResNet vai utilizar um sinal que é a soma do sinal produzido pelas duas camadas convolucionais anteriores somado ao sinal transmitido diretamente do ponto anterior a estas camadas, juntando um sinal processado com um sinal de uma etapa anterior no processamento (WANGENHEIM, 2018).

Figura 14 – Bloco Residual



Fonte: Adaptada de Jordan (2018).

A arquitetura da ResNet, ilustrada na Figura 15, é formada por vários blocos residuais empilhados. A ResNet possui as seguintes variações: ResNet-18, ResNet-34,

ResNet-50, ResNet-101 e ResNet-152. Uma ResNet-50 significa que a rede possui cinquenta camadas usando esses blocos.

Figura 15 – Arquitetura ResNet-34



Fonte: Adaptada de He et al. (2015).

Em ResNets mais profundas, como a ResNet-50, ResNet-101 ou ResNet-152, é utilizada uma variante do bloco residual, chamado de bloco residual de gargalo (*Bottleneck Residual Block*), que utiliza convoluções 1x1 para criar um gargalo. O uso de um gargalo reduz o número de parâmetros e multiplicações de matrizes. A ideia é fazer blocos residuais tão finos quanto possível para aumentar a profundidade e ter menos parâmetros (SIK-HO TSANG, 2018).

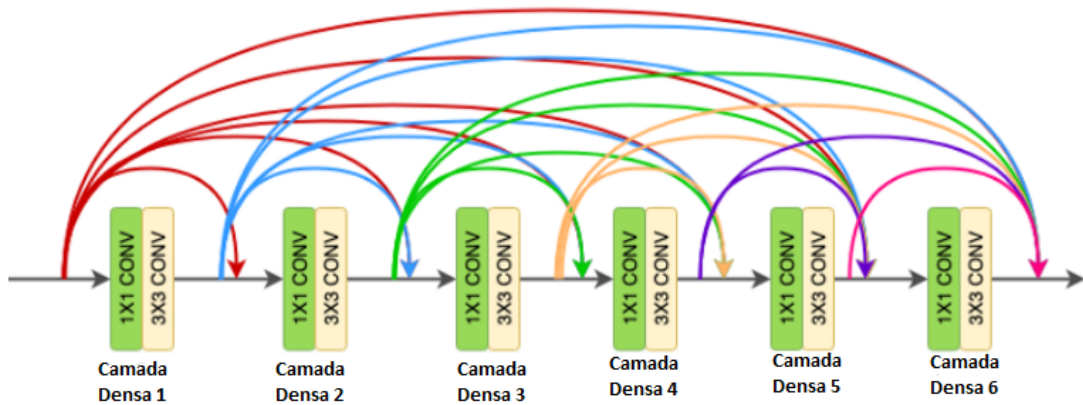
3.3.2.6 Rede Convolutiva DenseNet

Assim como a ResNet, a rede convolutiva Densa (DenseNet) (HUANG et al., 2017) introduz um novo módulo para facilitar o treinamento em redes neurais profundas: os blocos densos.

Um bloco denso, que pode ser observado na Figura 16, é um módulo que conecta todas as camadas diretamente entre si. Um bloco denso compreende “n” camadas densas. Cada camada densa consiste em duas operações de convolução: uma convolução 1x1 seguida por uma convolução 3x3. Essas camadas densas são conectadas de modo que cada camada densa receba mapas de características de todas as camadas anteriores e passe seus mapas de características para todas as camadas subsequentes (KHANNA, 2019). Os mapas de

características recebidos de outras camadas são agregados por meio de concatenação, e não por meio de soma, como em ResNets. Essas conexões reduzem o número de parâmetros, uma vez que estão reutilizando informações de mapas de características anteriores.

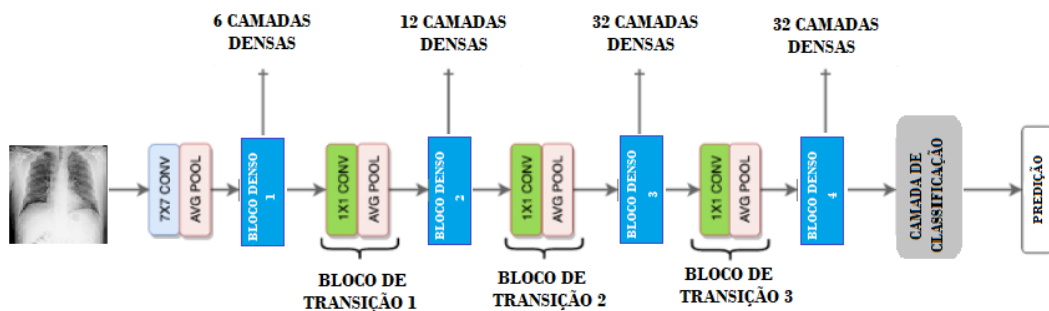
Figura 16 – Bloco Denso



Fonte: Adaptada de Khanna (2019).

Originalmente, a DenseNet é composta por 4 blocos densos, com camadas de transição, que fazem convolução e pooling, e a camada de classificação. A DenseNet possui as seguintes variações: DenseNet-121, DenseNet-169, DenseNet-201 e DenseNet-264. A única diferença entre as versões é a profundidade da rede, definida pelo tamanho dos blocos densos.

Figura 17 – Arquitetura DenseNet-169



Fonte: Adaptada de Khanna (2019).

3.3.3 Hiperparâmetros

No contexto de redes neurais artificiais, o treinamento corresponde ao processo de ajuste dos parâmetros da rede através de um mecanismo de apresentação de estímulos

ambientais, conhecidos como dados de entrada ou de treinamento. Os parâmetros são essenciais para algoritmos de aprendizado de máquina, sendo a parte do modelo que é aprendida com os dados de treinamento.

Um modelo treinado é uma função matemática particular, pertencente a um certo tipo de algoritmo de aprendizado de máquina, que foi determinado por uma tupla particular de parâmetros. Os parâmetros que permitem a customização da função são os parâmetros do modelo ou simplesmente parâmetros e são exatamente o que a máquina vai aprender com os dados. No decorrer do treinamento, os parâmetros do modelo são ajustados automaticamente (APORRAS, 2020).

Porém, no processo de construção de um modelo treinado, mais parâmetros são necessários para definir como o algoritmo fará isso. Em aprendizado de máquina, usamos hiperparâmetros para denotar esse tipo específico de parâmetro. Os hiperparâmetros não podem ser aprendidos usando o algoritmo que precisa deles, mas devem ser ajustados antes do estágio de treinamento, manual ou automaticamente (APORRAS, 2020). Os hiperparâmetros são utilizados para configurar diversos aspectos do algoritmo de aprendizagem e podem ter efeitos variados no modelo resultante e na sua performance (CLAESEN; MOOR, 2015).

Em resumo, os parâmetros do modelo são estimados a partir dos dados automaticamente. Já os hiperparâmetros do modelo são ajustados antes do treinamento e são usados para ajudar a estimar os parâmetros do modelo.

A Tabela 3 lista as principais diferenças entre parâmetros e hiperparâmetros de um modelo (GUPTA, 2020).

Tabela 3 – Parâmetros x Hiperparâmetros

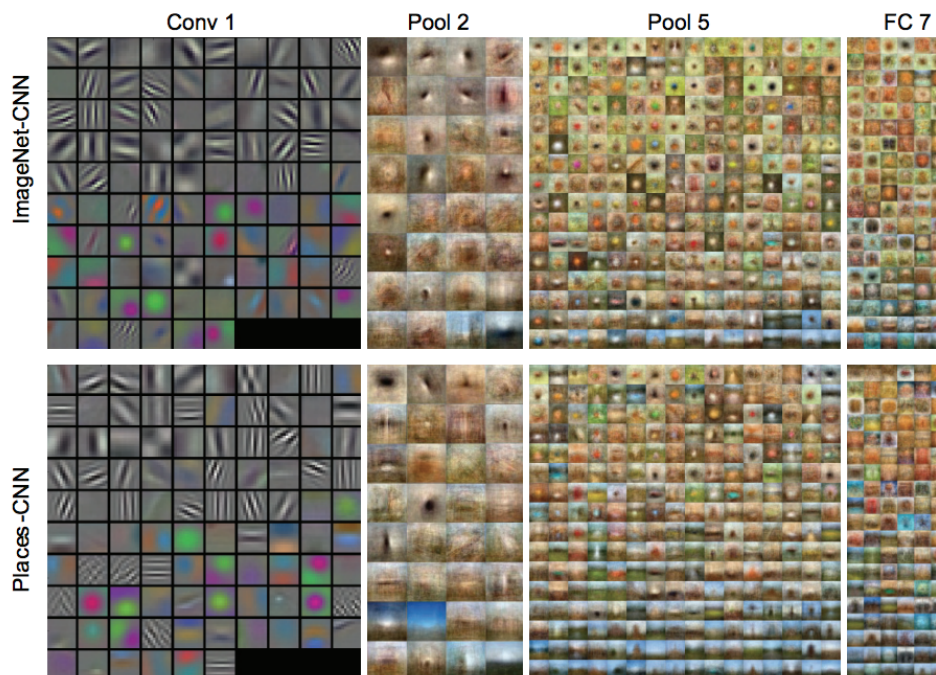
Parâmetros	Hiperparâmetros
São necessários para fazer previsões.	São necessários para estimar os parâmetros do modelo.
São estimados por algoritmos de otimização (Gradiente Descente, Adam, etc).	São estimados por ajuste de hiperparâmetro.
Não são setados manualmente.	São setados manualmente ou automaticamente.
Os parâmetros finais encontrados após o treinamento decidirão como o modelo será executado em dados não vistos.	A escolha dos hiperparâmetros decide o quão eficiente é o treinamento.

3.3.4 *Transfer Learning*

Redes neurais, especialmente redes neurais convolucionais, possuem uma alta capacidade discriminativa. O poder de modelagem não linear faz dessas redes não apenas um bom classificador (MIRUNALINI et al., 2017), mas também um bom extrator de características (YU et al., 2017).

Na Figura 18 são demonstradas as características de CNNs pré-treinadas para tarefas de reconhecimento visual. Cada camada exemplifica a visualização dos campos receptivos dos diferentes filtros aplicados, para duas arquiteturas de CNNs diferentes. As respostas das camadas de nível mais alto das CNNs provaram ser recursos genéricos eficazes com desempenho de última geração em vários conjuntos de dados de imagens (ZHOU et al., 2014). Pode-se notar que as características passam de simples detectores de contornos e bordas para representação de textura e forma mais complexas de acordo com o avanço da imagem através da rede.

Figura 18 – Visualização dos campos receptivos dos filtros.



Fonte: Zhou et al. (2014).

Essas características extraídas por meio de arquiteturas de CNNs pré-treinadas podem ser utilizadas não só em uma rede neural artificial simples completamente conectada, mas também para treinar outros classificadores como Regressão Logística e *Random Forest*.

CNNs são o estado-da-arte para classificação de imagens (MENEGOLA et al., 2017), mas seu uso para imagens médicas é desafiador, já que esses modelos exigem conjuntos de treinamento muito grandes (de dezenas de milhares até milhões de imagens) (SHIN et al., 2016). Para contornar essa dificuldade, a literatura mais atual emprega o chamado *Transfer Learning*, uma técnica em que um modelo treinado para uma determinada tarefa de origem é parcialmente “reciclado” para uma nova tarefa de destino. *Transfer Learning* varia de simplesmente usar a saída de uma CNN de origem como um vetor de características e treinar um modelo completamente novo (por exemplo, um SVM) para uma determinada tarefa; até usar uma CNN inicial pré-treinada para inicializar alguns dos pesos de uma CNN de destino e continuar o processo de treinamento normalmente (YOSINSKI et al., 2014).

3.4 Meta-aprendizagem

A aprendizagem pode ser definida como o ato de aprender ou adquirir conhecimento através da experiência ou de um método de ensino (PEREIRA, 2015). O termo meta-aprendizagem (do inglês *meta-learning*) se refere a capacidade de “aprender a aprender”.

De acordo com (THRUN; PRATT, 1998), um algoritmo é capaz de aprender se o seu desempenho em uma tarefa melhora com a experiência e, dado

1. um conjunto de tarefas
2. um treinamento de experiência para cada uma dessas tarefas e
3. um conjunto de medidas de desempenho (uma medida para cada tarefa),

um algoritmo é capaz de “aprender a aprender” se o seu desempenho em cada tarefa melhora com a experiência e com o número de tarefas.

O meta-aprendizado difere do aprendizado básico no escopo do nível de adaptação. Enquanto o aprendizado básico é baseado no acúmulo de experiência em uma tarefa específica, a aprendizagem em nível meta é baseada na acumulação de experiência no desempenho de múltiplas aplicações. Meta-aprendizagem é, então, importante para compreender a interação entre o mecanismo de aprendizagem e os contextos em que esse mecanismo é aplicável (VILALTA; GIRAUD-CARRIER; BRAZDIL, 2009).

Um conceito frequentemente usado em meta-aprendizado é o *bias*, que, neste contexto, refere-se a um conjunto de suposições que influenciam a escolha das hipóteses. O

bias no aprendizado básico é fixo, enquanto que no meta-aprendizado a escolha do *bias* certo é feita dinamicamente (BRAZDIL; SOARES, 2000).

O trabalho de Lemke, Budka e Gabrys (2015) enumera dois requisitos que definem um sistema de meta-aprendizagem:

1. Precisa incluir um subsistema de aprendizagem.
2. A experiência é obtida pela exploração da metainformação extraída:
 - em um episódio anterior de aprendizagem em um único dataset e/ou
 - de diferentes domínios ou problemas.

A ideia de meta-aprendizagem não é nova. Uma das primeiras contribuições foi feita por Rice (1976), na década de 70. Contudo, somente em 1990 o termo começou a aparecer na literatura do aprendizado de máquina.

A meta-aprendizagem no aprendizado de máquina geralmente se refere aos algoritmos de aprendizado de máquina que aprendem a partir da saída de outros algoritmos de aprendizado de máquina. Se em aprendizado de máquina o algoritmo aprende como usar melhor as informações nos dados para fazer previsões, na meta-aprendizagem o algoritmo aprende como usar melhor as previsões dos algoritmos de aprendizado de máquina para fazer previsões (BROWNLEE, 2020).

A meta-aprendizagem pode ajudar a reduzir a quantidade de experimentação ao fornecer conselhos dinâmicos na forma de assistentes, diminuir o tempo que deve ser gasto em ajuste e manutenção de modelos e ajuda a promover o aprendizado de máquina fora de um ambiente acadêmico (LEMKE; BUDKA; GABRYS, 2015).

3.4.1 Otimização Bayesiana

A otimização bayesiana é uma abordagem baseada em modelo sequencial (do inglês *Sequential Model-Based Optimization* - SMBO) para otimizar funções objetivo que são muito custosas ou demoradas para avaliar (FRAZIER, 2018). Esse método é especialmente útil quando tais avaliações são caras, quando não se tem acesso a derivadas ou quando o assunto em questão não é convexo (BROCHU; CORA; FREITAS, 2010).

As estatísticas bayesianas vinculam a chance a algum grau de credibilidade ou incerteza sobre um caso, previsão ou interesse aleatório. Esta abordagem permite a determinação de probabilidades *a priori*, que resultam de informações passadas de uma

entidade, sem o conhecimento de qualquer outro evento, e probabilidades *a posteriori*, que é a probabilidade condicionada a algum outro evento conhecido.

Na otimização bayesiana, *priori* representa nossa crença sobre o espaço de funções objetivo possíveis. Embora a função de custo seja desconhecida, é razoável supor que exista conhecimento prévio sobre algumas de suas propriedades, e isso torna algumas funções objetivo possíveis mais plausíveis do que outras (BROCHU, 2010).

Tradicionalmente, a busca por hiperparâmetros adequados é feita manualmente (HINTON, 2012) (HSU; CHANG; LIN, 2003) ou testando conjuntos de hiperparâmetros em uma grade de opções predefinidas (PEDREGOSA et al., 2011). No entanto, tais abordagens são insuficientes em termos de reprodução e são impraticáveis quando o número de hiperparâmetros é extenso (CLAESEN; MOOR, 2015). Por causa dessas falhas, a ideia de automatizar a pesquisa de hiperparâmetros tem recebido muita atenção recentemente. Os processos de busca automatizada já demonstraram melhor capacidade quando comparados à busca manual para vários problemas (BERGSTRA et al., 2011), (BERGSTRA; BENGIO, 2012).

