

Universidade Católica de Moçambique
Faculdade de Gestão de Turismo e Informática

Uso de Deep Learning para Identificação de Pneumonia em Imagens de
Raio-X

Nazir Abibo Francisco Daúdo

Pemba, Abril de 2023

Universidade Católica de Moçambique

Faculdade de Gestão de Turismo e Informática

Uso de Deep Learning para Identificação de Pneumonia em Imagens de
Raio-X

Nazir Abibo Francisco Daúdo

Monografia Submetida a Faculdade de
Gestão de Turismo e Informática, da
Universidade Católica de Moçambique,
como requisito parcial para obtenção do
grau de Licenciatura em Tecnologias de
Informação.

Supervisor: Nacir Ibraimo

Pemba, Abril de 2023

ÍNDICE

DECLARACÇÃO	I
AGRADECIMENTO.....	II
DEDICATÓRIA	IV
EPÍGRAFE	V
LISTA DE FIGURAS.....	VI
LISTA DE TABELAS.....	VIII
LISTA DE ABREVIATURAS	IX
RESUMO.....	X
ABSTRACT.....	XI
CAPÍTULO I - INTRODUÇÃO.....	1
1. Contextualização	1
1.1. Problematização	2
1.2. Justificativa	2
1.3. Objectivos	3
1.3.1. Geral.....	3
1.3.2. Específicos	3
CAPÍTULO II – FUNDAMENTAÇÃO TEÓRICA	4
2.1. Anatomia do Pulmão Humano	4
2.2. Pneumonia.....	6
2.2.1. Tipos de Pneumonia.....	6
2.2.2. Causas da Pneumonia.....	7
2.2.3. Sintomas da Pneumonia	8
2.2.4. Tratamento de Pneumonia.....	8
2.3. Radiografia do Tórax	9
2.4. Python	9
2.5. Machine Learning	10
2.5.1. Divisão de Machine Learning	11
2.5.1.1. Aprendizado supervisionado	12
2.5.1.2. Aprendizado não supervisionado.....	12
2.5.1.3. Aprendizado por reforço.....	13
2.5.2. Deep Learning	14
2.5.2.1. Deep Convolution Neural Networks	15
2.6. Arquitecturas Implementadas.....	18
2.6.1. DenseNet121	18

2.6.2. Inception_V3.....	19
2.6.3. ResNet50.....	20
2.6.4. Modelo_1	21
2.7. Trabalhos relacionados.....	22
2.8. Considerações gerais	25
CAPÍTULO III – METODOLOGIA	27
3.1. Tipo de pesquisa.....	27
3.1.1. Quanto a abordagem	27
3.1.2. Quanto a natureza.....	27
3.2. Técnicas e instrumentação de recolha de dados.....	27
3.3. População	28
3.4. Amostra	28
3.5. Como os dados recolhidos serão tratados	28
4. Metodologia	28
3.1. Construção do conjunto de dados (ou dataset).....	29
4.1.1. Pré-processamento do dataset	30
4.2. Implementação	32
4.2.1.1. Materiais	32
4.2.1.2. Métodos de Validação de classificação	34
4.3. Treinamento das CNN's.....	36
4.4. Avaliação dos Classificadores.....	36
CAPÍTULO IV: ANÁLISE E DISCUSSÃO DOS RESULTADOS.....	38
4.1. Desempenho do Inception_V3	39
4.2. Desempenho do DenseNet121	43
4.3. Desempenho do Resnet50	45
4.4. Desempenho do Modelo_1	46
4.5. Discussão.....	49
CAPÍTULO V: CONCLUSÃO	51
5.1. Limitações da Pesquisa	51
5.2. Sugestões.....	52
6. Referências Bibliográficas	54

DECLARACÇÃO

Eu, **Nazir Abibo Francisco Daúdo**, filho de Nuro Daúdo António e Ana Violante Arnaldo Francisco, declaro que este trabalho é da minha autoria e nunca foi submetido para qualquer outra instituição para obtenção de grau académico.

Data ____/____/____

Assinatura_____

Nome do Supervisor:

dr. Nacir Ibraimo

Data ____/____/____

Assinatura_____

AGRADECIMENTO

Em nome de *Allah*, O Clemente e Misericordioso, é com grande humildade e gratidão que expresso meus agradecimentos por ter concluído esta monografia. Agradeço primeiramente a Allah, o Senhor dos Mundos, por Sua graça e misericórdia, e por me conceder a força e o conhecimento necessários para enfrentar os desafios e as dificuldades que encontrei durante este processo.

A monografia que concluí é fruto de um esforço conjunto e cooperação de muitas pessoas. Em primeiro lugar, gostaria de agradecer a minha família, que sempre me apoiou e encorajou em minha jornada acadêmica. Sem o amor, a paciência e o apoio de meus pais, irmãos e avós, certamente não teria chegado tão longe. Sou muito grato aos meus pais por nunca terem desistido de mim mesmo quando tinham tudo para desistir, palavras de gratidão nunca serão suficiente para expressar.

Sou grato aos meus irmãos, Hawa Aboobacar, Arcélia Ana Arnaldo pelo apoio incondicional e por me ter amparado num dos momentos mais turbulentos da minha vida, Arcélio Ana Arnaldo, Urbay Daúdo (em memória) pelos inúmeros conselhos, pela ajuda que me deu e pelas vezes que me defendeu mesmo eu não tendo razão, para depois me corrigir no privado, Nuro Francisco Daúdo, pelas vezes que me tirou de muitos sufocos, e por tudo que já passamos juntos, Aiuba Aboobacar, por ser esse homem de grande coração, pelo companheirismo e pelo suporte dado neste tempo todo.

Gostaria de agradecer também aos meus amigos, que me acompanharam ao longo deste processo e me apoiaram em momentos difíceis. Augusto António, Elísio Paúnde, Elton Gimo, Guerzeneldo Marcelino, Nelson Muhoho, Elana Raul, Abdala Falume, Nurdine Maulana, Chafim Purulia, Ednilson Sarmento, Aniceto de Carvalho, agradeço pela companhia, pelas conversas e pelas risadas, que me ajudaram a manter a sanidade e a perspectiva durante os períodos mais intensos da monografia e nos momentos da vida. Não me esquecendo, agradecer ao meu sobrinho Buquine Buquine por me estender a mão durante esta longa jornada, e minha tia Stelia Selemane.

Gostaria de expressar minha sincera gratidão à minha segunda família aqui em Pemba. Quando cheguei aqui, senti-me um pouco perdido e sozinho, mas vocês me acolheram de braços abertos e fizeram com que eu me sentisse em casa, dona Mariana (em memória), dona

Fátima António Antipa, nenhum agradecimento será suficiente para todo o apoio que vocês me deram. Sem a ajuda de vocês, não sei como teria conseguido me adaptar a esta nova vida em Pemba. Agradeço também pelo carinho e amizade que vocês me mostraram. Compartilhamos momentos maravilhosos juntos. Sinto-me abençoado por ter encontrado uma família tão amorosa e acolhedora aqui em Pemba.

Aproveito também este espaço para expressar minha mais profunda gratidão a Clytie de Fátima Simba. Você é uma mulher incrível e uma inspiração para todos que a conhecem. Sua bondade, generosidade e amor incondicional são verdadeiramente admiráveis. Eu me sinto extremamente abençoado por tê-la em minha vida. Quando passei pelo momento mais difícil de minha vida, você foi a minha rocha, meu porto seguro e minha força. Você me levantou quando eu caí e me ajudou a superar as adversidades. Agradeço por todo o apoio emocional que você me deu, por suas palavras de sabedoria e pelos conselhos valiosos que você compartilhou comigo. Sua presença em minha vida foi um verdadeiro presente de Allah, e eu nunca vou esquecer o que você fez por mim. Você é uma pessoa extraordinária, e espero que a vida lhe retribua todas as bênçãos que você tem dado aos outros. Obrigado por ser uma companheira tão maravilhosa e por tudo o que você fez por mim. Que Allah abençoe você e sua família sempre.

Agradeço aos meus professores e orientadores, em especial ao meu supervisor Nacir Ibraimo, que foram uma fonte de inspiração e orientação ao longo de minha jornada acadêmica. Sou grato por sua sabedoria, paciência e dedicação em ajudar-me a desenvolver minhas habilidades acadêmicas e a aprimorar minha capacidade de pensar criticamente e analiticamente. Seu ensinamento foi fundamental para o meu sucesso neste projeto e em toda a minha carreira acadêmica. Agradeço também aos meus colegas de classe, com quem tive a oportunidade de compartilhar experiências e aprendizados durante todo o curso. Seu apoio e encorajamento foram fundamentais para que eu pudesse superar os desafios que surgiram ao longo do caminho.

Por fim, agradeço a todos os que de alguma forma contribuíram para a realização desta monografia, incluindo funcionários da biblioteca, colaboradores e profissionais que forneceram suporte técnico e outros recursos necessários ao longo deste processo. Espero que esta monografia possa contribuir de alguma forma para o avanço do conhecimento em minha área de estudo, e para o bem da humanidade em geral. Mais uma vez, expresso minha profunda gratidão

DEDICATÓRIA

Dedico este trabalho em memória da minha irmã, Urbay da Virgínia Daúdo (Que Allah tenha misericórdia de ti, e que te dê o Jan'nat), A minha filha Nura Daúdo e aos meus pais, Nuro Daúdo António e Ana Violante Arnaldo Francisco.

EPÍGRAFE

"Em verdade, com a dificuldade vem a facilidade." (Qur'An 94:5-6)

LISTA DE FIGURAS

Figura 1. Anatomia do Pulmão Humano.....	5
Figura 2. Chest X-Ray Images (Pneumonia).....	6
Figura 3. Esquema do aprendizado supervisionado.	12
Figura 4. Fluxo do aprendizado não supervisionado.....	13
Figura 5. Fluxo do Aprendizado por reforço.....	14
Figura 6. Redes Neurais Convolucionais.	15
Figura 7. Operações de Pooling.....	16
Figura 8. Ilustração de uma camada totalmente conectada.	17
Figura 9. Estrutura do DenseNet121.	18
Figura 10. Estrutura do Inception_V3.	19
Figura 11. Estrutura do ResNet50.	20
Figura 12. Imagens de Raio-X de tórax depois de aplicadas as técnicas de pré – processamento.	31
Figura 13. Métodos propostos.	28
Figura 14. Precisão do modelo Inception_V3.	40
Figura 15. Perda do modelo Inception_V3.....	40
Figura 16. Matriz de confusão Inception_V3.....	41
Figura 17. Gráfico de precisão do Inception_V3.	42
Figura 18. Gráfico de perda do Inception_V3.....	42
Figura 19. Matriz de confusão Inception_V3.....	43
Figura 20. Matriz de confusão DenseNet121.	44

Figura 21. Gráfico de precisão DenseNet121.....	44
Figura 22. Gráfico de perda DenseNet121.	45
Figura 23. Matriz de confusão ResNet50.	46
Figura 24. Gráfico de acurácia ResNet50.....	46
Figura 25. Gráfico de acurácia Modelo_1.....	48
Figura 26. Gráfico de perda Modelo_1.	48
Figura 27. Matriz de confusão Modelo_1.	49

LISTA DE TABELAS

Tabela 1. Trabalhos relacionados.	22
Tabela 2. Descrição dos modelos implementados.....	32
Tabela 3. Matriz de confusão.	34
Tabela 4.Resultados classificação geral.	38
Tabela 5.Resultado do teste do modelo.....	41
Tabela 6. Resultados do teste do modelo Inception_V3.	43
Tabela 7. Resultados do teste do modelo Inception_V3.	44
Tabela 8. Resultados teste do modelo ResNet50.....	45
Tabela 9. Resultados teste do modelo Modelo_1.....	47

LISTA DE ABREVIATURAS

ADAM – Adaptive Moment Estimation

CNN – Convolution Neural Network

CPU – Central Processing Unit

DCNN - Deep Convolution Neural Network

DL – Deep Learning

GB – Giga Bytes

GPU – Graphic Processing Unit

IA – Inteligência Artificial

IoT - Internet of Things

ML – Machine Learning

OMS – Organização Mundial da Saúde

RAM - Random Access Memory

TPU – Tensor Processing Unit

WHO – World Health Organization

RESUMO

Em Moçambique a pneumonia é responsável por 16% de morte de crianças menores de 5 anos de idade, segundo a OMS, sendo responsável pela morte de cerca de 10 mil crianças por ano. Este problema de saúde afecta a qualidade de vida dos pacientes, entretanto o diagnóstico precoce evita a morte e permite que o paciente tenha um tratamento adequado. Entretanto, em Moçambique são utilizados os métodos convencionais para o diagnóstico de Pneumonia, há escassez de profissionais, e também falta de equipamentos nas unidades sanitárias, o que pode levar a uma alta taxa de diagnósticos errados desta doença. Portanto, o diagnóstico precoce torna-se um grande desafio para as autoridades de saúde. O presente trabalho, propõe o uso de arquitecturas modernas de Deep Convolutional Neural Networks (CNN's), para a classificação de Pneumonia baseado em imagens de Raio-X do tórax.

Palavras-Chave: Pneumonia, Redes Neurais Convolucionais, Imagem de Raio-X

ABSTRACT

In Mozambique, pneumonia is responsible for 16% of deaths of children under 5 years of age, according to the WHO, being responsible for the death of about 10 thousand children per year. This health problem affects the quality of life of patients, however early diagnosis prevents death and allows the patient to have an adequate treatment. However, in Mozambique conventional methods of this disease are used for professional diagnosis, as well as the lack of equipment in health facilities, which can lead to an erroneous diagnosis rate. Therefore, early diagnosis becomes a major challenge for health authorities. The present work, planning modern architectures of Deep Convolutional Neural Networks (CNN's), for a classification of Pneumonia based on chest X-ray images.

Keywords: Pneumonia, Convolutional Neural Networks, X-Ray Imaging

CAPÍTULO I - INTRODUÇÃO

1. Contextualização

A pneumonia é uma das principais causas de mortalidade infantil em países em desenvolvimento, incluindo Moçambique. O diagnóstico precoce e preciso da pneumonia é fundamental para o tratamento eficaz da doença e, consequentemente, para a redução da mortalidade.

Recentemente, tem havido um interesse crescente em utilizar técnicas de aprendizado profundo (Deep Learning) para ajudar no diagnóstico da pneumonia a partir de imagens de raio-x do tórax. Essas técnicas são capazes de analisar grandes quantidades de dados de forma rápida e precisa, permitindo a detecção de anomalias e padrões sutis que podem ser difíceis para os médicos identificarem.

Em Moçambique, onde muitas comunidades não têm acesso fácil a serviços de saúde e recursos limitados, o uso de tecnologias de diagnóstico automatizado pode ser particularmente benéfico. Os sistemas de aprendizado profundo poderiam ser usados para triar pacientes, permitindo que aqueles com maior probabilidade de ter pneumonia sejam identificados rapidamente e recebam tratamento prioritário.

No entanto, é importante lembrar que o uso de tecnologias de aprendizado profundo requer uma infraestrutura adequada e um treinamento extensivo para garantir que os resultados sejam precisos e confiáveis. Além disso, é importante que a implementação dessas tecnologias ocorra em colaboração com médicos e profissionais de saúde locais para garantir que o sistema seja adaptado às necessidades específicas da população de Moçambique.