O Método SMBO difere de métodos como *Grid Search* e *Random Search* porque considera em sua pesquisa desempenhos anteriores dos hiperparâmetros, enquanto que nos outros dois métodos a pesquisa é independente das avaliações anteriores. Os métodos SMBO funcionam procurando o próximo conjunto de hiperparâmetros a ser avaliado na função objetivo, selecionando os hiperparâmetros que se destacam em uma função probabilística substituta, que é menos dispendiosa de avaliar. Se os valores avaliados também apresentarem resultados promissores na função objetivo, eles serão incorporados ao conjunto dos melhores hiperparâmetros.

Uma etapa importante do método SMBO são os critérios pelos quais o próximo conjunto de hiperparâmetros é escolhido a partir da função substituta. A escolha mais comum de critérios é o *Expected Improvement*:

$$EI_{y^*}(x) = \int_{-\infty}^{y^*} (y^* - y)p(y|x)dy \quad (3)$$

onde y^* é um valor limite da função objetivo, x é o conjunto proposto de hiperparâmetros, y é o valor real da função objetivo usando hiperparâmetros x e $p(y|x)$ é o modelo de probabilidade substituto expressando a probabilidade de y dado x .

Uma implementação eficiente para SMBO é o *Tree-Structured Parzen Estimator* (*TPE*) (BERGSTRA et al., 2011), um método que em suas iterações coleta novas observações e, no final da iteração, decide quais conjuntos de parâmetros que deve testar em sequência. Por ser um método SMBO, ele emprega sua própria função substituta e possui o *Expected Improvement* como seu critério de avaliação.

O algoritmo TPE difere do SMBO original na forma como constrói o modelo substituto. O TPE aplica a regra de Bayes em vez de representar diretamente $p(y|x)$:

$$p(x|y) = \frac{p(x|y) * p(y)}{p(x)} \quad (4)$$

As observações coletadas até então são divididas em dois grupos: o primeiro grupo contém aqueles com os melhores *scores*, chamado $l(x)$, enquanto o segundo grupo contém o resto das observações, chamado $g(x)$. O objetivo é encontrar um conjunto de hiperparâmetros com maior probabilidade de estar no primeiro grupo e menor probabilidade de estar no segundo grupo (BERGSTRA et al., 2011).

Esses processos modelam as variáveis de domínio quando a função objetivo está abaixo e acima de um quantil y^* especificado. A probabilidade dos hiperparâmetros dada a pontuação na função objetivo é expressa por:

$$p(x|y) = \begin{cases} l(x) & \text{if } y < y^* \\ g(x) & \text{if } y \geq y^* \end{cases} \quad (5)$$

onde $l(x)$ é a distribuição dos hiperparâmetros quando a pontuação é inferior ao limite y^* e $g(x)$ é a distribuição quando a pontuação é superior a y^* .

Aplicando a regra de Bayes e substituições, o *Expected Improvement* para TPE torna-se (BERGSTRA et al., 2011):

$$EI_{y^*}(x) = \frac{\gamma y^* l(x) \int_{-\infty}^{y^*} p(y) dy}{\gamma l(x) + (1 - \gamma) g(x)} \propto (\gamma + \frac{g(x)}{l(x)} (1 - \gamma))^{-1} \quad (6)$$

A cada iteração do algoritmo, as distribuições se tornam mais esparsas. Os pontos com alta probabilidade em $l(x)$ e baixa probabilidade em $g(x)$ devem ser escolhidos como o próximo hiperparâmetro. A seleção dos melhores hiperparâmetros forma um espaço em árvore (que dá o nome ao algoritmo), no qual a descida dos caminhos no grafo forma o caminho de decisão tomado pelo algoritmo.

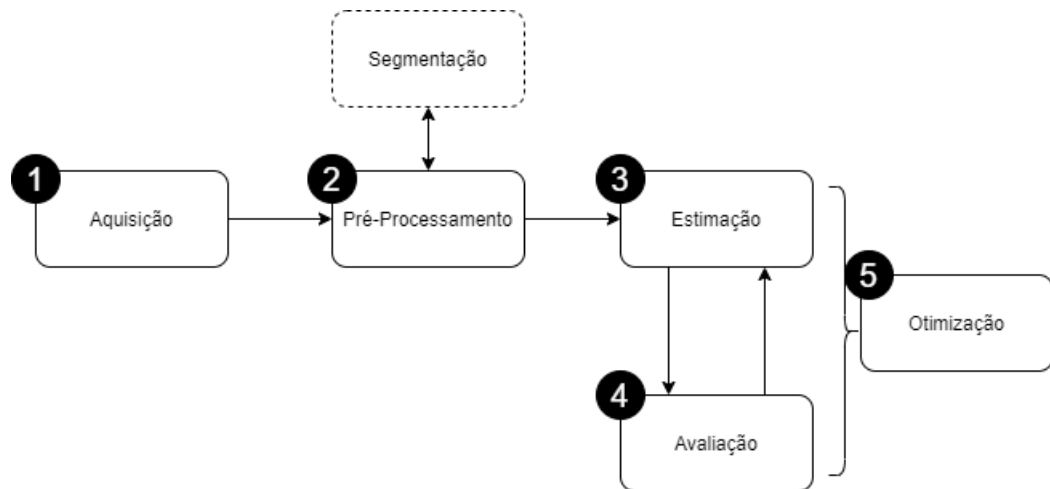
Em cada análise, o TPE acumula conhecimento usado na escolha do próximo conjunto de hiperparâmetros. Após um certo número de iterações, espera-se que ele seja capaz de estimar bons conjuntos de parâmetros.

Este capítulo apresentou os conceitos necessários para melhor entendimento da metodologia proposta. A pneumonia, que é a doença alvo desta tese, foi detalhada explicando seus sintomas e tratamentos. Detalhou-se também a radiografia torácica e a técnica de aprendizado profundo, que foi a técnica utilizada para a classificação automática. Por fim, discorreremos sobre o meta-aprendizado destacando a otimização bayesiana, tipo este utilizado para a otimização das arquiteturas geradas.

4 Metodologia proposta

Esta tese visa a proposição de uma metodologia para estimar redes neurais convolucionais específicas para detecção de pneumonia e classificação quanto aos tipos viral, bacteriano e COVID-19, através de imagens de raio-X. A questão de pesquisa consiste em demonstrar a viabilidade da construção de arquiteturas de redes neurais convolucionais, estimadas automaticamente, que tenham desempenho satisfatório quando aplicadas ao problema. Para tanto, foi construída a metodologia composta por 5 etapas principais, apresentada na Figura 19. A primeira etapa é a aquisição da base de dados. A segunda etapa é a de pré-processamento da imagens. A terceira e quarta etapas são a de estimação e avaliação respectivamente. São nessas etapas que a arquitetura especializada será estimada. A quinta e última etapa é a etapa de otimização, onde a arquitetura estimada na etapa anterior será otimizada. Essas três últimas etapas trabalham em conjunto na busca pela melhor arquitetura de CNN baseada em métricas predeterminadas de avaliação.

Figura 19 – Metodologia Proposta



Fonte: Elaborada pela autora.

As subseções a seguir apresentam em detalhes os passos da metodologia proposta.

4.1 Aquisição das Imagens

Os experimentos realizados neste trabalho utilizam duas bases de imagens: a base de imagens de raio-X publicada por Kermayn et al. (2018) e a base de raio-X publicada por Rahman (2020) com casos de COVID-19.

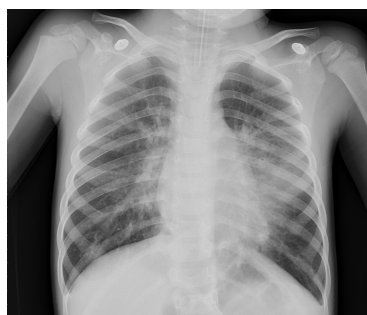
4.1.1 Base de imagens de Pneumonia

As imagens de raio-X disponibilizada por Kermany et al. (2018) foram coletadas de pacientes com idade entre um e cinco anos no *Guangzhou Woman and Children's Medical Center*, em Guangzhou, China, como parte de sua rotina de atendimento clínico. Trata-se de um conjunto de imagens de raio-X do tórax relativamente grande e heterogêneo, com imagens com altos e baixos níveis de contraste, formas pulmonares deformadas devido a infecções por pneumonia, posicionamento incorreto do paciente e ruídos como objetos médicos, mãos de adultos, cabeça e/ou braços do paciente e outras áreas da imagem fora dos pulmões, como o estômago. A rotulagem dessas imagens foi realizada por dois médicos especialistas.

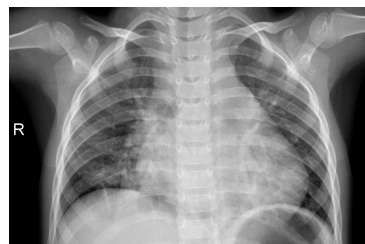
A base é composta por dois conjuntos, um de treino e outro de teste, obedecendo a proporção de 90% para treino e 10% para teste. O conjunto de treinamento possui um total de 5.232 raio-X do tórax de crianças, no qual 3.883 são caracterizadas como tendo pneumonia e 1.349 não tendo pneumonia. Das imagens rotuladas como pneumonia, 2.538 são diagnosticadas como bacteriana e 1.345 como viral. No conjunto de teste estão 234 normais e 390 com pneumonia, onde destes 242 são casos bacterianos e 148 virais.

A base apresenta o rótulo por imagem, mas não apresenta a localização da lesão. A Figura 20 apresenta exemplos de imagens da base. As imagens possuem resolução diversificada, sempre superiores a 1024x768, mantendo essa razão de aspecto.

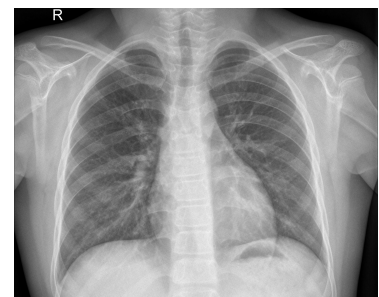
Figura 20 – Exemplos de imagens da base fornecida por Kermany et al. (2018)



(a) Pneumonia Viral



(b) Pneumonia Bacteriana



(c) Normal

Para fins de simplificação na comparação de trabalhos, foram utilizados os mesmos conjuntos de treinamento e teste propostos na base como os conjuntos para os experimentos realizados pelo método proposto nesse trabalho.

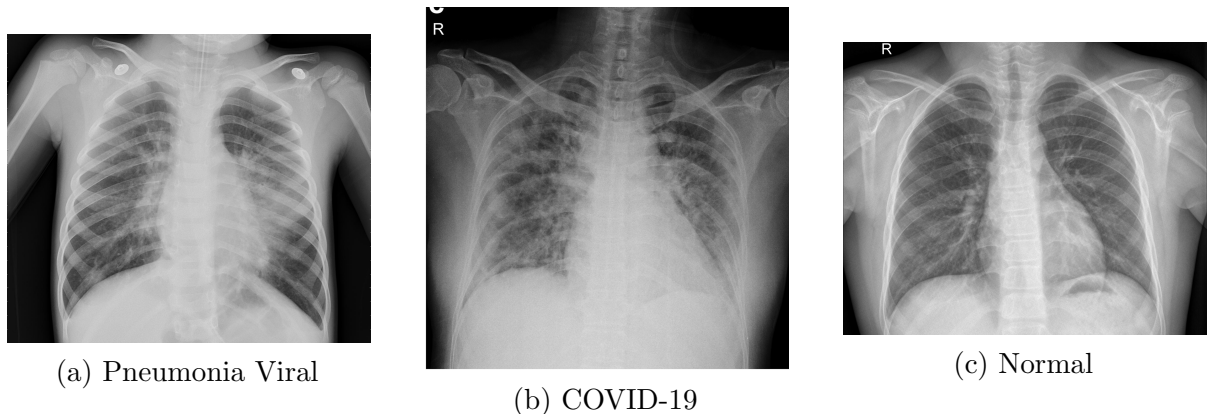
4.1.2 Base de imagens de COVID-19

A base com casos de COVID-19 disponibilizada em Rahman (2020) foi criada por equipes de pesquisadores da *Qatar University*, em Doha e da *University of Dhaka*, em Bangladesh, juntamente com colaboradores do Paquistão e da Malásia. A base é formada por imagens de raios-X de tórax com casos positivos de COVID-19, imagens de pneumonia viral e imagens de pulmão normal. As imagens que compõem a base foram coletadas de diferentes conjuntos de dados acessíveis ao público. As imagens de COVID-19 foram coletados de diferentes artigos publicados, da *Italian Society of Medical and Interventional Radiology COVID-19 Database (SIRM)* (SIM, 2020) e de *Novel Corona Virus 2019 Dataset* (COHEN; MORRISON; DAO, 2020). As imagens de casos normais e de pneumonia viral foram coletadas da base fornecida por Kermany et al. (2018). A base está sendo constantemente atualizada. Os experimentos realizados nesta tese utilizaram a primeira versão da base, que possuía 219 casos de COVID-19, 1341 imagens de raios-X de tórax normais e 1345 casos de pneumonia viral.

Assim como a base de pneumonia, a base de COVID-19 apresenta o rótulo por imagem, mas não apresenta a localização da lesão. A Figura 21 apresenta exemplos de imagens da base. Por ser uma base formada de imagens de diferentes lugares, as imagens possuem resolução diversificadas e não há detalhamento sobre os pacientes, como idade ou sexo.

A imagens foram separadas aleatoriamente em conjuntos de treino e teste, sendo que 80% das imagens foram para o conjunto de treinamento e 20% para o conjunto de teste.

Figura 21 – Exemplos de imagens da base fornecida por Rahman (2020)



4.2 Pré-Processamento

Com a finalidade de balancear a base de dados de treinamento, foi realizado o *data augmentation*. Esse tipo de técnica provoca variações ou perturbações nas imagens para gerar indivíduos que simulem situações parecidas com a realidade, facilitando o treinamento de redes neurais convolucionais ao gerar maior variabilidade de dados, permitindo a rede a aprender com maior eficiência os padrões.

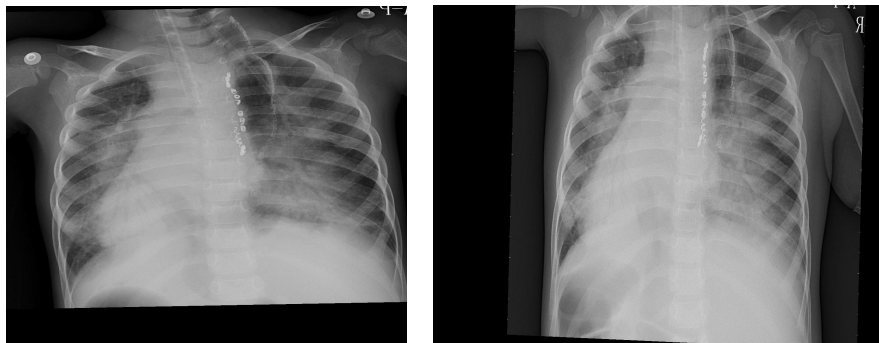
Com o *data augmentation* foram realizadas pequenas variações de rotação, escala e espelhamento nas imagens, de maneira aleatória, até que a base estivesse balanceada. Foram aplicadas as transformações:

- Espelhamento
- Rotação entre 5 e 15 graus, com passo de 5 graus
- Escala entre 0.90 e 0.95

Na base de pneumonia, o processo de *data augmentation* foi realizado apenas para a tarefa de diagnóstico entre as pneumonias. Assim, a base de treino para diagnóstico ficou com 5.234 imagens, sendo 2.540 casos de pneumonia bacteriana e 2.694 casos de pneumonia viral. Na base de COVID-19 foi feito *data augmentation* nas três categorias: normal, pneumonia viral e COVID-19. Após o balanceamento, a base de treino ficou com 2.339 casos normais, 2.278 casos de COVID-19 e 2.023 casos de pneumonia viral.

A Figura 22 apresenta alguns exemplos de imagens geradas pelo processo.

Figura 22 – Dois exemplos de *data augmentation* da base para o balanceamento da mesma.



O segundo pré-processamento realizado foi a uniformização da resolução das imagens para as dimensões 640x450 na base de pneumonia e 450x450 na base de COVID-19, tamanho este escolhido por ser suficiente para determinar detalhes da imagens e limite para a carga de memória imposta pela metodologia de treinamento de redes neurais

convolucionais. A mudança de resolução manteve sempre a razão de aspecto das imagens originais, preenchendo os espaços vazios com zero quando necessário.