O presente trabalho, propõe uso de arquiteturas modernas de Deep Learning (Convolutional Neural Networks (CNN's)) para diagnóstico automático da pneumonia. As CNN's apresentam desempenhos exemplares na classificação de imagens em relação aos outros algoritmos de Machine Learning. As arquiteturas modernas apresentam maiores profundidades, tornando-se um factor importante no aumento da eficiência da rede, bem como na obtenção de desempenho ideal. O modelo a ser desenvolvido usará imagens de Raio-x para fazer a classificação automática de Pneumonia cujo propósito é atingir uma proporção percentual média de precisão de classificação.

1.1. Problematização

Pneumonia é uma infecção aguda dos pulmões, que pode produzir sinais e sintomas respiratórios, como tosse, respiração curta e rápida, produção de secreção e dores no peito, além de sintomas sistêmicos não-específicos, incluindo febre, fadiga, dores musculares e falta de apetite (Raghavendran et al., 2007). Os médicos costumam realizar o diagnóstico da pneumonia não só através dos sintomas que caracterizam esta patologia, mas também com a realização de um exame ao tórax do doente. A radiografia ou RX de tórax é realizada para confirmar o diagnóstico. (Costa, 2022)

Entretanto, em Moçambique são utilizados os métodos convencionais para o diagnóstico de Pneumonia, há escassez de profissionais, e também falta de equipamentos nas unidades sanitárias, o que pode levar a uma alta taxa de diagnósticos errados desta doença. O presente trabalho, propõe o uso de arquiteturas modernas de Deep Convolutional Neural Networks (CNN's), para a classificação de Pneumonia baseado em imagens de Raio-X do tórax. Deste modo, surge a seguinte questão de partida: **“Até que proporção percentual média de previsão as DCNN's podem alcançar na identificação de pneumonia baseado em imagem de Raio-X?”**.

1.2. Justificativa

O diagnóstico de pneumonia é um processo que só pode ser confirmado por médicos e que pode resultar em diagnósticos diferentes ou errados dependendo do especialista. Realizar um Raio- X do tórax é uma tarefa comum em hospitais, sendo que por dia podem ser geradas imagens que precisam ser analisadas com muito cuidado. Pelo tempo que a análise e o diagnóstico podem demorar, a automatização e precisão no processo de diagnóstico é algo desejável, então o uso de métodos avançados deve ser considerado. O uso desses métodos podem diminuir a carga dos especialistas e ainda oferecer uma segunda opinião mais rápida e precisa.

Deste modo, este processo pode ser automatizado com auxílio das Deep Convolution Neural Networks (DCNN's), sendo que as CNN's provaram ter melhor desempenho em tarefas de classificação de imagens em relação aos outros algoritmos de Machine Learning, e este método irá melhorar a eficácia e eficiência no diagnóstico da pneumonia.

1.3. Objectivos

1.3.1. Geral

- Desenvolver um modelo preditivo automático baseado em Deep Convolution Neural Networks (DCNN's) para a identificação de pneumonia usando imagem de Raio-x de tórax.

1.3.2. Específicos

- Contruir um dataset para o treinamento dos algoritmos;
- Treinar os algoritmos das Deep Convolutional Neural Networks (DCNN's);
- Comparar os resultados dos treinamentos dos algoritmos;
- Escolher o algoritmo com melhor desempenho;

CAPÍTULO II – FUNDAMENTAÇÃO TEÓRICA

2.1. Anatomia do Pulmão Humano

O pulmão humano é um órgão vital responsável pelas trocas gasosas entre o ar inspirado e o sangue. A anatomia do pulmão é complexa e composta por diversas estruturas, incluindo brônquios, bronquíolos, alvéolos e vasos sanguíneos. Segundo Tortora e Derrickson (2017), os pulmões humanos são dois órgãos em formato de cone, localizados no tórax, protegidos pelas costelas e envolvidos por duas membranas chamadas de pleuras. Cada pulmão é subdividido em lobos, sendo o pulmão direito dividido em três lobos e o esquerdo em dois lobos.

Os brônquios são as primeiras estruturas que transportam o ar para dentro dos pulmões, ramificando-se em bronquíolos e, posteriormente, em alvéolos, que são os responsáveis pelas trocas gasosas. Segundo Moore, Dalley e Agur (2018), cada alvéolo é envolvido por uma rede de capilares sanguíneos, que transportam o sangue rico em dióxido de carbono para ser expelido do corpo e o sangue rico em oxigênio para ser distribuído para o restante do organismo. Essas trocas gasosas são essenciais para o bom funcionamento do corpo humano.

Além disso, o pulmão humano é um órgão que apresenta diversas adaptações que permitem a respiração em diferentes condições ambientais e fisiológicas. De acordo com Marieb e Hoehn (2018), os pulmões têm a capacidade de se expandir e contrair de acordo com a demanda respiratória do corpo, além de apresentarem cílios e muco, que ajudam a filtrar e remover partículas e micro-organismos presentes no ar inspirado. Essas adaptações garantem que o pulmão humano possa realizar suas funções de forma eficiente e sem comprometer a saúde do indivíduo.

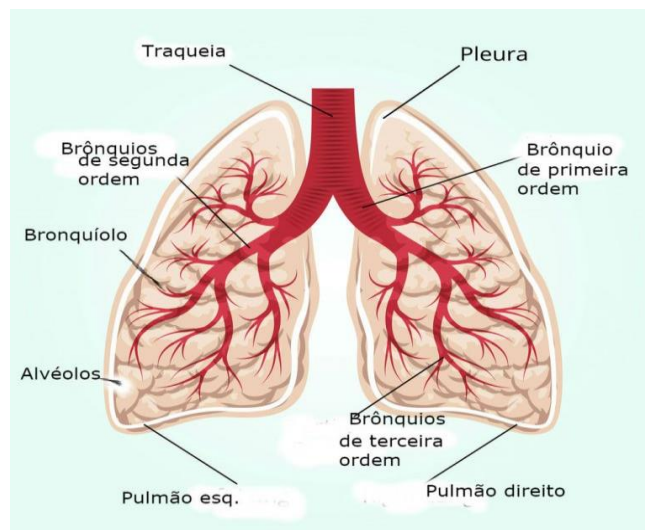


Figura 1 Anatomia do Pulmão Humano.

Fonte: Mariana Varella. (2017).

Os brônquios, bronquíolos, alvéolos e vasos sanguíneos são estruturas importantes do sistema respiratório humano.

Os **brônquios** são tubos que transportam o ar para dentro dos pulmões e se ramificam em bronquíolos, que por sua vez se dividem em bronquíolos terminais e respiratórios. Os bronquíolos respiratórios, por sua vez, são os responsáveis por conduzir o ar para os alvéolos, onde ocorrem as trocas gasosas.

Os **alvéolos** são pequenos sacos de ar localizados nas extremidades dos bronquíolos respiratórios. São revestidos por uma fina camada de células, chamada de epitélio alveolar, e envolvidos por uma rede de capilares sanguíneos. É nesses capilares que ocorrem as trocas gasosas entre o ar inspirado e o sangue, permitindo a oxigenação do sangue e a eliminação do dióxido de carbono.

Os **vasos sanguíneos**, por sua vez, são estruturas que transportam o sangue para os pulmões e de volta para o coração. Os capilares sanguíneos que envolvem os alvéolos são responsáveis pelas trocas gasosas, enquanto as artérias e veias pulmonares transportam o sangue para e do pulmão, respectivamente.

Em conjunto, essas estruturas formam um sistema complexo e altamente especializado para permitir a respiração adequada e a manutenção da homeostase do corpo humano.