Adicionalmente, como parte do pré-processamento, foi realizada a segmentação da região do pulmão como medida para destacar estruturas e direcionar o método de aprendizado. A seção seguinte trata da segmentação do pulmão realizada.

4.2.1 Segmentação

A segmentação de imagem é o processo de delimitação da imagem em regiões relacionadas. Essa separação da região de interesse permite extrair as características e anomalias do tecido desejado (VITAL; SAIS; MORAES, 2018). Consequentemente, é possível quantificar e analisar características difíceis de detectar ou distinguir a olho nu (SOKASHE, 2013).

Dessa forma, a utilização ou criação de métodos de segmentação pulmonar para radiografias de tórax torna-se importante para um melhor diagnóstico (CHEN; SUZUKI, 2013). Uma vez que a região pulmonar é isolada, resta uma quantidade reduzida de informações a serem pesquisadas. Assim, a segmentação pode melhorar a busca de doenças na inspeção visual, além de ser usada como uma etapa inicial para construir outro método computacional dedicado a segmentar ou extrair uma doença específica (VITAL; SAIS; MORAES, 2018).

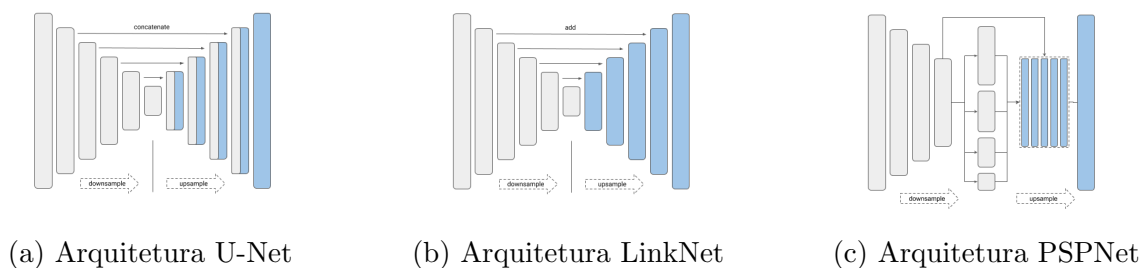
As bases de imagens utilizadas nesta proposta não possuem marcação do especialista. Quando não existe informação de localização da lesão, mas apenas o diagnóstico geral, os métodos empregados para extração de característica precisam também localizar onde estão as informações mais importantes para o reconhecimento. Esse é um cenário viável para redes neurais convolucionais pois elas são capazes de aliar o treinamento a extração de características invariante a localização e escala.

Todavia, embora possível, o processo de treinamento pode ser simplificado quando somente a região de interesse está disponível para processamento. Para realizar a segmentação na base em questão, esta metodologia avalia uma proposta de aplicação de *transfer learning*, onde redes totalmente convolucionais serão treinadas com imagens anotadas. O modelo treinado é então aplicado às imagens que se deseja gerar marcação, sem informação prévia dessas imagens durante o treinamento.

Para realizar a segmentação pulmonar na base fornecida por Kermay et al. (2018), avaliou-se diferentes redes de segmentação semântica: U-Net (RONNEBERGER; FISCHER; BROX, 2015), LinkNet (CHAURASIA; CULURCIELLO, 2017) e PSPNet (ZHAO et al., 2017), em conjunto com quatro redes para extração de características: ResNet (HE et al., 2015), Inception (SZEGEDY et al., 2014), DenseNet (HUANG et al., 2017) e EfficientNet (TAN; LE, 2019).

A arquitetura U-Net é flexível e aplicável a vários problemas de segmentação biomédica. É uma rede totalmente convolucional que pode ser treinada com muito poucas imagens e, ainda assim, produz uma segmentação precisa se combinada com o aumento de dados. Como apresentado na Figura 23a, ela consiste em duas partes: uma unidade de codificação, chamada *downsample*, e uma unidade de decodificação, chamada *upsample*. No *downsample*, cada etapa consiste em duas convoluções 3x3 e uma operação de maxpooling 2x2 com passo de 2 para redução da amostra. No *downsample* a U-Net aprende características da imagem. No *upsample*, a rede transforma essas características em ativações localizadas que levam à máscara desejada.

Figura 23 – Arquiteturas de Redes de Segmentação Semântica



A Linknet também possui arquitetura codificador-decodificador (Figura 23b) e se diferencia da U-Net por utilizar blocos residuais (add) entre os mapas de ativação ao invés de concatená-los com camadas paralelas. A PSPNet se baseia no uso de informações globais, através da coleta de informações de contexto da cena e relacionamento entre classes. Como apresentado na Figura 23c, a arquitetura básica da PSPNet consiste em duas etapas: uma rede neural convolucional pré-treinada que extrai o mapa de recursos a partir de uma imagem de entrada. Esse mapa de recursos é posteriormente concatenado com a saída do módulo de pool de pirâmide. E uma última convolução gera a máscara de segmentação prevista.

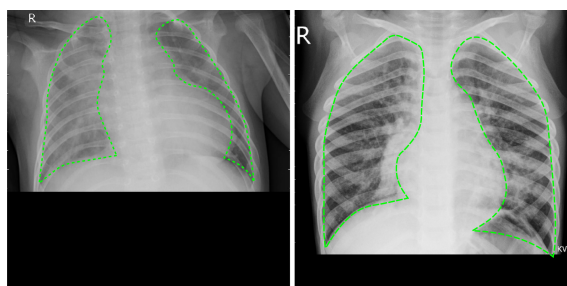
Apoiando-se na técnica de *transfer learning*, o treinamento das redes usadas para a segmentação do pulmão foi realizado em diferentes bases que possuem a marcação do especialista.

As bases de imagens utilizadas para treinar o modelo foram JSRT (SHIRAISHI et al., 2000), Montgomery County Dataset (JAEGER et al., 2014), Shenzhen CXR (JAEGER et al., 2014) e NIH dataset (WANG et al., 2017). Com os quatro conjuntos de dados anotados, chegou-se a um total de 1.147 imagens de radiografia torácicas, incluindo pósterio-anterior e anteroposterior, adulto e pediátrico, níveis de alto e baixo contraste e imagens normais e anormais de radiografia de tórax. Para manter a consistência nesse novo conjunto de imagens, foram adicionadas bordas às imagens com diferentes dimensões para manter a forma quadrada e, em seguida, todas as imagens foram redimensionadas para 512 x 512 pixels.

Cada rede (U-Net, LinkNet e PSPNet) foi treinada com o conjunto de dados gerado usando cada uma das redes propostas para extração de características (ResNet, Inception, DenseNet e EfficientNet). O processo de treinamento foi realizado usando opções padrão de otimizador, taxa de aprendizado e configurações de camadas fornecidas por Yakubovskiy (2019).

Com o modelo treinado, foi feita a segmentação do pulmão sobre a base do *Guangzhou Woman and Children's Medical Center*. Para validar o método, 521 imagens foram marcadas manualmente, sendo 10% de cada classe (normal, bacteriana e viral). As imagens foram escolhidas aleatoriamente sobre o conjunto de classes. Inspirado por Ginneken, Stegmann e Loog (2006), no processo de anotar essas imagens, um observador usou um mouse para marcar pontos nos limites dos pulmões. Para cada imagem, gerou-se uma única imagem com máscaras de pulmão esquerdo e direito. Alguns exemplos dessas imagens marcadas manualmente são fornecidos na Figura 24.

Figura 24 – Dois exemplos de marcação manual do pulmão.



4.3 Estimação e Avaliação

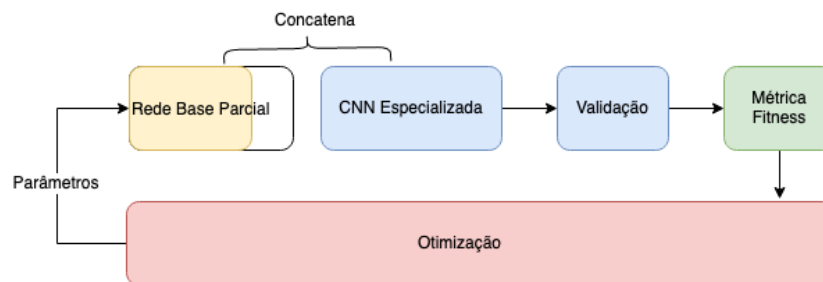
Esta etapa descreve como é realizada a identificação e reconhecimento dos padrões de pneumonia. Este é um processo fortemente dependente de características discriminantes dos padrões relacionados. Em abordagens tradicionais, utiliza-se a marcação do especialista como forma de localizar a lesão e depois se extrai informações sobre essas regiões, mesmo que apenas a informação de classe. Essas características são atribuídas a um classificador com a função de gerar um modelo discriminatório entre elas para classificação.

Treinar redes neurais convolucionais exige um grande volume de imagens visto a quantidade de pesos a serem estimados na rede. Por este motivo, técnicas de *transfer learning* podem ser utilizadas para inicializar o treinamento da rede de um ponto aceitável, partindo do pressuposto de que as características aprendidas de maneira hierárquica pela rede são mais complexas quanto mais próximas ao topo e primitivas quanto próximas da base. Características primitivas tendem a ser compartilhadas por qualquer problema no domínio de processamento de imagens. Outro desafio consiste em ajustar adequadamente a arquitetura da rede para o problema, visto a variabilidade de topologias e parâmetros. Dado os fatores destacados, esta metodologia propõe a construção otimizada de arquiteturas de redes neurais convolucionais a partir da combinação de camadas de redes pré-treinadas na forma de uma rede base, esta que é refinada através de uma nova rede especializada para o problema, conectada após a rede base, com topologia e hiperparâmetros estimados.

A Figura 25 resume os elementos propostos para atingir o objetivo dessa etapa. Inicialmente será usado *transfer learning* de uma rede base, pré-treinada com pesos da ImageNet.

Para esta metodologia, a camada de classificação da rede base foi retirada. Adicionalmente, não serão utilizadas todas as camadas remanescentes de convolução. Essa decisão será tomada durante a otimização da arquitetura completa da rede de maneira a escolher o melhor conjunto inicial de camadas da rede base juntamente com a rede estimada de maneira a gerar a melhor topologia para o problema. Essa decisão se fundamenta na informação de que as redes foram previamente treinadas para a ImageNet e nem todas as camadas, principalmente as que estão no topo, são adequadas para o problema específico. Assim, para reduzir o tamanho da arquitetura estimada e também deixá-la mais eficiente, será estimada automaticamente qual a camada de saída mais adequada da rede base utilizada.

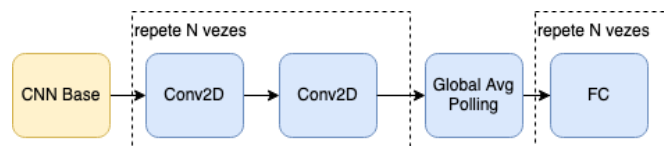
Figura 25 – Diagrama que ilustra os elementos do processo de estimação de topologia e parâmetros da rede convolucional proposta por esta metodologia



Fonte: Elaborado pela autora.

Após as redes bases, é acoplada uma nova arquitetura CNN, especializada para o problema. Essa nova rede é acoplada ao topo da rede base estimada anteriormente. São avaliadas duas abordagens de estimação da arquitetura dessa rede. Na primeira abordagem, denominada Sequencial, a nova rede é composta por uma sequência de blocos convolucionais, com filtros 3x3, seguido por um bloco de *pooling* e uma sequência de blocos totalmente conectados (Figura 26). Essa primeira abordagem é inspirada no conceito da VGG-16 onde pode-se adicionar camadas convolucionais sucessivas de maneira a melhorar a representação e complexidade dos padrões.

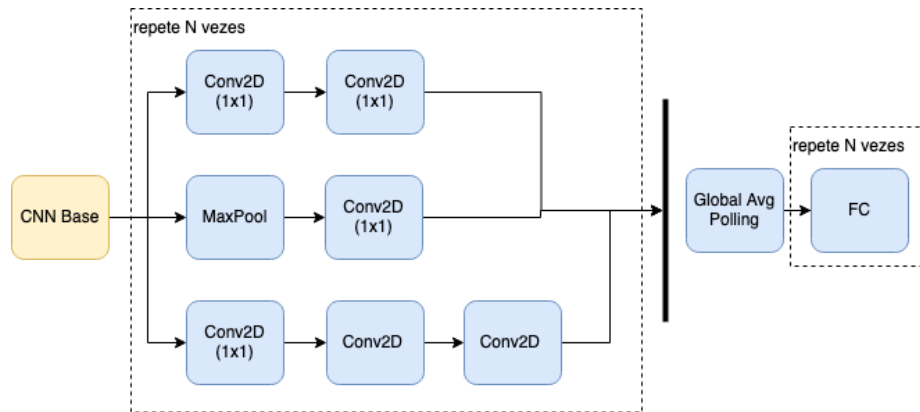
Figura 26 – Ilustra o modelo de estimação da rede especializada baseado em uma sequência de blocos convolucionais concatenados.



Fonte: Elaborado pela autora.

Na segunda abordagem de estimação, a rede é composta por uma sequência de blocos Inception, seguido por um bloco de *pooling* e uma sequência de blocos totalmente conectados (Figura 27). O princípio por trás da segunda abordagem consiste em utilizar blocos especializados em aprendizado rápido de informações, propostos nas redes Inception (SZEGEDY et al., 2017). Os blocos convolucionais 1x1 são usados para aumentar a complexidade do mapeamento das características, em paralelo a blocos de *pooling* e blocos convolucionais com filtros 3x3 e 5x5. Esses blocos com filtros maiores são usados para estimar, mesmo usando uma quantidade de pesos menor, uma maior quantidade de informação.

Figura 27 – Ilustra o modelo de estimação da rede especializada baseado em uma sequência de blocos Inception concatenados.



Fonte: Elaborado pela autora.

A decisão de quantos blocos (Convolação/Inception/Totalmente Conectados), quantidade de filtros, quantidade de neurônios, dentre outros parâmetros é tomada durante processo de otimização. A arquitetura gerada é avaliada e seu retorno auxiliará o otimizador a decidir o próximo conjunto de hiperparâmetros.

4.4 Otimização

O otimizador (em vermelho na Figura 25) é responsável por orquestrar a operação e o correto funcionamento do treinamento que tem como finalidade estimar o melhor conjunto de parâmetros e hiperparâmetros da rede. Este método propõe a utilização dos conceitos do *Sequential Model-Based* (SMBO) (PELIKAN; GOLDBERG; CANTÚ-PAZ, 1999) para realizar não só a estimação de hiperparâmetros mas também da topologia da rede.

Uma implementação eficiente para o SMBO é o *Tree-structured Parzen Estimator* (TPE) (BERGSTRA et al., 2011), um método que em suas iterações realiza a coleta de novas observações, e ao fim da iteração decide de forma otimizada qual conjunto de parâmetros ele deve testar na sequência. Sendo um método *SMBO*, ele emprega uma função *surrogate* própria e tem como critério de avaliação o *Expected Improvement*.

A cada análise de resultado da escolha de parâmetros, o TPE acumula conhecimentos utilizados na escolha do próximo conjunto de parâmetros. O que se espera é que após um certo número de iterações, o otimizador seja capaz de estimar bons conjuntos de parâmetros. A métrica de *fitness* utilizada para a escolha foi a de acurácia do modelo.

Portanto, ao fim da otimização, tem-se a estimação da arquitetura e hiperparâmetros completos da CNN projetada para o problema, composta pela utilização de algumas camadas de uma rede base concatenadas a uma CNN totalmente estimada de maneira a maximizar os resultados.

A Tabela 4 sumariza as decisões arquiteturais a serem tomadas pelo otimizador. A otimização é executada 100 vezes para cada problema e abordagem de estimação de rede. Todas as redes são treinadas por 20 épocas iniciais para adaptar os pesos da rede especializada e depois por 100 épocas com todas as camadas em treinamento. A cada 3 épocas sem melhoria na taxa de perda da validação, é realizada uma redução de 10% na taxa de aprendizado do otimizador, que inicia com o valor 0.0001. Todos os problemas foram ajustados com ponderação de classes, onde a classe minoritária recebeu peso 2 e a classe majoritária recebeu peso 1, de maneira adicional ao balanceamento realizado na etapa de pré-processamento. A camada dense de classificação para todos os problemas foi ajustada com a função de ativação *sigmoid*.

Tabela 4 – Decisões arquiteturais para serem avaliadas pelo otimizador ao estimar o melhor modelo.