2.2.Pneumonia

A pneumonia é uma inflamação que acomete os pulmões. Para ser mais exato, ela atinge principalmente os bronquíolos (pequenos tubos que transportam o ar dos brônquios para os alvéolos, onde ocorre a troca gasosa) e o interstício (tecido mais interno do órgão). Geralmente, a pneumonia é causada por agentes infecciosos — inclusive o coronavírus —, mas também pode decorrer da ação de substâncias químicas. Seus sintomas incluem falta de ar, dor no peito e tosse. E o tratamento, claro, depende muito da causa. (Santos M. T., pneumonia-o-que-e, 2021)

A pneumonia é um tipo de infecção pulmonar (no pulmão) devida à proliferação de micro-organismos infecciosos ao nível dos alvéolos (pequenos sacos de ar). A pneumonia é uma doença grave, sendo mais frequente nos adultos com mais de 60 anos de idade, afetando, no entanto, de igual modo ambos os sexos. Na criança ou bebê (pneumonia infantil) as causas são ligeiramente diferentes das do adulto, bem como o quadro clínico. (Costa, pneumonia, 2022)

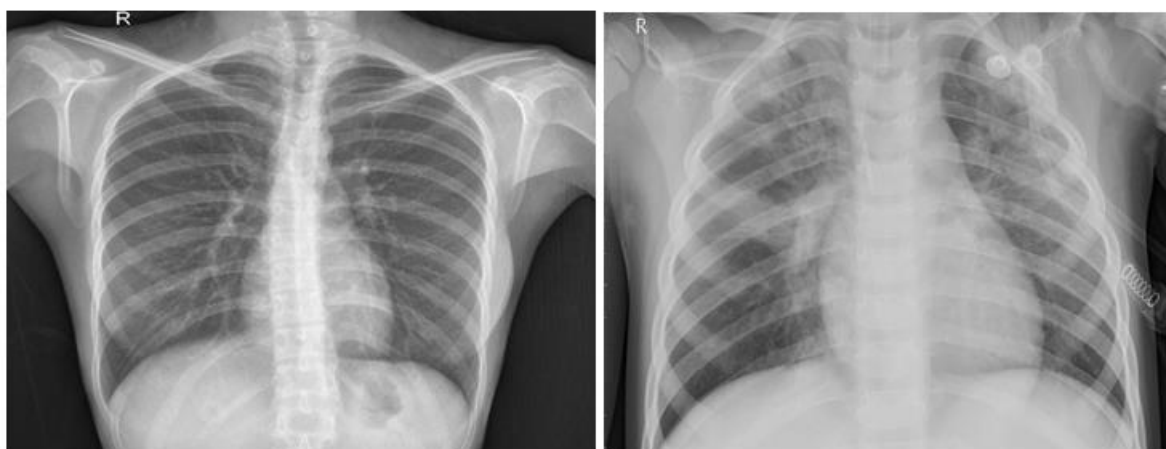


Figura 2. Chest X-Ray Images (Pneumonia).

Fonte: Kaggle. (2018)

a) Pulmão normal

b) Pulmão com Pneumonia

2.2.1. Tipos de Pneumonia

Existem vários tipos de pneumonia, e a classificação pode ser feita de diferentes maneiras, levando em consideração fatores como a causa da infecção, a gravidade da doença, a idade do

paciente e outros fatores.

De acordo com a American Thoracic Society e a Infectious Diseases Society of America, a pneumonia pode ser classificada em quatro categorias principais: pneumonia adquirida na comunidade (PAC), pneumonia hospitalar (PH), pneumonia associada à ventilação mecânica (PAVM) e pneumonia em pacientes imunocomprometidos (PIC) (Mandell et al., 2019).

Outra classificação comum é baseada no agente infeccioso que causa a pneumonia. Por exemplo, a pneumonia bacteriana pode ser causada por *Streptococcus pneumoniae*, *Haemophilus influenzae*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa* e outros agentes bacterianos. A pneumonia viral pode ser causada por vírus respiratórios como o vírus da gripe, o vírus sincicial respiratório (VSR) e outros. A pneumonia fúngica pode ser causada por fungos como *Pneumocystis jirovecii* e *Candida albicans* (Long et al., 2020).

2.2.2. Causas da Pneumonia

A pneumonia é uma doença infecciosa do trato respiratório inferior que pode ser causada por vários agentes infecciosos, como bactérias, vírus, fungos e parasitas. As causas específicas da pneumonia podem variar dependendo do tipo de agente infeccioso e da condição do paciente.

De acordo com a American Thoracic Society e a Infectious Diseases Society of America (2019), as causas mais comuns de pneumonia adquirida na comunidade são:

- *Streptococcus pneumoniae*
- *Haemophilus influenzae*
- *Mycoplasma pneumoniae*
- *Chlamydia pneumoniae*
- *Legionella pneumophila*
- Respiratory syncytial virus (RSV)
- Influenza A e B vírus

Além disso, outras causas menos comuns incluem bactérias como *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa* e outros vírus, como adenovírus e vírus parainfluenza.

A pneumonia pode ser causada por diferentes agentes infecciosos, sendo as bactérias e os vírus os mais comuns. A pneumonia bacteriana pode ser causada por *Streptococcus pneumoniae*, *Mycoplasma pneumoniae*, *Legionella pneumophila* e outras bactérias. Já a pneumonia viral é causada por vírus como o vírus sincicial respiratório (VSR), o vírus da gripe (influenza) e outros. Além disso, a pneumonia pode ser causada por fungos, como o *Pneumocystis jirovecii*, que é comum em pacientes com AIDS, e por parasitas, como o *Strongyloides stercoralis*, que é comum em áreas rurais.

2.2.3. Sintomas da Pneumonia

Os sintomas da pneumonia podem variar de acordo com a causa da infecção, mas geralmente incluem tosse com muco ou pus, febre, calafrios, falta de ar, dor no peito ao respirar ou tossir, fadiga e fraqueza. Além disso, em casos mais graves, a pneumonia pode causar confusão mental, taquicardia, sudorese excessiva e lábios ou unhas azuladas devido à falta de oxigênio no sangue. (Mayo Clinic, 2021)

Para a American Thoracic Society (ATS), os sintomas da pneumonia incluem febre, tosse com produção de muco ou catarro, falta de ar, dor no peito ao respirar profundamente ou tossir, fadiga e fraqueza.

2.2.4. Tratamento de Pneumonia

O tratamento varia de acordo com a causa da doença e pode incluir o uso de antibióticos, antivirais, antifúngicos, entre outras medidas terapêuticas.

Segundo Ferreira et al. (2020), o tratamento da pneumonia causada por bactérias geralmente inclui o uso de antibióticos que apresentem eficácia contra os agentes identificados. Já no caso da pneumonia viral, o tratamento é principalmente de suporte, com o uso de antivirais em casos selecionados. Além disso, medidas como hidratação adequada, oxigenoterapia, fisioterapia respiratória e repouso são importantes para a recuperação do paciente.

Outro autor que aborda o tema é Torres et al. (2021), que destacam a importância da identificação precoce da causa da pneumonia para a escolha adequada do tratamento. Além

disso, ressaltam a necessidade de monitoramento clínico e radiológico dos pacientes para avaliar a resposta ao tratamento e ajustar a terapia quando necessário.

2.3.Radiografia do Tórax

Uma radiografia do tórax ou Raio X (RX) do tórax é um exame que serve para auxiliar o médico no diagnóstico ou avaliação da resposta aos tratamentos em várias patologias (doenças), como por exemplo a pneumonia, tuberculose, doença pulmonar obstrutiva crônica (DPOC), fraturas, entre outros. O tórax é a região anatômica situada entre a região do pescoço (cervical) e o abdômen, sendo separada desta última pelo diafragma (músculo interno) e vulgarmente conhecida como a região do “peito”. (Oliveira, 2020)

Um raio X de tórax, ou radiografia de tórax, é um exame utilizado para produzir imagens do tórax através de pequenas doses de radiação ionizante. O exame é recomendado para avaliar a caixa torácica, pulmões e coração. A radiografia de tórax pode diagnosticar condições que vão desde uma simples falta de ar ou tosse persistente até uma pneumonia ou câncer. É um exame rápido, muito útil em tratamentos de emergência e um dos mais solicitados na Medicina. (Posses, s.d.)

2.4.Python

Python é uma linguagem de programação de alto nível e de código aberto amplamente utilizada em áreas como ciência de dados, inteligência artificial, desenvolvimento web e automação de tarefas. De acordo com van Rossum e Drake (2009), Python foi criado em 1991 com o objetivo de ser uma linguagem de programação fácil de aprender, legível e de sintaxe simples. Desde então, Python tem sido uma das linguagens de programação mais populares no mundo.

Python é uma das linguagens de programação mais populares no campo da Inteligência Artificial (IA) devido à sua facilidade de uso, grande variedade de bibliotecas e comunidade ativa. Uma das bibliotecas mais conhecidas é o TensorFlow, desenvolvido pela Google, que é uma plataforma para criação e treinamento de modelos de IA, incluindo redes neurais profundas. Segundo Géron (2019), o TensorFlow é amplamente utilizado em aplicações de processamento de imagens e linguagem natural.

Outra biblioteca popular em Python é o PyTorch, desenvolvido pelo Facebook, que é uma biblioteca de aprendizado profundo com suporte para operações de tensor em CPU e GPU. De

acordo com Paszke et al. (2019), o PyTorch é fácil de usar, escalável e permite a criação de modelos complexos de forma eficiente.

O Keras é outra biblioteca popular de aprendizado de máquina em Python, que permite a criação e treinamento de modelos de forma rápida e intuitiva. Segundo Chollet (2018), o Keras é uma biblioteca de alto nível que facilita a criação de modelos de IA, permitindo a implementação de redes neurais profundas em apenas algumas linhas de código.

Finalmente, a biblioteca scikit-learn é uma das bibliotecas mais conhecidas de aprendizado de máquina em Python, que oferece uma ampla variedade de algoritmos de aprendizado supervisionado e não supervisionado, além de ferramentas para avaliação de modelos e pré-processamento de dados. De acordo com Pedregosa et al. (2011), a scikit-learn é fácil de usar, bem documentada e pode ser utilizada tanto em aplicações acadêmicas quanto comerciais.

2.5. Machine Learning

Machine Learning ou aprendizado de máquina é uma área da inteligência artificial voltada para descoberta de padrões usando um conjunto de dados para posteriormente com base no conjunto de dados usados para detectar padrões prever dados de classes futuras ou tomar decisões sobre dados novos ou incertezas (Murphy, 2012).

O Machine Learning tem se tornado cada vez mais relevante na era da informação e vem transformando diversos setores da sociedade, desde a saúde e a segurança pública até a indústria e os negócios. Segundo Alpaydin (2010), Machine Learning é uma área de Inteligência Artificial que se preocupa em desenvolver algoritmos e modelos capazes de aprender padrões e realizar previsões a partir de dados, sem serem explicitamente programados para isso. O autor destaca que o aprendizado automático pode ser supervisionado, não supervisionado ou por reforço, e que tem sido aplicado em diversas áreas, como reconhecimento de fala, visão computacional, processamento de linguagem natural, entre outras.

Outro autor importante que contribuiu para o desenvolvimento do Machine Learning é Hastie (2009), que define a área como "a ciência de fazer computadores aprenderem a partir de dados" (p. 1). Ele destaca que o aprendizado a partir de dados pode ser visto como uma forma de generalização, ou seja, a capacidade de um modelo de fazer previsões precisas em novas situações com base no que foi aprendido a partir de dados anteriores. O

autor também ressalta que o Machine Learning é uma área multidisciplinar, que envolve conceitos de estatística, computação, matemática e outras áreas relacionadas.

Por fim, Jordan e Mitchell (2015) destacam a importância do Machine Learning para o desenvolvimento da Inteligência Artificial como um todo, e afirmam que "o aprendizado a partir de dados é o coração da IA moderna" (p. 1). Os autores destacam que a área tem evoluído rapidamente, com novos algoritmos e modelos sendo propostos constantemente, e que tem sido aplicada em diversas áreas, desde a previsão do tempo até a detecção de fraudes financeiras. Eles enfatizam, no entanto, que a área ainda tem desafios a serem superados, como a interpretabilidade dos modelos e a garantia de que os resultados obtidos são éticos e justos.

2.5.1. Divisão de Machine Learning

O Machine Learning pode ser dividido em três categorias principais: o aprendizado supervisionado, o aprendizado não supervisionado e o aprendizado por reforço. Segundo Alpaydin (2010), o aprendizado supervisionado consiste em aprender a partir de exemplos rotulados, ou seja, dados que já foram classificados previamente. Nesse tipo de aprendizado, o objetivo é encontrar uma função que mapeie os dados de entrada para uma saída correta. Já o aprendizado não supervisionado, segundo o mesmo autor, consiste em aprender a partir de dados não rotulados, buscando identificar padrões e estruturas nos dados. Por fim, o aprendizado por reforço é uma forma de aprendizado em que o agente aprende a partir da interação com um ambiente, recebendo recompensas ou punições por suas ações.

Hastie, Tibshirani e Friedman (2009) também dividem o Machine Learning em três categorias, mas apresentam uma abordagem um pouco diferente. Para esses autores, o aprendizado supervisionado e o aprendizado não supervisionado são categorias distintas, mas o aprendizado por reforço é incluído em uma categoria chamada de "aprendizado de máquina por feedback". Essa categoria inclui tanto o aprendizado por reforço quanto outras formas de aprendizado que envolvem a interação do agente com o ambiente.

Outros autores, como Bishop (2006), também apresentam uma divisão do Machine Learning em diferentes categorias, mas com uma abordagem um pouco mais ampla. Para Bishop, o Machine Learning pode ser dividido em quatro categorias principais: aprendizado supervisionado, aprendizado não supervisionado, aprendizado semi-supervisionado e aprendizado por reforço. O aprendizado semi-supervisionado é uma forma de aprendizado

que combina exemplos rotulados e não rotulados para melhorar o desempenho do modelo.

Para este trabalho foram abordadas as 3 categorias principais: o aprendizado supervisionado, o aprendizado não supervisionado e o aprendizado por reforço.

2.5.1.1. Aprendizado supervisionado

O aprendizado supervisionado é uma das principais categorias de Machine Learning e consiste em aprender a partir de exemplos rotulados, ou seja, dados que já foram classificados previamente. Segundo Alpaydin (2010), o objetivo do aprendizado supervisionado é encontrar uma função que mapeie os dados de entrada para uma saída correta, ou seja, aprender a partir de exemplos para poder generalizar para novos exemplos.

Existem diversas técnicas de aprendizado supervisionado, como as árvores de decisão, as redes neurais artificiais e os algoritmos de regressão e classificação. Cada uma dessas técnicas possui vantagens e desvantagens e pode ser mais adequada para diferentes tipos de problemas.