Parâmetros	Valores Sequencial	Valores Inception
Número de camadas	3 a 7	1 a 6
Quantidade de filtros	128, 256, 512	32, 64, 128, 256
Camada de saída	cinco últimas camadas da rede base utilizada	
Número de camadas Dense	1 a 3	
Quantidade de neurônios	1024, 512, 256, 128	
Dropout	uniforme(de 0.4 a 0.6)	
Otimizador	SGD, RMSprop, Adam	

Neste capítulo apresentou-se a metodologia proposta para a otimização de redes neurais convolucionais, principal contribuição desta tese. A metodologia proposta é composta por cinco etapas, que trabalham em conjunto para estimação tanto da arquitetura quanto de seus hiperparâmetros. Para a estimação das arquiteturas foram pensadas duas abordagens, denominadas Sequencial e *Inception*. Por fim, foi apresentado o espaço de busca dos hiperparâmetros, valores esses que serão considerados para a otimização das arquiteturas.

5 Resultados e Discussão

Nesta seção são apresentados os resultados obtidos pelo método proposto para segmentação do pulmão e pela metodologia proposta para classificação das imagens de raio-X. Na tarefa de classificação, foram realizados experimentos com dois objetivos: detecção e diagnóstico de pneumonia em imagens de raio-x. Especificamente na base de pneumonia, nos experimentos de detecção, deseja-se identificar a presença de pneumonia na imagem de raio-X ou simplesmente classificá-la como normal. Nos experimentos de diagnóstico, deseja-se identificar em imagens de raio-x com pneumonia se o tipo de pneumonia é viral ou bacteriana. Já na base de COVID-19, nos experimentos de detecção, deseja-se identificar a presença da COVID-19 na imagem de raio-X ou simplesmente classificá-la como normal. Nos experimentos de diagnóstico, deseja-se identificar se dado raio-X é classificado como normal, como pneumonia viral ou como COVID-19.

Os experimentos de classificação foram conduzidos utilizando a biblioteca Keras (CHOLLET et al., 2015). A implementação do TPE utilizada foi a fornecida pelo HyperOpt (BERGSTRA et al., 2013). As bases utilizadas foram fornecidas por Kermany et al. (2018) e Rahman (2020), descritas no Capítulo 4. Em cada um dos experimentos, testou-se três arquiteturas de CNN como rede base: VGG-16, ResNet-50 e DenseNet-169. Os testes foram realizados em um computador com o processador core i5, 16 gigabyte de RAM, e placa de vídeo GeForce 1080TI.

Como forma de avaliação, foram utilizadas as métricas de precisão, sensibilidade, especificidade, F-score e acurácia. Todas as métricas são obtidas diretamente da matriz de confusão, onde nos experimentos de detecção, a classe positiva é representada pela pneumonia e a classe negativa representada por um exame normal. Nos experimentos de diagnóstico, a classe positiva é representada pela pneumonia bacteriana e a classe negativa representada pela pneumonia viral. Sensibilidade e especificidade referem-se a taxa de acerto nas classes positivas e negativas, respectivamente. Precisão é definida como a parcela de acerto dentro do conjunto de casos informados pelo método como positivos e F-Score é definida como a média harmônica entre precisão e sensibilidade. Essas métricas possuem a seguinte formulação sobre a matriz de confusão:

Acurácia:

$$\frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

Precisão:

$$\frac{TP}{TP + FP} \quad (8)$$

Sensibilidade:

$$\frac{TP}{TP + FN} \quad (9)$$

Especificidade:

$$\frac{TN}{TN + FP} \quad (10)$$

F-score:

$$\frac{2 \times TP}{2 \times TP + FP + FN} \quad (11)$$

onde TP (*True Positive*) significa os exemplos positivos que foram corretamente classificados como tal, TN (*True Negative*) significa os exemplos negativos corretamente classificados, FP (*False Positive*) e FN (*False Negative*) são respectivamente exemplos negativos classificados incorretamente como positivos e vice-versa. A Tabela 5 contém esses conceitos e é chamada de matriz de confusão.

Tabela 5 – Matriz de confusão para duas classes.

		Predição	
		Negativo	Positivo
Classes	Negativo	TN	FP
	Positivo	FN	TP

5.1 Resultados da Segmentação do Pulmão

Cada um dos modelos de segmentação aqui apresentados foi treinado por 20 épocas com os conjuntos de dados fornecido por Kermany et al. (2018) e seus desempenhos foram avaliados com o subconjunto anotado por nós.

Para treinamento, foi realizada a divisão de forma aleatória do conjunto de imagens anotadas em 2 conjuntos, treinamento e validação, com 918 e 229 imagens, respectivamente. O conjunto de imagens de testes utilizado foram as imagens anotadas no banco de dados de radiografia de tórax.

As métricas usadas nesta etapa de segmentação foram o coeficiente de Dice (DSC), sensibilidade e especificidade. O coeficiente de Dice mede a semelhança entre a região segmentada com o *ground truth* e pode ser calculado por:

$$DSC = \frac{2TP}{2TP + FP + FN} \quad (12)$$

A Tabela 6 apresenta os resultados do método na fase de treinamento no conjunto de validação. E a Tabela 7 apresenta os resultados do aprendizado de transferência para o conjunto de testes. Os resultados são ordenados por DSC. Nas duas fases, a LinkNet e a U-Net tiveram desempenhos semelhantes e promissores, atingindo 0,966 de DSC médio na fase de treinamento.

Na fase de teste, o melhor resultado foi encontrado pela U-net com 0,917 de DSC médio. A U-Net alcança uma pequena vantagem de sensibilidade sobre a LinkNet nos dois problemas. O desvio padrão relativamente baixo no DSC indica que esses modelos podem alcançar uma alta generalização na tarefa de segmentação pulmonar.

Como apresentado nas Tabelas 6 e 7, a PSPNet não foi capaz de aprender com nosso conjunto de treinamento. Os resultados da validação e do conjunto de testes demonstram que essa rede não é aplicável à segmentação pulmonar, pois a sensibilidade máxima e a especificidade mínima indicam que este modelo prevê uma máscara quadrada plana, segmentando toda a imagem de entrada.

Tabela 6 – Resultados de segmentação no conjunto de validação, fase de treinamento do método

Rede	Backbone	DSC	Sensibilidade	Especificidade
U-Net	DenseNet 169	0,966±0,024	0,971	0,988
LinkNet	ResNet 101	0,966±0,024	0,967	0,990
U-Net	Efficientnet B3	0,966±0,026	0,970	0,989
U-Net	ResNet 101	0,965±0,02	0,972	0,987
LinkNet	Inception V3	0,965±0,025	0,972	0,987
LinkNet	DenseNet 169	0,964±0,024	0,974	0,986
U-net	Inception V3	0,964±0,025	0,974	0,986
LinkNet	Efficientnet B3	0,963±0,025	0,977	0,985
PSPNet	ResNet 101	0,390±0,079	1	0
PSPNet	DenseNet 169	0,390±0,079	1	0
PSPNet	Inception V3	0,390±0,079	1	0
PSPNet	Efficientnet B3	0,390±0,079	1	0

A Figura 28 apresenta as curvas DSC para cada combinação de redes de segmentação e *backbone*. O eixo x representa o coeficiente de dados para imagem e o eixo y representa o

Tabela 7 – Resultados de segmentação no conjunto de testes, fase de aprendizado de transferência do método

Rede	Backbone	DSC	Sensibilidade	Especificidade
U-Net	Inceptionv3	$0,917 \pm 0,052$	0,912	0,986
U-Net	Densenet169	$0,917 \pm 0,053$	0,903	0,988
Linknet	Resnet101	$0,916 \pm 0,052$	0,904	0,988
Linknet	Inceptionv3	$0,914 \pm 0,050$	0,927	0,981
U-Net	Resnet101	$0,914 \pm 0,051$	0,923	0,983
U-Net	Efficientnetb3	$0,914 \pm 0,052$	0,918	0,984
Linknet	Densenet169	$0,914 \pm 0,064$	0,909	0,986
Linknet	Efficientnetb3	$0,912 \pm 0,051$	0,924	0,981
U-Net	None	$0,871 \pm 0,108$	0,847	0,985
PSPNet	Inceptionv3	$0,286 \pm 0,071$	1	0
PSPNet	Efficientnetb3	$0,286 \pm 0,071$	1	0
PSPNet	Densenet169	$0,286 \pm 0,071$	1	0
PSPNet	Resnet101	$0,286 \pm 0,071$	1	0

quantitativo de imagens, onde, para cada combinação, o conjunto de testes foi classificado de dados mais baixos para mais altos. Percebe-se a semelhança entre os gráficos para U-Net e LinkNet.

Em 75,6% das imagens segmentadas pela U-Net o DSC foi superior a 90% e em 99,94% tiveram DSC maior que 80%. Apenas uma pequena fração de 5,3% obteve DSC menor que 80%.

- (a) Distribuição do DSC sobre todo o conjunto de imagens de teste para a U-Net (b) Distribuição do DSC sobre todo o conjunto de imagens de teste para a LinkNet

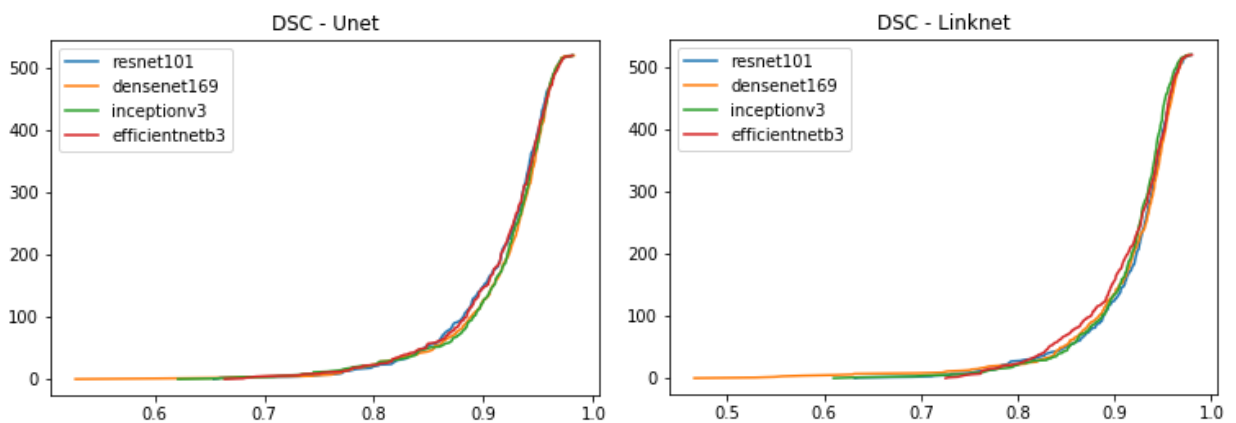


Figura 28 – Comparação do quantitativo de imagens por DSC para todas as redes

Apresentamos alguns casos de estudos nas Figuras 29 e 30 que possuem exemplos de imagens resultantes no conjunto de testes para U-Net e LinkNet, respectivamente. Em cada exemplo, a imagem da esquerda é um caso de sucesso e a imagem da direita é um

Figura 29 – Caso de estudo para o melhor resultado com a U-Net. A esquerda, um caso de sucesso com $DSC= 0,980$ e a direita, um caso de erro com $DSC=0,663$.

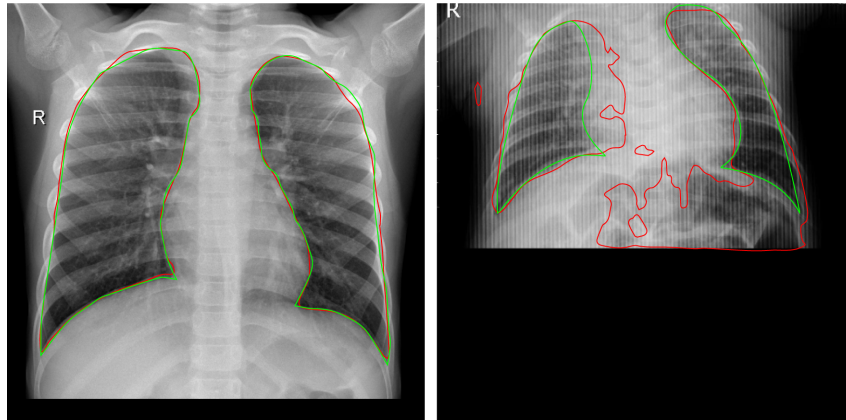
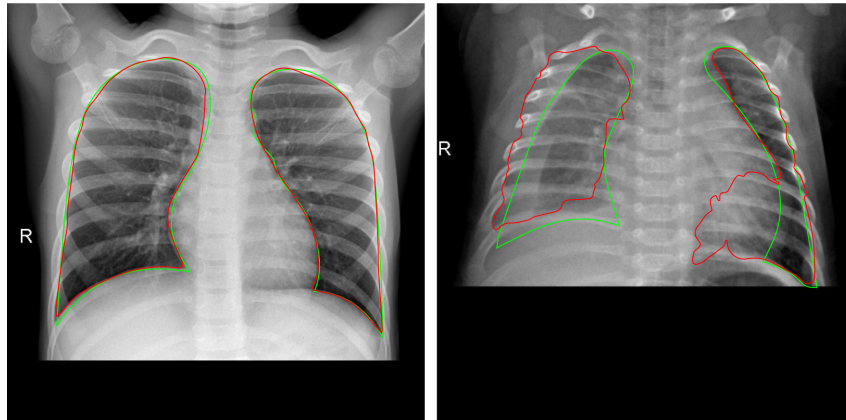


Figura 30 – Caso de estudo para o melhor resultado com a LinkNet. A esquerda, um caso de sucesso com $DSC= 0,978$ e a direita, um caso de erro com $DSC=0,725$.



caso falha. Para ambos, a linha verde indica a segmentação correta e a linha vermelha o resultado obtido.

Nos dois casos de sucesso apresentados, as imagens têm uma localização regular do pulmão, bem como suas estruturas bem identificadas. No caso apresentado como exemplo para a U-Net, pode-se observar que uma região muito semelhante ao pulmão está presente na imagem de uma maneira imediatamente adjacente a ela. Isso fez com que a rede perdesse a precisão do local nesse caso. Na imagem de falha da LinkNet, pode-se ver que o contraste da imagem é muito diferente da imagem bem-sucedida e, portanto, que as estruturas do plano de fundo são muito mais semelhantes às estruturas do próprio pulmão. Nesse caso, levando a rede a perder precisão.

5.2 Experimentos para detecção e diagnóstico de Pneumonia

Os experimentos realizados nesta seção utilizaram a base de imagens de raio-X publicada em Kermany et al. (2018). A base está descrita de forma mais detalhada no Capítulo 4, Seção 2.

5.2.1 Resultados para a detecção de Pneumonia

Nesta seção são apresentados os resultados encontrados pela CNN estimada para a detecção de pneumonia, tanto utilizando a segmentação de pulmão, como sem utilizar a segmentação. Todas as imagens da base foram segmentadas pelo método proposto, de forma automática, e utilizadas no processo de classificação. A etapa de construção e otimização da CNN para detecção de pneumonia foi conduzida por 30 épocas utilizando segmentação e 100 épocas sem a segmentação, para cada uma das abordagens projetadas: Sequencial e Inception. Como as bases de detecção e diagnóstico foram previamente balanceadas, a métrica de *fitness* utilizada para otimização foi a acurácia. Os resultados estão separados de acordo com a rede base utilizada: VGG-16, ResNet-50 e DenseNet-169.

Após determinar o melhor conjunto de hiperparâmetros, o melhor modelo em termos de F-score foi escolhido para cada uma das abordagens e executado com a base de teste, gerando os resultados apresentados nas Tabelas 8, 9 e 10.

Analisando as Tabelas 8, 9 e 10, observa-se que a segmentação de pulmão melhorou os resultados apenas quando se utiliza a VGG-16 como rede base. Apesar da ResNet-50 e DenseNet-169 apresentarem melhor resultado na taxa de sensibilidade com a segmentação, ao se observar também a taxa especificidade verifica-se que o modelo tendenciou somente a classe positiva.

Em relação às abordagens utilizadas, a rede VGG-16 apresenta resultados parecidos utilizando ambas abordagens. Com a ResNet-50 os resultados são melhores utilizando a abordagem Sequencial, já com a DenseNet-169 os melhores resultados encontrados são utilizando a abordagem Inception. De forma geral, a rede VGG-16 foi a que obteve melhor resultado considerando o F-Score, mas destaca-se aqui a rede DenseNet-169 utilizando a abordagem Inception, que obteve resultado parecido em termo de F-Score e uma alta taxa de Sensibilidade, atingindo 99%.