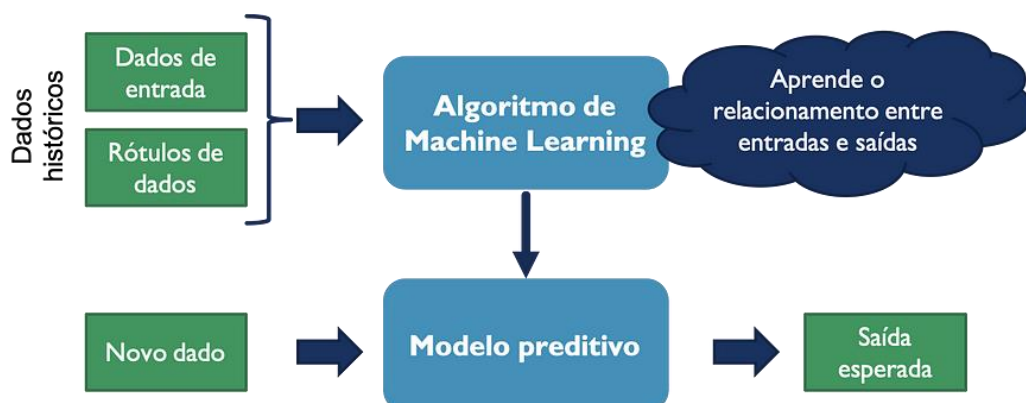


Figura 3. Esquema do aprendizado supervisionado.

Fonte: Escovedo. (2020).

De acordo com Hastie, Tibshirani e Friedman (2009), o aprendizado supervisionado é uma das áreas mais estudadas e desenvolvidas em Machine Learning, com aplicações em diversas áreas, como reconhecimento de voz, reconhecimento de imagem e processamento de linguagem natural.

2.5.1.2. Aprendizado não supervisionado

O aprendizado não supervisionado é outra categoria importante de Machine Learning, que

consiste em aprender a partir de dados não rotulados, buscando identificar padrões e estruturas nos dados. Segundo Alpaydin (2010), nesse tipo de aprendizado, o objetivo é encontrar regularidades nos dados sem a necessidade de um conjunto de dados rotulado.

Existem diversas técnicas de aprendizado não supervisionado, como o clustering, que é uma técnica de agrupamento de dados em grupos com base em suas similaridades, e a análise de componentes principais (PCA), que é uma técnica de redução de dimensionalidade que busca identificar os principais componentes dos dados.

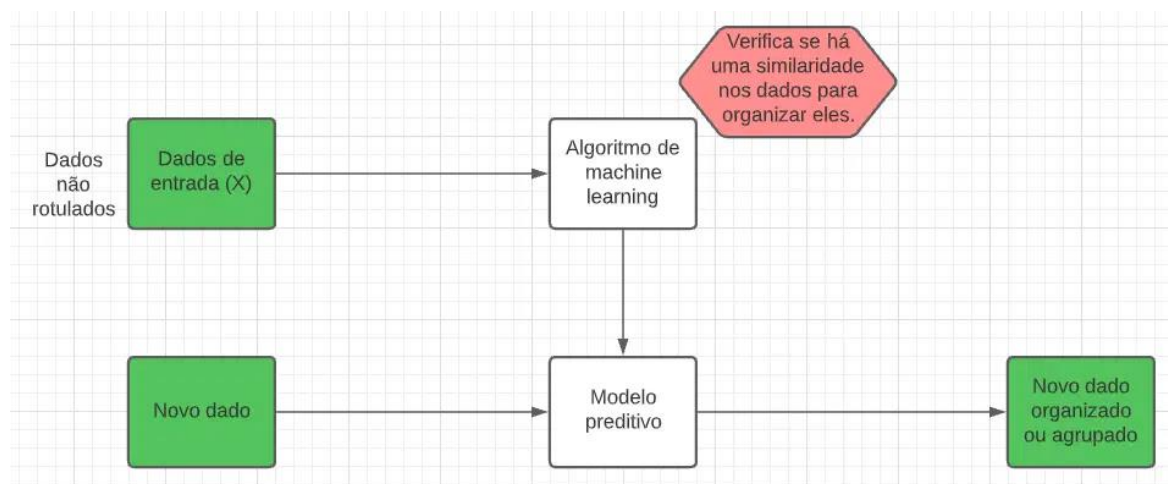


Figura 4. Fluxo do aprendizado não supervisionado.

Fonte: Marchiori. (2022).

De acordo com Hastie, Tibshirani e Friedman (2009), o aprendizado não supervisionado é útil em diversas áreas, como análise de dados, reconhecimento de padrões e mineração de dados. Além disso, pode ser usado em conjunto com o aprendizado supervisionado, por exemplo, para pré-processamento de dados e seleção de características.

2.5.1.3. Aprendizado por reforço

O aprendizado por reforço é uma categoria de Machine Learning que envolve um agente que interage com um ambiente através de tentativa e erro, e busca aprender a tomar ações que maximizem uma recompensa. Segundo Sutton e Barto (2018), no aprendizado por reforço, o agente aprende a partir da experiência, através do recebimento de feedback em forma de recompensas ou punições.

Em geral, o aprendizado por reforço envolve três componentes principais: o ambiente, o

agente e a recompensa. O ambiente é o sistema no qual o agente está interagindo, e é definido por um conjunto de estados e ações possíveis. O agente é responsável por escolher uma ação a partir do estado atual do ambiente, e seu objetivo é maximizar uma recompensa acumulada ao longo do tempo. A recompensa é um sinal que indica se a ação tomada pelo agente foi boa ou ruim.

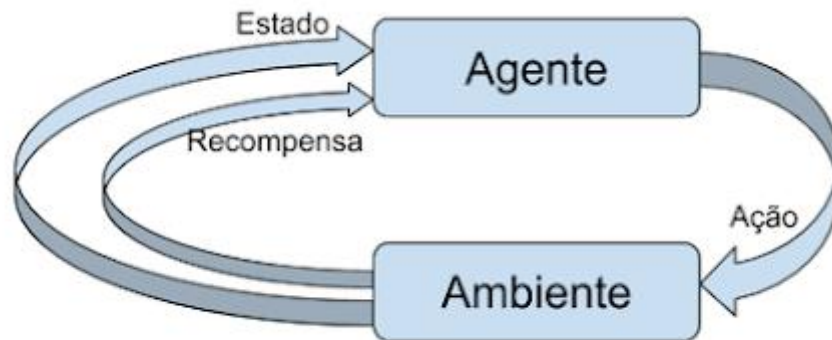


Figura 5. Fluxo do Aprendizado por reforço.

Fonte: Infodados. (2019).

De acordo com Kaelbling, Littman e Moore (1996), o aprendizado por reforço tem aplicações em diversas áreas, como robótica, jogos, controle de processos e sistemas de recomendação. Além disso, nos últimos anos, tem havido um grande avanço na aplicação do aprendizado por reforço em áreas como inteligência artificial e veículos autônomos.

2.5.2. Deep Learning

Deep Learning é uma subárea de Machine Learning que tem como objetivo treinar redes neurais profundas para aprender a partir de dados complexos, tais como imagens, sons e textos. Segundo Goodfellow, Bengio e Courville (2016), as redes neurais profundas são compostas por várias camadas de neurônios artificiais, que permitem a extração de características cada vez mais abstratas e complexas a partir dos dados.

As redes neurais profundas são treinadas utilizando algoritmos de otimização que ajustam os pesos e os vieses dos neurônios a partir de um conjunto de dados de treinamento. Esse processo é conhecido como backpropagation, e permite que a rede neural aprenda a partir dos exemplos de entrada.

De acordo com LeCun, Bengio e Hinton (2015), o Deep Learning tem tido um grande

impacto em diversas áreas, como reconhecimento de imagens, processamento de linguagem natural, robótica, entre outras. Além disso, o surgimento de arquiteturas de redes neurais profundas, como as Redes Neurais Convolucionais (CNNs) e as Redes Neurais Recorrentes (RNNs), tem permitido a criação de sistemas capazes de superar o desempenho humano em várias tarefas.