Tabela 8 – Resultados obtidos para a Detecção de Pneumonia - VGG-16

Abordagem	Segmentação	Precisão	Sens.	Espec.	Acurácia	F-Score
Sequencial	Não	0,954	0,976	0,923	0,956	0,965
	Sim	0,964	0,984	0,940	0,968	0,974
Inception	Não	0,964	0,979	0,940	0,964	0,972
	Sim	0,971	0,976	0,953	0,968	0,974

Tabela 9 – Resultados obtidos para a Detecção de Pneumonia - ResNet-50

Abordagem	Segmentação	Precisão	Sens.	Espec.	Acurácia	F-Score
Sequencial	Não	0,844	0,943	0,709	0,855	0,891
	Sim	0,624	1	0	0,624	0,768
Inception	Não	0,697	0,994	0,282	0,727	0,820
	Sim	0,624	1	0	0,624	0,768

Tabela 10 – Resultados obtidos para a Detecção de Pneumonia - DenseNet-169

Abordagem	Segmentação	Precisão	Sens.	Espec.	Acurácia	F-Score
Sequencial	Não	0,840	0,930	0,705	0,846	0,883
	Sim	0,639	0,997	0,068	0,648	0,779
Inception	Não	0,955	0,992	0,923	0,966	0,973
	Sim	0,632	1	0,034	0,636	0,774

Considerando o F-Score, a rede VGG-16 foi a que apresentou o melhor resultado, em ambas abordagens. Os hiperparâmetros escolhidos pela estimação Bayesiana em cada abordagem são apresentados na Tabela 11. Inicialmente, a quantidade de blocos escolhidos pela abordagem Sequencial é ligeiramente superior do que na abordagem Inception. Entretanto, a abordagem Sequencial gerou uma CNN bem mais profunda por utilizar 14 camadas iniciais da rede base, contra apenas 10 na abordagem Inception. Ambas diferem também no otimizador utilizado. Enquanto que a abordagem Sequencial utilizou o otimizador RMSprop, na Inception utilizou-se o Adam.

Tabela 11 – Hiperparâmetros estimados do melhor modelo - Detecção de pneumonia

Hiperparâmetro	Abordagem	
	Sequencial	Inception
Número de Camadas	6	4
Quantidade de filtros	128	64
Camada de Saída	14	10
Número de camadas Dense	2	2
Quantidade de neurônios	128	256
Dropout	0,5121	0,5407
Otimizador	RMSprop	Adam

Utilizando a arquitetura otimizada proposta nesta tese, o melhor modelo gerado para a detecção de pneumonia foi a VGG-16 utilizando a abordagem Inception, ilustrada na Figura 31. Esse modelo tem um total de 13.767.105 parâmetros treináveis. Comparada com a VGG-16 original (Tabela 12), houve redução de mais de 90% de parâmetros.

Figura 31 – Arquitetura otimizada (VGG-16) - Detecção de Pneumonia

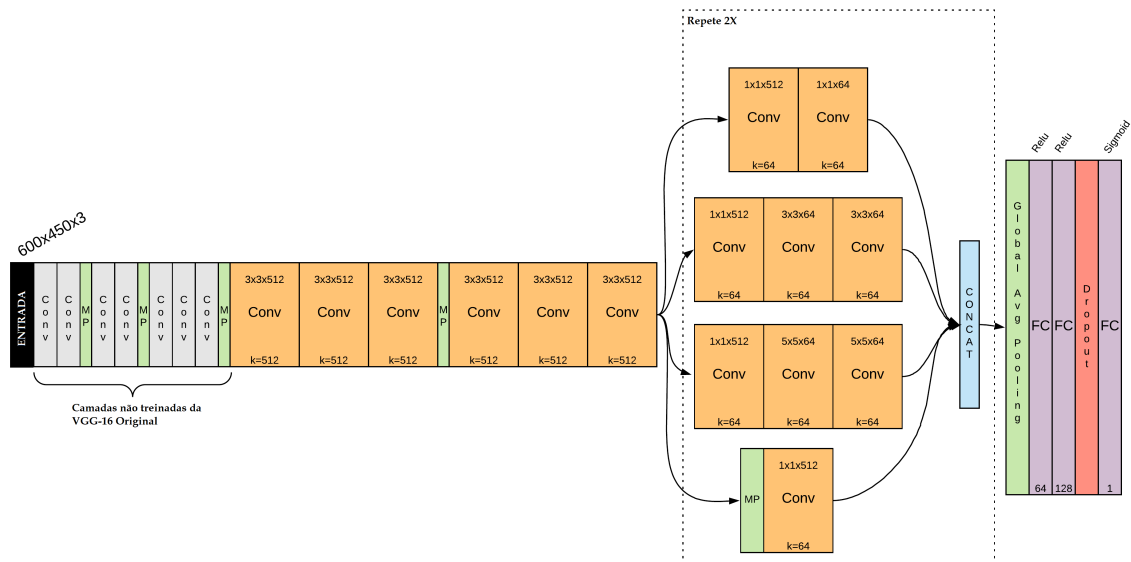


Tabela 12 – Comparação de Parâmetros - Detecção de Pneumonia

Rede	Total de Parâmetros	Parâmetros treináveis	Parâmetros Não treináveis
VGG-16 original	138.357.544	138.357.544	0
VGG-16 otimizada	15.502.593	13.767.105	1.735.488

5.2.2 Resultados para o diagnóstico entre Pneumonias

Assim como na detecção de pneumonia, foram executadas 100 épocas sem segmentação do pulmão e 30 épocas utilizando a segmentação para cada uma das abordagens projetadas: Sequencial e Inception. Todas as imagens da base foram segmentadas pelo método proposto, de forma automática, e utilizadas no processo de classificação. A acurácia foi usada como *fitness* na base de validação.

Após determinar o melhor conjunto de hiperparâmetros, o melhor modelo em termos de F-score foi escolhido para cada uma das abordagens e executado com a base de teste, gerando os resultados apresentados nas Tabelas 13, 14 e 15.

Analisando as Tabelas 13, 14 e 15, verifica-se que todas as redes apresentaram resultados inferiores ao utilizarem a segmentação de pulmão. Assim como na tarefa de detecção de pneumonia, as redes ResNet-50 e DenseNet-169 apresentaram alta taxa de sensibilidade com baixíssima especificidade. Um dos motivos para os resultados utilizando segmentação terem sido inferiores pode ser o menor número de execuções: como o diagnóstico de pneumonia é uma tarefa mais específica, o otimizador não conseguiu encontrar o melhor modelo dentro do número estimado de execuções. Ainda assim, os resultados demonstram que o *transfer learning* somado à estimação de rede especializada adicional gera resultados promissores e robustos.

Em relação às abordagens utilizadas, observa-se que a abordagem Inception superou a abordagem Sequencial em todos os testes. Assim como na detecção, na tarefa de diagnóstico entre pneumonias a rede VGG-16 também obteve os melhores resultados, com a taxa de Sensibilidade alcançando 95.50% .

Tabela 13 – Resultados obtidos para o Diagnóstico de Pneumonia - VGG-16

Abordagem	Segmentação	Precisão	Sens.	Espec.	Acurácia	F-Score
Sequencial	Não	0,947	0,966	0,912	0,946	0,957
	Sim	0,913	0,958	0,851	0,917	0,935
Inception	Não	0,947	0,975	0,912	0,951	0,961
	Sim	0,920	0,958	0,864	0,923	0,939

Tabela 14 – Resultados obtidos para o Diagnóstico de Pneumonia - ResNet-50

Abordagem	Segmentação	Precisão	Sens.	Espec.	Acurácia	F-Score
Sequencial	Não	0,805	0,756	0,884	0,838	0,780
	Sim	0,624	1	0	0,624	0,768
Inception	Não	0,850	0,689	0,925	0,835	0,761
	Sim	0,728	0,921	0,439	0,738	0,813

Tabela 15 – Resultados obtidos para o Diagnóstico de Pneumonia - DenseNet-169

Abordagem	Segmentação	Precisão	Sens.	Espec.	Acurácia	F-Score
Sequencial	Não	0,896	0,878	0,938	0,620	0,887
	Sim	0,620	1	0	0,620	0,765
Inception	Não	0,909	0,885	0,946	0,923	0,897
	Sim	0,623	1	0,013	0,625	0,768

Considerando o F-Score, a rede VGG-16 foi a que apresentou o melhor resultado, em ambas abordagens. Analisando os hiperparâmetros escolhidos pela estimação Bayesiana

em cada abordagem, apresentados na Tabela 16, verificamos que para ambos os casos foi escolhida a mesma camada de saída da VGG-16. Este fato demonstra inicialmente que nem todas as camadas da rede base eram interessantes para o problema, e que ambas abordagens tiveram otimizações que levaram ao mesmo ponto de máximo. Adicionalmente, a abordagem Inception exigiu uma menor quantidade de camadas e filtros por convolução, o que torna o modelo todo mais leve para treinamento e menos sujeito a *overfit*. A abordagem Sequencial usou apenas uma camada densa, significando que foi apenas a camada de classificação com ativação sigmoid. Ambas abordagens tiveram como melhor resultado o otimizador SGD, fato possivelmente relacionado com a velocidade de atualização dos pesos da rede, que neste caso, por existir a reutilização de camadas pré-treinadas, é recomendado que seja mais lento, característica obedecida pelo SGD em comparação com os demais otimizadores.

Tabela 16 – Hiperparâmetros estimados do melhor modelo - Diagnóstico de pneumonia

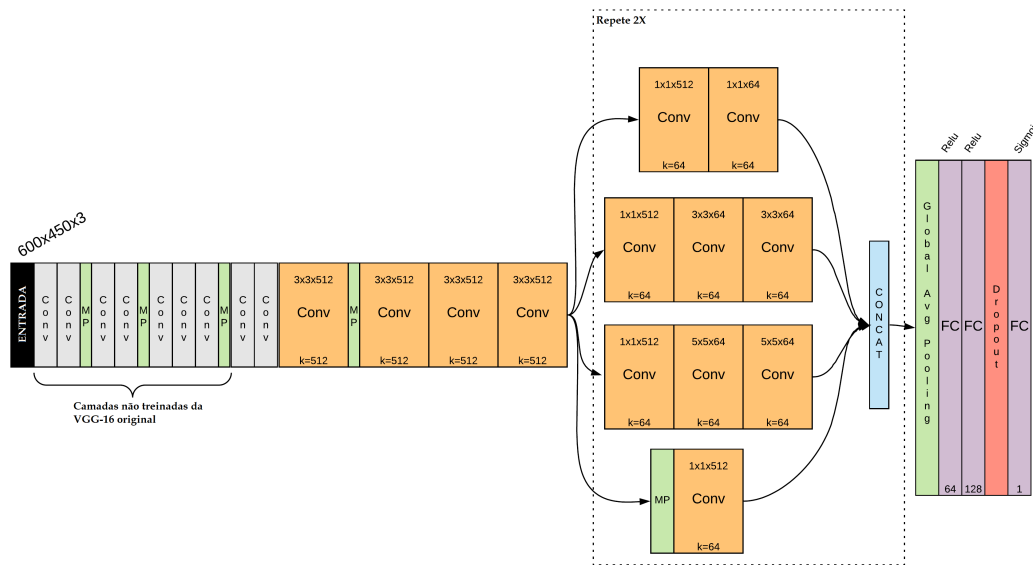
Hiperparâmetros	Abordagem	
	Sequencial	Inception
Número de camadas	4	2
Quantidade de filtros	512	64
Camada de Saída	12	12
Número de camadas Dense	1	2
Quantidade de neurônios	128	128
Dropout	0,4643	0,5724
Otimizador	SGD	SGD

Através da arquitetura otimizada proposta nesta tese, o melhor modelo gerado para o diagnóstico de pneumonia foi a VGG-16 utilizando a abordagem Inception, ilustrada na Figura 32. Esse modelo tem um total de 12.586.945 parâmetros treináveis. Comparada com a VGG-16 original (Tabela 17), houve redução de mais de 90% de parâmetros.

Tabela 17 – Comparação de Parâmetros - Diagnóstico de Pneumonia

Rede	Total de Parâmetros	Parâmetros treináveis	Parâmetros Não treináveis
VGG-16 original	138.357.544	138.357.544	0
VGG-16 otimizada	15.502.593	12.586.945	2.915.648

Figura 32 – Arquitetura otimizada (VGG-16) - Diagnóstico de Pneumonia



5.2.3 Comparação com outros trabalhos

Esta seção apresenta a comparação dos resultados obtidos com outros trabalhos estado da arte na literatura. O comparativo é feito na Tabela 18, onde todos os trabalhos apresentados utilizam a mesma base de imagens e apresentam diferentes soluções para o problema.

Pode-se observar o sucesso do uso de técnicas de aprendizado profundo para detectar pneumonia analisando os resultados apresentados na Tabela 18. A maioria dos trabalhos listados obteve acurácia maior que 90%. No diagnóstico de pneumonia, que é a tarefa de classificar entre pneumonia viral e bacteriana, a abordagem proposta nesta tese é superior às demais. Em relação ao diagnóstico, o trabalho de Rahman et al. (2020) apresenta melhor taxa de precisão e especificidade, enquanto o método proposto apresenta melhores taxas de sensibilidade, acurácia e F-Score. Porém, podemos reforçar que a arquitetura proposta nesta tese é especializada ao problema, o que garante maior adaptabilidade à variância dos exames. Essa afirmativa ocorre porque as características da rede são específicas ao problema. Ainda podemos destacar como vantagem o fato da arquitetura ser otimizada para não consumir recursos desnecessários e, portanto, supõe-se ser computacionalmente eficiente para o problema.

Tabela 18 – Comparação com outros trabalhos

Trabalho	Problema	Precisão	Sens.	Espec.	Acurácia	F-Score
Kermany et al. (2018)	Detecção	-	0,932	0,901	0,928	-
	Diagnóstico	-	0,886	0,909	0,907	-
Rajaraman et al. (2018)	Detecção	0,977	0,962	0,962	0,962	0,970
	Diagnóstico	0,920	0,984	0,860	0,936	0,951
Sousa, Fernandes e Paiva (2019)	Detecção	-	0,997	0,880	0,953	-
	Diagnóstico	-	0,913	0,696	0,831	-
Liang e Zheng (2019)	Detecção	0,891	0,967	-	0,905	0,927
	Diagnóstico	-	-	-	-	-
Saraiva et al. (2019)	Detecção	-	-	-	0,953	-
	Diagnóstico	-	-	-	-	-
Stephen et al. (2019)	Detecção	-	-	-	0,930	-
	Diagnóstico	-	-	-	-	-
Ayan e Ünver (2019)	Detecção	-	0,820	0,910	0,87	-
	Diagnóstico	-	-	-	-	-
Rahman et al. (2020)	Detecção	0,970	0,990	0,970	0,980	0,981
	Diagnóstico	0,950	0,960	0,950	0,950	0,952
Hashmi et al. (2020)	Detecção	0,982	0,990	-	0,984	-
	Diagnóstico	-	-	-	-	-
Toğaçar et al. (2020)	Detecção	0,992	0,996	0,992	0,994	0,994
	Diagnóstico	-	-	-	-	-
Moujahid et al. (2020)	Detecção	0,910	0,970	-	0,968	0,940
	Diagnóstico	-	-	-	-	-
Método Proposto	Detecção	0,971	0,984	0,953	0,968	0,974
	Diagnóstico	0,947	0,975	0,912	0,951	0,961

5.3 Experimentos para detecção e diagnóstico de COVID-19

Os experimentos realizados nesta seção utilizaram a base de imagens de raio-X publicada em Rahman (2020), que possui casos de COVID-19. A base está descrita de forma mais detalhada no Capítulo 4, Seção 2.

5.3.1 Resultados para a detecção de COVID-19

Nesta seção são apresentados os resultados encontrados pela CNN estimada para a detecção de COVID-19. Optou-se por utilizar as imagens sem a segmentação de pulmão, já que de acordo com os resultados anteriores não obteve-se melhora significativa. A etapa de construção e otimização da CNN para detecção de pneumonia foi conduzida por 50 épocas para cada uma das abordagens projetadas: Sequencial e Inception. Como as bases de detecção e diagnóstico foram previamente balanceadas, a métrica de *fitness* utilizada para otimização foi a acurácia. Os resultados estão separados de acordo com a rede base utilizada: VGG-16, ResNet-50 e DenseNet-169

Após determinar o melhor conjunto de hiperparâmetros, o melhor modelo em termos de F-score foi escolhido para cada uma das abordagens e executado com a base de teste, gerando os resultados apresentados a seguir.

Observa-se que os resultados obtidos com as abordagens Inception e Sequencial variam de acordo com a rede base utilizada. Enquanto a ResNet-50 obteve melhor resultado utilizando a abordagem Inception, a DenseNet-169 apresentou resultados melhores utilizando a abordagem Sequencial. Destaca-se a rede VGG-16 que obteve 100% de acertos em todas as métricas, utilizando ambas abordagens.

Tabela 19 – Resultados obtidos para a Detecção de COVID-19 - VGG-16

Abordagem	Precisão	Sensibilidade	Especificidade	Acurácia	F-Score
Sequencial	1	1	1	1	1
Inception	1	1	1	1	1

Tabela 20 – Resultados obtidos para a Detecção de COVID-19 - ResNet-50

Abordagem	Precisão	Sensibilidade	Especificidade	Acurácia	F-Score
Sequencial	0,975	1	0,996	0,996	0,987
Inception	1	1	1	1	1

Tabela 21 – Resultados obtidos para a Detecção de COVID-19 - DenseNet-169

Abordagem	Precisão	Sensibilidade	Especificidade	Acurácia	F-Score
Sequencial	1	0,923	1	0,990	0,960
Inception	1	0,743	1	0,968	0,852

Utilizando a VGG-16 como rede base, o otimizador conseguiu encontrar várias combinações de modelo ótimos (modelos com 100% de acertos). Os hiperparâmetros encontrados por um desses melhores modelos estão descritos na Tabela 22. A ResNet-50 otimizada utilizando a abordagem Inception também encontrou um modelo com 100% de acerto em todas as taxas. Os hiperparâmetros encontrados por esse modelo ótimo de ResNet-50 estão listados na Tabela 23.

Utilizando a arquitetura otimizada proposta nesta tese, o melhor modelo gerado para a detecção da COVID-19 foi tanto a rede VGG-16 utilizando ambas abordagens, quanto a ResNet-50 utilizando a abordagem Inception. As Tabelas 24 e 25 mostram o comparativo da redução de parâmetros.

Tabela 22 – Hiperparâmetros estimados pelos melhores modelos da VGG-16 - Detecção COVID-19.

Hiperparâmetros	Abordagem	
	Sequencial	Inception
Número de Camadas	2	3
Quantidade de filtros	32	16
Camada de Saída	14	10
Número de camadas Dense	2	1
Quantidade de neurônios	256	256
Dropout	0,4836	0,4180
Otimizador	RMSprop	RMSprop

Tabela 23 – Hiperparâmetros estimados pelo melhor modelo da ResNet-50 - Detecção COVID-19.

Hiperparâmetros	Abordagem Inception
Número de Camadas	0
Quantidade de filtros	256
Camada de Saída	49
Número de camadas Dense	2
Quantidade de neurônios	256
Dropout	0,5993
Otimizador	RMSprop

Tabela 24 – Comparação de Parâmetros - Detecção COVID-19 - VGG-16.

Abordagem	Rede	Total de Parâmetros	Parâmetros treináveis	Parâmetros Não treináveis
Sequencial	VGG-16 original	138.357.544	138.357.544	0
	VGG-16 otimizada	14.922.273	207.585	14.714.688
Inception	VGG-16 original	138.357.544	138.357.544	0
	VGG-16 otimizada	14.825.969	111.281	14.714.688

5.3.2 Resultados para o diagnóstico de pneumonia (incluindo COVID-19)

Nesta seção são apresentados os resultados encontrados pela CNN estimada para o diagnóstico de pneumonia utilizando a base de COVID-19. Nesta tarefa, a rede irá classificar as imagens de raio-X em normal, pneumonia viral ou COVID-19. Da mesma forma que na tarefa de detecção da COVID-19, optou-se por utilizar as imagens sem a segmentação de pulmão, já que de acordo com os resultados anteriores não obteve-se melhora significativa. A etapa de construção e otimização da CNN para o diagnóstico também foi conduzida por 50 épocas para cada uma das abordagens projetadas: Sequencial

Tabela 25 – Comparação de Parâmetros - Detecção COVID-19 - ResNet-50.

Rede	Total de Parâmetros	Parâmetros treináveis	Parâmetros Não treináveis
ResNet-50 original	25.636.712	25.636.712	0
ResNet-50 otimizada	23.883.265	295.553	23.587.712

e Inception. Como as bases de detecção e diagnóstico foram previamente balanceadas, a métrica de *fitness* utilizada para otimização foi a acurácia. Os resultados estão separados de acordo com a rede base utilizada: VGG-16, ResNet-50 e DenseNet-169

Após determinar o melhor conjunto de hiperparâmetros, o melhor modelo em termos de Acurácia foi escolhido para cada uma das abordagens e executado com a base de teste, gerando os resultados apresentados nas tabelas a seguir.

O melhor resultado encontrado para a classificação multiclasse (COVID-19, Pneumonia Viral e Normal) foi utilizando a abordagem Sequencial com a DenseNet-169 como rede base, atingindo acurácia de 0,993. A abordagem Sequencial apresentou melhores resultados quando utilizada com as redes VGG-16 e DenseNet-169. Já com a ResNet-50, os resultados obtidos foram superiores com a abordagem Inception. Destaca-se a alta taxa de sensibilidade no diagnóstico da COVID-19 com a abordagem Sequencial, que foi de 100% em todas as redes base utilizadas.

Tabela 26 – Resultados obtidos para o Diagnóstico (COVID-19) - VGG-16

Abordagem	Diagnóstico	Acurácia	Precisão	Sensibilidade	F-Score
Sequencial	COVID-19	0,989	1	1	1
	Pneumonia Viral		1	0,977	0,988
	Normal		0,978	1	0,989
Inception	COVID-19	0,987	1	1	1
	Pneumonia Viral		0,996	0,977	0,986
	Normal		0,978	0,996	0,987

Tabela 27 – Resultados obtidos para o Diagnóstico (COVID-19)- ResNet-50

Abordagem	Diagnóstico	Acurácia	Precisão	Sensibilidade	F-Score
Sequencial	COVID-19	0,951	0,906	1	0,951
	Pneumonia Viral		0,928	0,981	0,954
	Normal		0,984	0,916	0,949
Inception	COVID-19	0,953	0,883	0,974	0,926
	Pneumonia Viral		0,980	0,924	0,951
	Normal		0,940	0,978	0,959

Tabela 28 – Resultados obtidos para o Diagnóstico (COVID-19) - DenseNet-169

Abordagem	Diagnóstico	Acurácia	Precisão	Sensibilidade	F-Score
Sequencial	COVID-19	0,993	1	1	1
	Pneumonia Viral		1	0,984	0,992
	Normal		0,985	1	0,992
Inception	COVID-19	0,702	0,279	0,923	0,428
	Pneumonia Viral		0,783	0,721	0,751
	Normal		0,869	0,652	0,745

O melhor modelo utilizando a abordagem Sequencial usou a DenseNet-169 como rede base, já o melhor modelo utilizando a abordagem Inception utilizou a VGG-16 como rede base. Os hiperparâmetros encontrados pelos melhores modelos das abordagens Inception e Sequencial estão descritos na Tabela 29.

Tabela 29 – Hiperparâmetros estimados pelo melhor modelo - Diagnóstico COVID-19

Hiperparâmetros	Abordagem	
	Sequencial	Inception
Número de Camadas	0	3
Quantidade de filtros	32	64
Camada de Saída	166	11
Número de camadas Dense	2	1
Quantidade de neurônios	256	256
Dropout	0,5203	0,4553
Otimizador	SGD	SGD

Através da arquitetura otimizada proposta nesta tese, o melhor modelo gerado para o diagnóstico de COVID-19 foi a DenseNet-169 utilizando a abordagem Sequencial. Esse modelo tem um total de 246.915 parâmetros treináveis. Comparada com a DenseNet-169 original (Tabela 30), houve redução em torno de 98% de parâmetros.

Tabela 30 – Comparação de Parâmetros - Diagnóstico COVID-19

Rede	Total de Parâmetros	Parâmetros treináveis	Parâmetros Não treináveis
DenseNet-169 original	14.307.880	14.307.880	0
DenseNet-169 otimizada	12.889.795	246.915	12.642.880

5.3.3 Comparação com outros trabalhos

Nesta seção é feita a comparação dos resultados obtidos com outros trabalhos estado da arte na literatura. O comparativo é feito na Tabela 31, onde a maioria dos trabalhos

apresentados utilizam a mesma versão da mesma base de imagens fornecida por Rahman (2020). A única exceção são os trabalhos de Gupta et al. (2021) e Quan et al. (2021). Gupta et al. (2021) utiliza duas bases: a base fornecida por Rahman (2020) e a base fornecida por Maguolo e Nanni (2020) e Quan et al. (2021) juntou três bases: Cohen, Morrison e Dao (2020), Praveen (2020) e Rahman (2020).

Observando os resultados listados na Tabela 31, verifica-se que tanto no problema de detecção de COVID-19 quanto no de diagnóstico, a abordagem proposta mostrou-se promissora. Na tarefa de detecção da COVID-19, a abordagem proposta alcançou a solução ótima, atingindo 100% em todas as métricas. Na tarefa de diagnóstico, que é classificar as imagens de raio-X em normal, pneumonia viral e COVID-19, o método proposto também apresenta resultados bastante satisfatórios, obtendo os melhores resultados em todas as métricas: Precisão, Sensibilidade, Acurácia e F-Score.

Dessa forma, os resultados encontrados com a realização dos experimentos de detecção e diagnóstico corroboram o sucesso do uso da inteligência artificial no auxílio aos profissionais da saúde, melhorando a velocidade e a acurácia em diagnósticos de doenças.

Tabela 31 – Comparação com outros trabalhos

Trabalho	Problema	Prec.	Sens.	Espec.	Acurácia	F-Score
Ouchicha e Ammor (2020)	Detecção	-	-	-	-	-
	Diagnóstico	96,72	96,84	-	96,69	96,68
Nour, Cömert e Polat (2020)	Detecção	-	-	-	-	-
	Diagnóstico	-	89,39	99,75	98,97	96,72
Chowdhury et al. (2020)	Detecção	99,70	99,70	99,55	99,70	99,70
	Diagnóstico	97,95	97,94	-	97,94	97,94
Aslan et al. (2021)	Detecção	-	-	-	-	-
	Diagnóstico	98,77	98,76	-	98,70	98,76
Quan et al. (2021)	Detecção	-	97,00	100	98,50	-
	Diagnóstico	91,10	90,70	-	90,70	90,90
Gupta et al. (2021)	Detecção	99,50	99,50	-	99,53	99,50
	Diagnóstico	99,00	99,00	-	99,08	99,00
Método Proposto	Detecção	100	100	100	100	100
	Diagnóstico	99,32	99,31	-	99,31	99,31

6 Conclusão

A pneumonia ainda é uma doença com alta taxa de mortalidade e morbidade em todo o mundo e o surgimento de um novo tipo de pneumonia viral agravou ainda mais essas taxas. A COVID-19 se espalhou rapidamente em todo os países, levando a um surto global. A rápida disseminação do vírus SARS-CoV-2 faz com que o diagnóstico precoce seja extremamente importante. O sucesso do uso de inteligência artificial em vários problemas na área da saúde na última década reforçam o potencial que a IA tem para contribuir no controle e diagnóstico da COVID-19.

Como forma de auxiliar a tomada de decisão médica, este trabalho apresentou uma metodologia para criação de arquiteturas otimizadas para a detecção e diagnóstico automático dos diferentes tipos de pneumonia.

A técnica utilizada neste trabalho foi de aprendizagem profunda, que tem sido utilizada com sucesso para a classificação e reconhecimento de imagens na literatura. Os resultados obtidos pelo método proposto são promissores, chegando a atingir mais de 95% de acurácia no diagnóstico entre os tipos normal, bacteriano, viral e COVID-19.

O método construído para atingir estes resultados avaliou a aplicação de otimização Bayesiana para estimação da topologia e hiperparâmetros da rede neural convolucional construída especificamente como solução proposta. A solução proposta possui taxa de assertividade promissora, o que indica que o método construído possui eficiência juntamente com desempenho.

A grande vantagem da metodologia proposta é a geração de arquiteturas eficientes e especializadas ao problema, mas é importante informar que a otimização dos hiperparâmetros foi limitada ao *hardware* disponível. As redes mais profundas aqui testadas, como ResNet e DenseNet, algumas vezes não conseguiram finalizar o treinamento com os hiperparâmetros selecionados, que necessitavam de um *hardware* mais poderoso. Os resultados alcançados por essas redes não necessariamente refletem o melhor conjunto estimado, mas sim o melhor conjunto que o *hardware* disponível conseguiu treinar. Ainda assim, comparativamente com outros trabalhos, os resultados obtidos são próximos ou melhores ao estado da arte.

Esta tese demonstrou a eficiência da estimação de arquiteturas de CNNs para detecção e diagnóstico entre pneumonias usando otimização bayesiana. Em trabalhos futuros, pretende-se combinar outras arquiteturas como forma de aprender relações residuais

de padrões em busca da melhoria de resultados. É válido também investigar outras abordagens de otimização diferentes da otimização bayesiana, além de validar o método de segmentação em outras bases que também não possuem a marcação do especialista.

Validando a pesquisa feita até aqui, a Tabela 32 lista o artigos científicos produzidos diretamente relacionados a Tese, considerando os resultados obtidos até o momento.

Tabela 32 – Artigos produzidos diretamente relacionados a Tese

Local	Artigo	Qualis	Status
Congresso	G. Sousa, V. Fernandes and A. Paiva "Optimized Deep Learning Architecture for the Diagnosis of Pneumonia through Chest X-Rays". 16th International Conference on Image Analysis and Recognition - ICIAR 2019	B1	Publicado
Revista	Vandecia Fernandes, Geraldo Braz Junior, Anselmo Paiva, Aristófanés Silva and Marcelo Gattass "Bayesian Convolutional Neural Network Estimation for Pediatric Pneumonia Detection and Diagnosis". Computer Methods and Programs in Biomedicine	A1	Publicado
Congresso	Gabriel Bras, Vandecia Fernandes, Anselmo Cardoso de Paiva, Geraldo Braz Júnior and Luis Rivero. "Transfer Learning Method Evaluation for Automatic Pediatric Chest X-Ray Image Segmentation". 27th International Conference on Systems, Signals and Image Processing - IWSSIP 2020	B1	Publicado
Revista	Vandecia Fernandes, Gabriel Bras, Lisle Faray de Paiva, Geraldo Braz Junior, Anselmo Cardoso de Paiva and Luis Rivero. "Transfer Learning Based Lung Segmentation and Pneumonia Detection for Pediatric Chest X-Ray Images". International Journal of Innovative Computing and Applications	B2	Aceito para publicação

Referências

- ABBAS, Q. Glaucoma-deep: detection of glaucoma eye disease on retinal fundus images using deep learning. *Int J Adv Comput Sci Appl*, v. 8, n. 6, p. 41–5, 2017.
- AHMAD, J.; FARMAN, H.; JAN, Z. Deep learning methods and applications. In: *Deep Learning: Convergence to Big Data Analytics*. USA: Now Publishers, 2019. p. 31–42.
- ALBUQUERQUE, C. *O impacto da COVID-19 nas internações do sistema único de saúde – SUS*. 2020. Disponível em: <<https://www.observatoriohospitalar.fiocruz.br/conteudo-interno/o-impacto-da-covid-19-nas-internacoes-do-sistema-unico-de-saude-sus>>. Acesso em: 24 de julho de 2021.
- AMERICAN LUNG ASSOCIATION. *Pneumonia*. 2019. Disponível em: <<https://www.lung.org/lung-health-and-diseases/lung-disease-lookup/pneumonia/>>. Acesso em: 15 de maio de 2019.
- AMERICAN THORACIC SOCIETY. *Top 20 Pneumonia Facts—2018*. 2018. Disponível em: <<https://www.thoracic.org/patients/patient-resources/resources/top-pneumonia-facts.pdf>>. Acesso em: 20 de maio de 2019.
- ANTIN, B.; KRAVITZ, J.; MARTAYAN, E. Detecting pneumonia in chest x-rays with supervised learning. *Semanticscholar.org*, 2017.
- APORRAS. *What is the difference between parameters and hyperparameters?* 2020. Disponível em: <<https://quantdare.com/what-is-the-difference-between-parameters-and-hyperparameters>>. Acesso em: 20 de março de 2021.
- ARDILA, D. et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nature medicine*, Nature Publishing Group, v. 25, n. 6, p. 954–961, 2019.
- ASAOKA, R. et al. Using deep learning and transfer learning to accurately diagnose early-onset glaucoma from macular optical coherence tomography images. *American journal of ophthalmology*, Elsevier, v. 198, p. 136–145, 2019.
- ASLAN, M. F. et al. Cnn-based transfer learning–bilstm network: A novel approach for covid-19 infection detection. *Applied Soft Computing*, v. 98, p. 106912, 2021. ISSN 1568-4946. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1568494620308504>>.
- ASNAOUI, K. E.; CHAWKI, Y. Using x-ray images and deep learning for automated detection of coronavirus disease. *Journal of Biomolecular Structure and Dynamics*, Taylor & Francis, p. 1–12, 2020.
- AYAN, E.; ÜNVER, H. M. Diagnosis of pneumonia from chest x-ray images using deep learning. *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*, p. 1–5, 2019.
- BENGIO, Y. Deep learning of representations for unsupervised and transfer learning. In: *Proceedings of ICML Workshop on Unsupervised and Transfer Learning*. Bellevue, Washington, USA: PMLR, 2012. (Proceedings of Machine Learning Research, v. 27), p. 17–36. Disponível em: <<http://proceedings.mlr.press/v27/bengio12a.html>>.