2.5.2.1. Deep Convolution Neural Networks

As Deep Convolutional Neural Networks (DCNNs) são um tipo de arquitetura de rede neural profunda que utiliza camadas convolucionais para extrair características de imagens de forma automatizada. Segundo LeCun, Bengio e Hinton (2015), as camadas convolucionais permitem que a rede aprenda a reconhecer padrões locais, como bordas, texturas e formas, que são importantes para a classificação de imagens.

As DCNNs têm sido amplamente utilizadas em tarefas de visão computacional, como reconhecimento de objetos, detecção de faces e segmentação de imagens. Uma das arquiteturas mais conhecidas de DCNNs é a Rede Neural Convolucional (CNN), que consiste em várias camadas convolucionais seguidas de camadas totalmente conectadas que realizam a classificação final.

De acordo com Krizhevsky, Sutskever e Hinton (2012), a utilização de DCNNs permitiu um grande avanço no desempenho em tarefas de reconhecimento de imagens, culminando com a vitória da rede neural profunda "AlexNet" no desafio ImageNet em 2012.

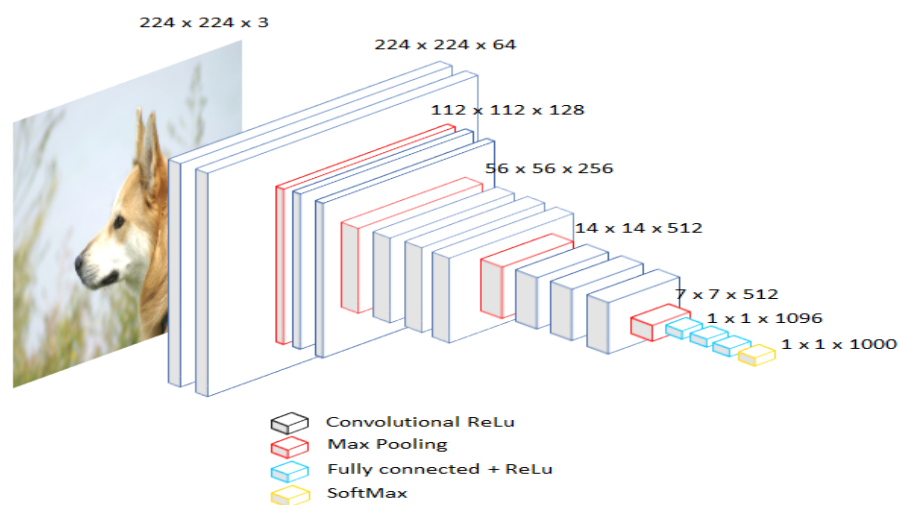


Figura 6. Redes Neurais Convolucionais.

Fonte: (Guimarães, 2021).

Segundo LeCun, Bengio e Hinton (2015), as CNNs são compostas por várias camadas, cada uma com uma função específica na extração e transformação das características das imagens de entrada. As principais camadas que compõem as CNNs são:

Convolutional Layer (Camada de Convolução): Esta camada é responsável pela aplicação de operações de convolução nas imagens de entrada. A convolução é uma operação matemática que aplica filtros ou kernels na imagem para extrair características importantes, como bordas, texturas e padrões. Cada filtro é aplicado em janelas (regiões) da imagem, e o resultado é obtido através de uma combinação linear dos valores dos pixels da janela, seguido da aplicação de uma função de ativação. A camada de convolução é essencial para a capacidade das CNNs de aprender automaticamente características relevantes das imagens.

Pooling Layer (Camada de Pooling): Também conhecida como camada de subamostragem, essa camada é responsável por reduzir a dimensão espacial das características extraídas pela camada de convolução. Ela faz isso agrupando localmente os valores de características em uma região, geralmente usando operações como média ou máximo. O pooling ajuda a reduzir a quantidade de dados processados pela rede, tornando-a mais eficiente computacionalmente e também auxiliando na invariância à translação das características aprendidas.

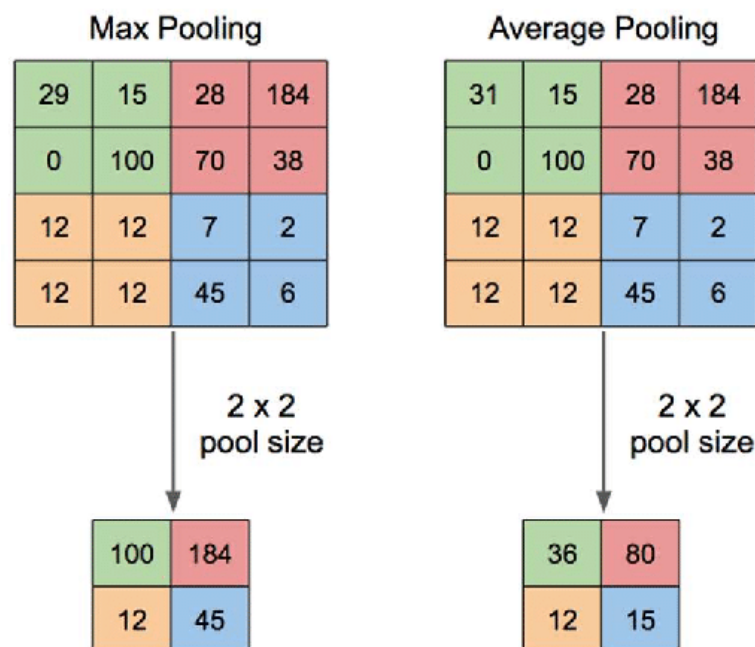


Figura 7. Operações de Pooling.

Fonte: (Rajput, 2022).

Activation Layer (Camada de Ativação): Essa camada é responsável por aplicar uma função de ativação aos resultados obtidos nas camadas anteriores. A função de ativação é usada para introduzir não-linearidades na rede, permitindo que a CNN aprenda relações complexas e não-lineares entre as características das imagens.

Fully Connected Layer (Camada Totalmente Conectada): Essa camada é semelhante à camada de ativação, mas é aplicada na saída das camadas anteriores. É uma camada densa em que cada neurônio está conectado a todos os neurônios da camada anterior. Ela é usada para combinar as características extraídas pelas camadas anteriores em uma representação final, e é frequentemente usada para a classificação final ou tarefas de regressão.

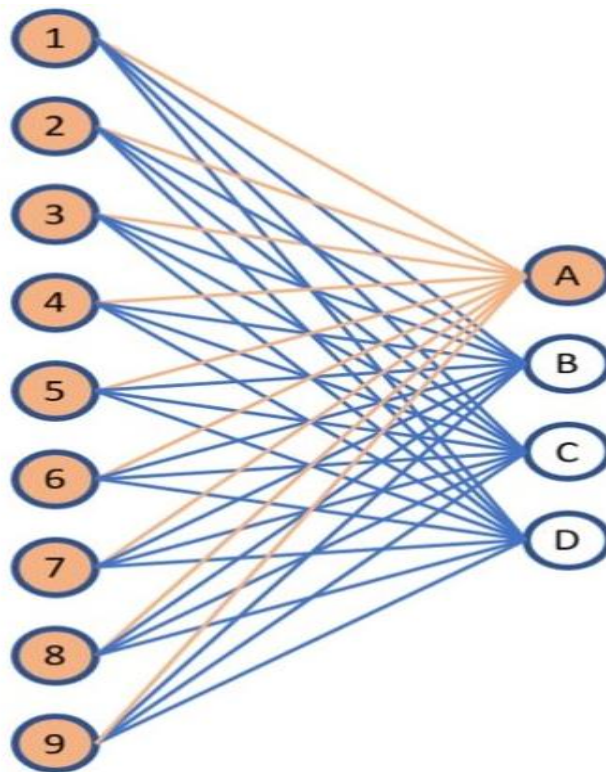


Figura 8. Ilustração de uma camada totalmente conectada.