- BENGIO, Y. et al. Learning deep architectures for ai. *Foundations and trends® in Machine Learning*, Now Publishers, Inc., v. 2, n. 1, p. 1–127, 2009.
- BERGSTRA, J.; BENGIO, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.*, JMLR.org, v. 13, p. 281–305, fev. 2012. ISSN 1532-4435. Disponível em: <<http://dl.acm.org/citation.cfm?id=2188385.2188395>>.
- BERGSTRA, J. et al. Hyperopt: a python library for model selection and hyperparameter optimization. *Computational Science & Discovery*, v. 8, n. 1, p. 014008, 2013. Disponível em: <<http://stacks.iop.org/1749-4699/8/i=1/a=014008>>.
- BERGSTRA, J. S. et al. Algorithms for hyper-parameter optimization. In: *Advances in Neural Information Processing Systems 24*. Curran Associates, Inc., 2011. p. 2546–2554. Disponível em: <<http://papers.nips.cc/paper/4443-algorithms-for-hyper-parameter-optimization.pdf>>.
- BERTRAND, H.; HASHIR, M.; COHEN, J. P. Do lateral views help automated chest x-ray predictions? *arXiv preprint arXiv:1904.08534*, 2019.
- BRAZDIL, P. B.; SOARES, C. A comparison of ranking methods for classification algorithm selection. In: SPRINGER. *European conference on machine learning*. [S.l.], 2000. p. 63–75.
- BRENTANO, V. B. et al. Interpretando a radiografia de tórax na emergência. *Acta méd.(Porto Alegre)*, v. 35, p. 8–8, 2014.
- BRILLIANT.ORG. *Feedforward Neural Networks*. 2018. Disponível em: <<https://brilliant.org/wiki/feedforward-neural-networks/>>. Acesso em: 15 de setembro de 2018.
- BROCHU, E. *Interactive Bayesian optimization: learning user preferences for graphics and animation*. Tese (Doutorado) — University of British Columbia, 2010.
- BROCHU, E.; CORA, V. M.; FREITAS, N. D. A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *arXiv preprint arXiv:1012.2599*, 2010.
- BROWNLEE, J. *What Is Meta-Learning in Machine Learning?* 2020. Disponível em: <<https://machinelearningmastery.com/meta-learning-in-machine-learning>>. Acesso em: 20 de março de 2021.
- BRUNESE, L. et al. Explainable deep learning for pulmonary disease and coronavirus covid-19 detection from x-rays. *Computer Methods and Programs in Biomedicine*, Elsevier, v. 196, p. 105608, 2020.
- BUSTOS, A. et al. Padchest: A large chest x-ray image dataset with multi-label annotated reports. *Medical Image Analysis*, v. 66, p. 101797, 2020. ISSN 1361-8415. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1361841520301614>>.
- CAPONE, D. et al. Diagnóstico por imagem na pneumonia por covid-19. *Pulmão RJ*, v. 29, n. 1, p. 22–27, 2020.

CERENTINI, A. et al. Automatic identification of glaucoma using deep learning methods. *Studies in Health Technology and Informatics*, PubMed, v. 2017, 2017. Disponível em: <<https://www.ncbi.nlm.nih.gov/pubmed/29295107>>.

CHAURASIA, A.; CULURCIELLO, E. Linknet: Exploiting encoder representations for efficient semantic segmentation. *2017 IEEE Visual Communications and Image Processing (VCIP)*, p. 1–4, 2017.

CHEN, S.; SUZUKI, K. Separation of bones from chest radiographs by means of anatomically specific multiple massive-training anns combined with total variation minimization smoothing. *IEEE transactions on medical imaging*, IEEE, v. 33, n. 2, p. 246–257, 2013.

CHOLLET, F. Xception: Deep learning with depthwise separable convolutions. *Proceedings of the IEEE conference on computer vision and pattern recognition*, p. 1251–1258, 2017.

CHOLLET, F. et al. *Keras*. GitHub, 2015. Disponível em: <<https://github.com/fchollet/keras>>.

CHOUHAN, V. et al. A novel transfer learning based approach for pneumonia detection in chest x-ray images. *Applied Sciences*, Multidisciplinary Digital Publishing Institute, v. 10, n. 2, p. 559, 2020.

CHOWDHURY, M. E. et al. Can ai help in screening viral and covid-19 pneumonia? *IEEE Access*, IEEE, v. 8, p. 132665–132676, 2020.

CHUNG, A. Actualmed covid-19 chest x-ray data initiative. 2020. Disponível em: <<https://github.com/agchung/Actualmed-COVID-chestxray-dataset>>.

CLAESEN, M.; MOOR, B. D. Hyperparameter search in machine learning. *CoRR*, abs/1502.02127, 2015.

COCCIA, M. Deep learning technology for improving cancer care in society: New directions in cancer imaging driven by artificial intelligence. *Technology in Society*, Elsevier, v. 60, p. 101198, 2020.

COHEN, J. P.; MORRISON, P.; DAO, L. Covid-19 image data collection. *arXiv 2003.11597*, 2020. Disponível em: <<https://github.com/ieee8023/covid-chestxray-dataset>>.

CORRÊA, R. d. A. et al. Recomendações para o manejo da pneumonia adquirida na comunidade 2018. *Jornal Brasileiro de Pneumologia*, SciELO Brasil, v. 44, n. 5, p. 405–423, 2018.

COSTA, A. *Pneumonia*. 2019. Disponível em: <<https://www.saudebemestar.pt/pt/medicina/pneumologia/pneumonia/>>. Acesso em: 20 de maio de 2019.

DADONAITE, B.; ROSER, M. *Pneumonia - Our World in Data*. 2020. Disponível em: <<https://ourworldindata.org/pneumonia>>. Acesso em: 10 de janeiro de 2020.

DAS, N. N. et al. Automated deep transfer learning-based approach for detection of covid-19 infection in chest x-rays. *Irbm*, Elsevier, 2020.

- DELRUE, L. et al. Difficulties in the interpretation of chest radiography. *Comparative Interpretation of CT and Standard Radiography of the Chest*, Springer, p. 27–49, 2011.
- ELAZIZ, M. A. et al. New machine learning method for image-based diagnosis of covid-19. *Plos one*, Public Library of Science San Francisco, CA USA, v. 15, n. 6, p. e0235187, 2020.
- ESTEVA, A. et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, Nature Publishing Group, v. 542, n. 7639, p. 115, 2017.
- ESTEVA, A. et al. A guide to deep learning in healthcare. *Nature medicine*, Nature Publishing Group, v. 25, n. 1, p. 24, 2019.
- FRAZIER, P. I. A tutorial on bayesian optimization. *arXiv preprint arXiv:1807.02811*, 2018.
- GINNEKEN, B. V.; STEGMANN, M. B.; LOOG, M. Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database. *Medical image analysis*, Elsevier, v. 10, n. 1, p. 19–40, 2006.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. Cambridge, Massachusetts, USA: MIT Press, 2016. Disponível em: <<http://www.deeplearningbook.org>>.
- GUPTA, A. et al. Instacovnet-19: A deep learning classification model for the detection of covid-19 patients using chest x-ray. *Applied Soft Computing*, v. 99, p. 106859, 2021. ISSN 1568-4946. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1568494620307973>>.
- GUPTA, N. *Difference Between Model Parameters VS HyperParameters*. 2020. Disponível em: <<https://www.geeksforgeeks.org/difference-between-model-parameters-vs-hyperparameters>>. Acesso em: 20 de março de 2021.
- HALAJI, M. et al. Viral etiology of acute respiratory infections in children in Southern Iran. *Revista da Sociedade Brasileira de Medicina Tropical*, scielo, v. 52, 2019. ISSN 0037-8682. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0037-86822019000100682&nrm=iso>.
- HASHMI, M. F. et al. Efficient pneumonia detection in chest xray images using deep transfer learning. *Diagnostics*, Multidisciplinary Digital Publishing Institute, v. 10, n. 6, p. 417, 2020.
- HAVAEI, e. a. M. Brain tumor segmentation with deep neural networks. *Medical Image Analysis*, v. 35, p. 18 – 31, 2017. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1361841516300330>>.
- HE, J. et al. The practical implementation of artificial intelligence technologies in medicine. *Nature medicine*, Nature Publishing Group, v. 25, n. 1, p. 30, 2019.
- HE, K. et al. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015. Disponível em: <<http://arxiv.org/abs/1512.03385>>.
- HINTON, G. E. A practical guide to training restricted boltzmann machines. *Neural Networks: Tricks of the Trade: Second Edition*, Springer Berlin Heidelberg, Berlin, Heidelberg, p. 599–619, 2012. Disponível em: <https://doi.org/10.1007/978-3-642-35289-8_32>.

HSU, C.-W.; CHANG, C.-C.; LIN, C.-J. A practical guide to support vector classification. Taipei, 2003.

HUANG, G. et al. Densely connected convolutional networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*, p. 4700–4708, 2017.

ICHI.PRO. *ResNet: uma compreensão simples das redes residuais*. 2020. Disponível em: <<https://ichi.pro/pt/resnet-uma-compreensao-simples-das-redes-residuais-210937146811463>>. Acesso em: 15 de março de 2021.

IRVIN, J. et al. Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. *Proceedings of the AAAI Conference on Artificial Intelligence*, v. 33, n. 01, p. 590–597, 2019.

JAEGER, S. et al. Two public chest x-ray datasets for computer-aided screening of pulmonary diseases. *Quantitative imaging in medicine and surgery*, AME Publications, v. 4, n. 6, p. 475, 2014.

JIA, G.; LAM, H.-K.; XU, Y. Classification of covid-19 chest x-ray and ct images using a type of dynamic cnn modification method. *Computers in Biology and Medicine*, v. 134, p. 104425, 2021. ISSN 0010-4825. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0010482521002195>>.

JOHNSON, A. E. et al. Mimic-cxr-jpg, a large publicly available database of labeled chest radiographs. *arXiv preprint arXiv:1901.07042*, 2019.

JORDAN, J. *Common architectures in convolutional neural networks*. 2018. Disponível em: <<https://www.jeremyjordan.me/convnet-architectures>>. Acesso em: 29 de março de 2021.

KATHER, J. N. et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nature medicine*, Nature Publishing Group, v. 25, n. 7, p. 1054–1056, 2019.

KERMANY, D. S. et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*, Elsevier, v. 172, n. 5, p. 1122–1131, 2018.

KHAN, A. I.; SHAH, J. L.; BHAT, M. M. Coronet: A deep neural network for detection and diagnosis of covid-19 from chest x-ray images. *Computer Methods and Programs in Biomedicine*, Elsevier, v. 196, p. 105581, 2020.

KHANNA, M. *Paper review: DenseNet -Densely Connected Convolutional Networks*. 2019. Disponível em: <<https://towardsdatascience.com/paper-review-densenet-densely-connected-convolutional-networks-acf9065dfefb>>. Acesso em: 30 de março de 2021.

KLAUS SCHWAB. *World Economic Forum*. 2016. Disponível em: <<https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/>>. Acesso em: 01 de fevereiro de 2020.

KOCAK, M. *Radiografias simples*. 2019. Disponível em: <<https://www.msmanuals.com/pt/casa/assuntos-especiais/exames-de-diagn%C3%B3stico-por-imagem-comuns/radiografias-simples>>. Acesso em: 10 de fevereiro de 2020.

- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, p. 1097–1105, 2012.
- KUMAR, P.; GREWAL, M.; SRIVASTAVA, M. M. Boosted cascaded convnets for multilabel classification of thoracic diseases in chest radiographs. *International Conference Image Analysis and Recognition*, p. 546–552, 2018.
- LAKHANI, P.; SUNDARAM, B. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology*, Radiological Society of North America, v. 284, n. 2, p. 574–582, 2017.
- LAKSHMANAPRABU, S. et al. Optimal deep learning model for classification of lung cancer on ct images. *Future Generation Computer Systems*, Elsevier, v. 92, p. 374–382, 2019.
- LECUN, Y.; BENGIO, Y. Convolutional networks for images, speech, and time-series. *The Handbook of Brain Theory and Neural Networks*, p. 14, 01 1995.
- LECUN, Y. et al. Backpropagation applied to handwritten zip code recognition. *Neural Computation*, v. 1, n. 4, p. 541–551, 1989.
- LECUN, Y. et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, IEEE, v. 86, n. 11, p. 2278–2324, 1998.
- LEMKE, C.; BUDKA, M.; GABRYS, B. Metalearning: a survey of trends and technologies. *Artificial Intelligence Review*, v. 44, n. 1, p. 117–130, Jun 2015. ISSN 1573-7462. Disponível em: <<https://doi.org/10.1007/s10462-013-9406-y>>.
- LI, Z. et al. Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. *Ophthalmology*, Elsevier, v. 125, n. 8, p. 1199–1206, 2018.
- LIANG, G.; ZHENG, L. A transfer learning method with deep residual network for pediatric pneumonia diagnosis. *Computer Methods and Programs in Biomedicine*, Elsevier, 2019.
- MA, J. et al. Group-attention single-shot detector (ga-ssd): Finding pulmonary nodules in large-scale ct images. *arXiv preprint arXiv:1812.07166*, 2018.
- MAGUOLO, G.; NANNI, L. *A Critic Evaluation of Methods for COVID-19 Automatic Detection from X-Ray Images*. 2020.
- MENEGOLA, A. et al. Knowledge transfer for melanoma screening with deep learning. In: IEEE. *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. Melbourne, VIC, Australia, 2017. p. 297–300.
- MINISTÉRIO DA SAÚDE. *Sistema de Informações Hospitalares do SUS (SIH/SUS)*. 2020. Disponível em: <<http://tabnet.datasus.gov.br/cgi/deftohtm.exe?sih/cnv/qiuf.def>>. Acesso em: 01 de fevereiro de 2020.
- MINISTÉRIO DA SAÚDE. *Coronavírus*. 2021. Disponível em: <<https://coronavirus.saude.gov.br/index.php/sobre-a-doenca>>. Acesso em: 15 de março de 2021.

- MIRUNALINI, P. et al. Deep learning for skin lesion classification. *Computer Vision and Pattern Recognition (cs.CV)*, 2017.
- MISRA, S. et al. Multi-channel transfer learning of chest x-ray images for screening of covid-19. *Electronics*, Multidisciplinary Digital Publishing Institute, v. 9, n. 9, p. 1388, 2020.
- MONSHI, M. M. A. et al. Covidxraynet: Optimizing data augmentation and cnn hyperparameters for improved covid-19 detection from cxr. *Computers in Biology and Medicine*, v. 133, p. 104375, 2021. ISSN 0010-4825. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0010482521001694>>.
- MORSCH, J. *Raio x Tórax com Laudo: entenda para que serve, indicação e resultados*. 2019. Disponível em: <<https://telemedicinamorsch.com.br/blog/raio-x-torax-com-laudo>>. Acesso em: 10 de fevereiro de 2020.
- MOUJAHID, H. et al. Convolutional neural network based classification of patients with pneumonia using x-ray lung images. *Transfer*, v. 2, n. 99.41, p. 16, 2020.
- MURPHY, K. et al. Computer aided detection of tuberculosis on chest radiographs: An evaluation of the cad4tb v6 system. *arXiv preprint arXiv:1903.03349*, 2019.
- NASRABADI, N. M. Pattern recognition and machine learning. *Journal of electronic imaging*, International Society for Optics and Photonics, v. 16, n. 4, p. 049901, 2007.
- NG, A. Y.-T. *What data scientists should know about deep learning*. 2015. Disponível em: <<https://www.slideshare.net/ExtractConf>>. Acesso em: 18 de setembro de 2018.
- NIELSEN, M. *Neural Networks and Deep Learning*. 2018. Disponível em: <<http://neuralnetworksanddeeplearning.com/index.html>>. Acesso em: 18 de Outubro de 2018.
- NISHIOKA, S. *Sete coronavírus causam doenças em humanos*. 2020. Disponível em: <<https://www.unasus.gov.br/especial/covid19/markdown/111>>. Acesso em: 24 de julho de 2021.
- NOUR, M.; CÖMERT, Z.; POLAT, K. A novel medical diagnosis model for covid-19 infection detection based on deep features and bayesian optimization. *Applied Soft Computing*, Elsevier, v. 97, p. 106580, 2020.
- NUNES, D. H. e. R. A. I. *Redes neurais artificiais: para engenharias aplicadas*. 2. ed. São Paulo: Artliber, 2016.
- OUCHICHA, C.; AMMOR, O. Cvdnet: A novel deep learning architecture for detection of coronavirus (covid-19) from chest x-ray images. *Chaos, Solitons & Fractals*, Elsevier, v. 140, p. 110245, 2020.
- 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.
- PARVEENA, M. M. S. N. R. S. Detection of pneumonia in chest x-ray images. *Journal of X-Ray Science and Technology* 19, 2011.

- PEDREGOSA, F. et al. Scikit-learn: Machine learning in python. *J. Mach. Learn. Res.*, JMLR.org, v. 12, p. 2825–2830, nov. 2011. ISSN 1532-4435. Disponível em: <<http://dl.acm.org/citation.cfm?id=1953048.2078195>>.
- PELIKAN, M.; GOLDBERG, D. E.; CANTÚ-PAZ, E. Boa: The bayesian optimization algorithm. In: *Proceedings of the 1st Annual Conference on Genetic and Evolutionary Computation - Volume 1*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1999. (GECCO'99), p. 525–532. ISBN 1-55860-611-4. Disponível em: <<http://dl.acm.org/citation.cfm?id=2933923.2933973>>.
- PEREIRA, L. *Aprendizagem*. 2015. Disponível em: <<https://www.infoescola.com/educacao/aprendizagem>>. Acesso em: 20 de março de 2021.
- PEREIRA, R. M. et al. Covid-19 identification in chest x-ray images on flat and hierarchical classification scenarios. *Computer Methods and Programs in Biomedicine*, Elsevier, v. 194, p. 105532, 2020.
- PRAVEEN. Coronahack: chest x-ray-dataset. 2020. Disponível em: <<https://www.kaggle.com/praveengovi/coronahack-chest-xraydataset>>.
- QUAN, H. et al. Denscapsnet: Detection of covid-19 from x-ray images using a capsule neural network. *Computers in Biology and Medicine*, v. 133, p. 104399, 2021. ISSN 0010-4825. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0010482521001931>>.
- RAHIMZADEH, M.; ATTAR, A. A modified deep convolutional neural network for detecting covid-19 and pneumonia from chest x-ray images based on the concatenation of xception and resnet50v2. *Informatics in Medicine Unlocked*, Elsevier, v. 19, p. 100360, 2020.
- RAHMAN, T. Covid-19 radiography database. 2020. Disponível em: <<https://www.kaggle.com/tawsifurrahman/covid19-radiography-database>>.
- RAHMAN, T. et al. Transfer learning with deep convolutional neural network (cnn) for pneumonia detection using chest x-ray. *Applied Sciences*, Multidisciplinary Digital Publishing Institute, v. 10, n. 9, p. 3233, 2020.
- RAJARAMAN, S. et al. Visualization and interpretation of convolutional neural network predictions in detecting pneumonia in pediatric chest radiographs. *Applied Sciences*, Multidisciplinary Digital Publishing Institute, v. 8, n. 10, p. 1715, 2018.
- RAJPURKAR, P. et al. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *CoRR*, abs/1711.05225, 2017. Disponível em: <<http://arxiv.org/abs/1711.05225>>.
- RAOOF, S. et al. Interpretation of plain chest roentgenogram. *Chest*, Elsevier, v. 141, n. 2, p. 545–558, 2012.
- RAWAT, W.; WANG, Z. Deep convolutional neural networks for image classification: A comprehensive review. *Neural Computation*, v. 29, n. 9, p. 2352–2449, 2017. PMID: 28599112.

- RICE, J. R. The algorithm selection problem. *Advances in computers*, Elsevier, v. 15, p. 65–118, 1976.
- ROHILLA, A.; HOODA, R.; MITTAL, A. Tb detection in chest radiograph using deep learning architecture. *ICETETSM-17*, p. 136–147, 2017.
- RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention*, p. 234–241, 2015.
- RSNA. RsnA pneumonia detection challenge. 2019. Disponível em: <<https://www.kaggle.com/c/rsna-pneumonia-detection-challenge>>.
- RUBIN, J. et al. Large scale automated reading of frontal and lateral chest x-rays using dual convolutional neural networks. *arXiv preprint arXiv:1804.07839*, 2018.
- SAHLSTEN, J. et al. Deep learning fundus image analysis for diabetic retinopathy and macular edema grading. *Scientific reports*, Nature Publishing Group, v. 9, n. 1, p. 1–11, 2019.
- SARAIVA, A. et al. Classification of images of childhood pneumonia using convolutional neural networks. p. 112–119, 2019.
- SAUDE BEM ESTAR. *Radiografia do Torax*. 2011. Disponível em: <<https://www.saudebemestar.pt/pt/exame/imagiologia/radiografia-de-torax/>>. Acesso em: 01 de maio de 2019.
- SAVE THE CHILDREN. *Pneumonia to Kill Nearly 11 Million Children by 2030*. 2018. Disponível em: <<https://www.savethechildren.org/us/about-us/media-and-news/2018-press-releases/pneumonia-to-kill-nearly-11-million-children-by-2030/>>. Acesso em: 20 de maio de 2019.
- SCHWARTZMANN, P. V. et al. Pneumonia comunitária e pneumonia hospitalar em adultos. *Medicina (Ribeirao Preto. Online)*, v. 43, n. 3, p. 238–248, 2010.
- SETHI, S. *Considerações gerais sobre pneumonia*. 2019. Disponível em: <<https://www.msmanuals.com/pt/casa/dist%C3%BArbios-pulmonares-e-das-vias-respirat%C3%B3rias/pneumonia/considera%C3%A7%C3%B5es-gerais-sobre-pneumonia>>. Acesso em: 01 de maio de 2019.
- SHARMA, A. *Is there a difference between neural networks and convolutional neural networks?* 2018. Disponível em: <<https://www.quora.com/Is-there-a-difference-between-neural-networks-and-convolutional-neural-networks>>. Acesso em: 16 de outubro de 2018.
- SHIN, H.-C. et al. Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning. *IEEE transactions on medical imaging*, IEEE, v. 35, n. 5, p. 1285–1298, 2016.
- SHIRAISHI, J. et al. Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. *American Journal of Roentgenology*, Am Roentgen Ray Soc, v. 174, n. 1, p. 71–74, 2000.

- SIK-HO TSANG. *Review: ResNet — Winner of ILSVRC 2015 (Image Classification, Localization, Detection)*. 2018. Disponível em: <<https://towardsdatascience.com/review-resnet-winner-of-ilsvrc-2015-image-classification-localization-detection-e39402bfa5d8>>. Acesso em: 15 de março de 2021.
- SIM. *Societa Italiana di Radiologia Medical Interventistica - COVID-19 database*. 2020. Disponível em: <<https://www.sirm.org/category/senza-categoria/covid-19/>>. Acesso em: 30 de março de 2021.
- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- SIRAZITDINOV, I. et al. Deep neural network ensemble for pneumonia localization from a large-scale chest x-ray database. *Computers & Electrical Engineering*, Elsevier, v. 78, p. 388–399, 2019.
- SOCIEDADE BRASILEIRA DE REUMATOLOGIA. *Radiografia Simples*. 2011. Disponível em: <<https://www.reumatologia.org.br/orientacoes-ao-paciente/radiografias-simples/>>. Acesso em: 20 de maio de 2019.
- SOKASHE, M. S. V. Computer assisted method for cervical vertebrae segmentation from x-ray images. *Computer*, v. 2, n. 11, p. 4387–8, 2013.
- SOUSA, G. G. B.; FERNANDES, V. R. M.; PAIVA, A. C. de. Optimized deep learning architecture for the diagnosis of pneumonia through chest x-rays. *International Conference on Image Analysis and Recognition*, p. 353–361, 2019.
- SRIVASTAVA, N. et al. Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, JMLR. org, v. 15, n. 1, p. 1929–1958, 2014.
- STEPHEN, O. et al. An efficient deep learning approach to pneumonia classification in healthcare. *Journal of healthcare engineering*, Hindawi, v. 2019, 2019.
- SZEGEDY, C. et al. Inception-v4, inception-resnet and the impact of residual connections on learning. *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.
- SZEGEDY, C. et al. Going deeper with convolutions. *CoRR*, abs/1409.4842, 2014. Disponível em: <<http://arxiv.org/abs/1409.4842>>.
- TAN, M.; LE, Q. V. Efficientnet: Rethinking model scaling for convolutional neural networks. *arXiv preprint arXiv:1905.11946*, 2019.
- TANG, H.; KIM, D. R.; XIE, X. Automated pulmonary nodule detection using 3d deep convolutional neural networks. *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, p. 523–526, 2018.
- TANG, H.; LIU, X.; XIE, X. An end-to-end framework for integrated pulmonary nodule detection and false positive reduction. *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, p. 859–862, 2019.
- TESINI, B. *Coronavírus e síndromes respiratórias agudas (covid-19, mers e sars)*. 2020. Disponível em: <<https://www.msmanuals.com/pt-br/profissional/doen%C3%A7as-infecciosas/v%C3%ADrus-respirat%C3%B3rios/coronav%C3%ADrus-e-s%C3%ADndromes-respirat%C3%B3rias-agudas-covid-19-mers-e-sars>>. Acesso em: 29 de março de 2021.

THRUN, S.; PRATT, L. Learning to learn: Introduction and overview. *Learning to Learn*, Springer US, Boston, MA, p. 3–17, 1998. Disponível em: <https://doi.org/10.1007/978-1-4615-5529-2_1>.

TING, D. S. et al. Deep learning in estimating prevalence and systemic risk factors for diabetic retinopathy: a multi-ethnic study. *NPJ digital medicine*, Nature Publishing Group, v. 2, n. 1, p. 1–8, 2019.

TOĞAÇAR, M. et al. A deep feature learning model for pneumonia detection applying a combination of mrmr feature selection and machine learning models. *Irbm*, Elsevier, v. 41, n. 4, p. 212–222, 2020.

TOPOL, E. J. High-performance medicine: the convergence of human and artificial intelligence. *Nature medicine*, Nature Publishing Group, v. 25, n. 1, p. 44, 2019.

UCAR, F.; KORKMAZ, D. Covidiagnosis-net: Deep bayes-squeezenet based diagnosis of the coronavirus disease 2019 (covid-19) from x-ray images. *Medical Hypotheses*, Elsevier, v. 140, p. 109761, 2020.

UNICEF; ORGANIZATION, W. H. *Pneumonia: The forgotten killer of children. 2006. 2008.*

VAYÁ, M. de la I. et al. *BIMCV COVID-19+: a large annotated dataset of RX and CT images from COVID-19 patients.* 2020.

VEJA SAÚDE. *Covid-19: as principais perguntas e respostas.* 2021. Disponível em: <<https://saude.abril.com.br/bem-estar/covid-19-as-principais-perguntas-e-respostas/>>. Acesso em: 20 de março de 2021.

VILALTA, R.; GIRAUD-CARRIER, C.; BRAZDIL, P. Meta-learning-concepts and techniques. *Data mining and knowledge discovery handbook*, Springer, p. 717–731, 2009.

VISION, S.; LAB, L. *Convolutional Neural Networks for Visual Recognition.* 2018. <http://cs231n.github.io/convolutional-networks/>. Online; acessado em 20 de Outubro de 2018.

VITAL, D. A.; SAIS, B. T.; MORAES, M. C. Robust pulmonary segmentation for chest radiography, combining enhancement, adaptive morphology and innovative active contours. *Research on Biomedical Engineering*, SciELO Brasil, v. 34, n. 3, p. 234–245, 2018.

WADA, D. T.; RODRIGUES, J. A. H.; SANTOS, M. K. Aspectos técnicos e roteiro de análise da radiografia de tórax. *Medicina (Ribeirão Preto. Online)*, v. 52, n. suppl., p. 5–15, 2019.

WANG, L.; LIN, Z. Q.; WONG, A. Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. *Scientific Reports*, Nature Publishing Group, v. 10, n. 1, p. 1–12, 2020.

WANG, X. et al. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. *Proceedings of the IEEE conference on computer vision and pattern recognition*, p. 2097–2106, 2017.

- WANGENHEIM, A. *Deep Learning:Reconhecimento de Imagens*. 2018. Disponível em: <<http://www.lapix.ufsc.br/ensino/visao/visao-computacionaldeep-learning/deep-learningreconhecimento-de-imagens>>. Acesso em: 15 de fevereiro de 2020.
- WEISBERG, S. P. et al. Distinct antibody responses to sars-cov-2 in children and adults across the covid-19 clinical spectrum. *Nature immunology*, Nature Publishing Group, v. 22, n. 1, p. 25–31, 2021.
- WORLD HEALTH ORGANIZATION. *Pneumonia*. 2019. Disponível em: <https://www.who.int/topics/pneumococcal_infections/en/>. Acesso em: 01 de fevereiro de 2020.
- WORLD HEALTH ORGANIZATION. *WHO Coronavirus (COVID-19) Dashboard*. 2020. Disponível em: <<https://covid19.who.int>>. Acesso em: 15 de março de 2021.
- XIE, S. et al. Aggregated residual transformations for deep neural networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*, p. 1492–1500, 2017.
- XIE, Z. Towards single-phase single-stage detection of pulmonary nodules in chest ct imaging. *arXiv preprint arXiv:1807.05972*, 2018.
- XU, Y. et al. Deep learning predicts lung cancer treatment response from serial medical imaging. *Clinical Cancer Research*, AACR, v. 25, n. 11, p. 3266–3275, 2019.
- YAKUBOVSKIY, P. *Segmentation Models*. GitHub, 2019. Disponível em: <https://github.com/qubvel/segmentation_models>. Acesso em: 20 de jan. de 2020.
- YALA, A. et al. A deep learning mammography-based model for improved breast cancer risk prediction. *Radiology*, Radiological Society of North America, v. 292, n. 1, p. 60–66, 2019.
- YAO, L. et al. Learning to diagnose from scratch by exploiting dependencies among labels. *arXiv preprint arXiv:1710.10501*, 2017.
- YAO, L. et al. Weakly supervised medical diagnosis and localization from multiple resolutions. *arXiv preprint arXiv:1803.07703*, 2018.
- YASAKA, K. et al. Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced ct: a preliminary study. *Radiology*, Radiological Society of North America, v. 286, n. 3, p. 887–896, 2018.
- YOSINSKI, J. et al. How transferable are features in deep neural networks? In: *Advances in neural information processing systems*. Washinton, DC, USA: NIPS, 2014. p. 3320–3328.
- YU, L. et al. Automated melanoma recognition in dermoscopy images via very deep residual networks. *IEEE transactions on medical imaging*, IEEE, v. 36, n. 4, p. 994–1004, 2017.
- ZHAO, H. et al. Pyramid scene parsing network. *Proceedings of the IEEE conference on computer vision and pattern recognition*, p. 2881–2890, 2017.
- ZHOU, B. et al. Learning deep features for scene recognition using places database. In: *Advances in neural information processing systems*. Montreal, Canada: Curran Associates, Inc., 2014. p. 487–495.

ZHOU, Y. . et al. Image restoration using a neural network. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, v. 36, n. 7, p. 1141–1151, July 1988. ISSN 0096-3518.