Fonte: (Unzueta, 2022).

Além das camadas principais, como camadas de convolução, camadas de pooling e camadas totalmente conectadas, as CNNs também podem incluir camadas adicionais, como camadas de normalização, camadas de dropout para regularização, entre outras, dependendo da arquitetura específica da rede. A combinação e a ordem dessas camadas podem variar em diferentes arquiteturas de CNNs, como AlexNet, VGG, ResNet, Inception, entre outras, que têm suas próprias características e desempenho em tarefas específicas de visão computacional

(Goodfellow et al., 2016).

2.6.Arquitecturas Implementadas

Nesse subcapítulo serão apresentadas de forma resumidas as características dessas arquiteturas, DenseNet121, Inception_V3, ResNet50 e Um Modelo genérico que são arquiteturas que são muito utilizadas na classificação de imagens.

2.6.1. DenseNet121

DenseNet121 é uma arquitetura de rede neural convolucional (CNN) proposta por Gao Huang, Zhuang Liu, Laurens van der Maaten e Kilian Q. Weinberger em seu artigo intitulado "Densely Connected Convolutional Networks" publicado na Conferência de Visão Computacional e Reconhecimento de Padrões (CVPR) de 2017.

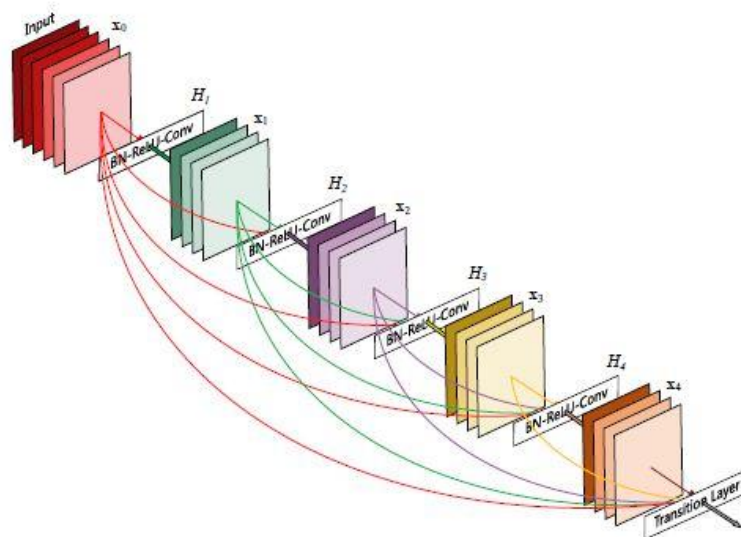


Figura 9. Estrutura do DenseNet121.

Fonte: (Ahmed, s.d.).

DenseNet121 é parte da família de arquiteturas DenseNet, que se destaca por sua abordagem de conexões densas, onde cada camada recebe as saídas de todas as camadas anteriores, permitindo um fluxo de informações mais eficiente e uma melhor utilização dos recursos de computação. DenseNet121 possui 121 camadas no total, incluindo camadas de convolução, normalização, ativação (geralmente ReLU) e pooling, seguidas por camadas de classificação para a saída final.

Uma das principais vantagens do DenseNet121 é sua capacidade de melhorar o fluxo de informações e mitigar o problema de degradação de desempenho à medida que a rede se aprofunda, tornando-o mais eficiente e preciso em comparação com outras arquiteturas de CNNs. DenseNet121 foi pré-treinado em grandes conjuntos de dados, como ImageNet, e tem sido amplamente utilizado para uma variedade de tarefas de visão computacional, como classificação de imagens, detecção de objetos e segmentação de imagens.

2.6.2. Inception_V3

Inception_V3 é uma arquitetura de rede neural convolucional (CNN) proposta por Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe e Jonathon Shlens em seu artigo intitulado "Rethinking the Inception Architecture for Computer Vision" publicado na Conferência de Visão Computacional e Reconhecimento de Padrões (CVPR) de 2016.

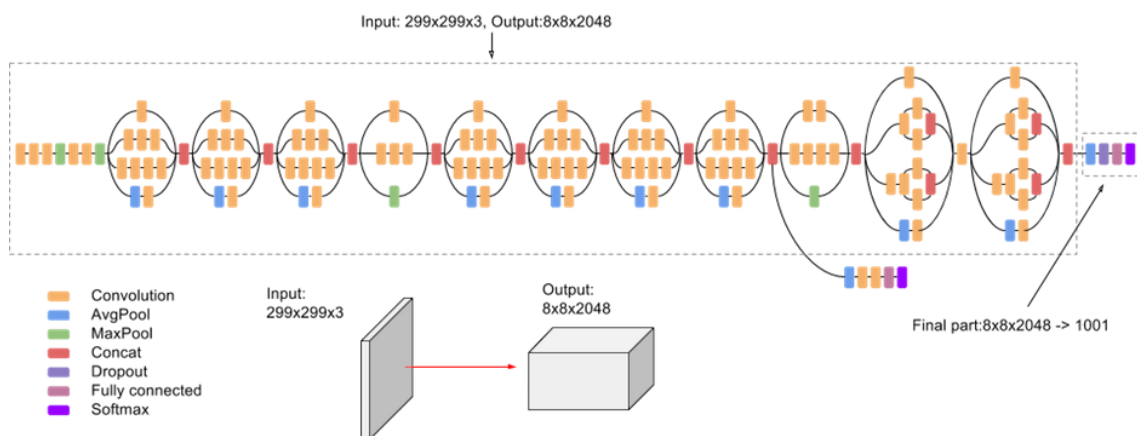


Figura 10. Estrutura do Inception_V3.

Fonte: (T, s.d.).

Inception_V3 é uma versão aprimorada da arquitetura Inception original, que é conhecida por seu uso de módulos de convolução paralela com diferentes tamanhos de filtro para capturar informações em várias escalas e reduzir o número de parâmetros na rede. Inception_V3 possui várias camadas de convolução, normalização, ativação (geralmente ReLU) e pooling, seguidas por camadas de classificação para a saída final.

Uma das principais inovações do Inception_V3 é o uso de técnicas de regularização, como normalização em lotes (Batch Normalization) e regularização L2, para melhorar a capacidade de generalização e evitar o overfitting. Além disso, Inception_V3 também incorpora recursos

de pré-processamento avançados, como o uso de componentes principais de coloração para normalização de cores, que ajuda a melhorar a robustez do modelo a variações de iluminação e cor nas imagens.

Inception_V3 foi pré-treinado em grandes conjuntos de dados, como ImageNet, e tem sido amplamente utilizado para uma variedade de tarefas de visão computacional, como classificação de imagens, detecção de objetos, segmentação de imagens e outras aplicações relacionadas.

2.6.3. ResNet50

ResNet50 é uma arquitetura de rede neural convolucional (CNN) proposta por Kaiming He, Xiangyu Zhang, Shaoqing Ren e Jian Sun em seu artigo intitulado "Deep Residual Learning for Image Recognition" publicado na Conferência de Visão Computacional e Reconhecimento de Padrões (CVPR) de 2016. ResNet50 é uma versão aprimorada da arquitetura ResNet original, que introduziu a ideia de conexões residuais para permitir que as redes mais profundas sejam treinadas com sucesso. ResNet50 possui 50 camadas no total, incluindo camadas de convolução, normalização, ativação (geralmente ReLU) e pooling, seguidas por camadas de classificação para a saída final.

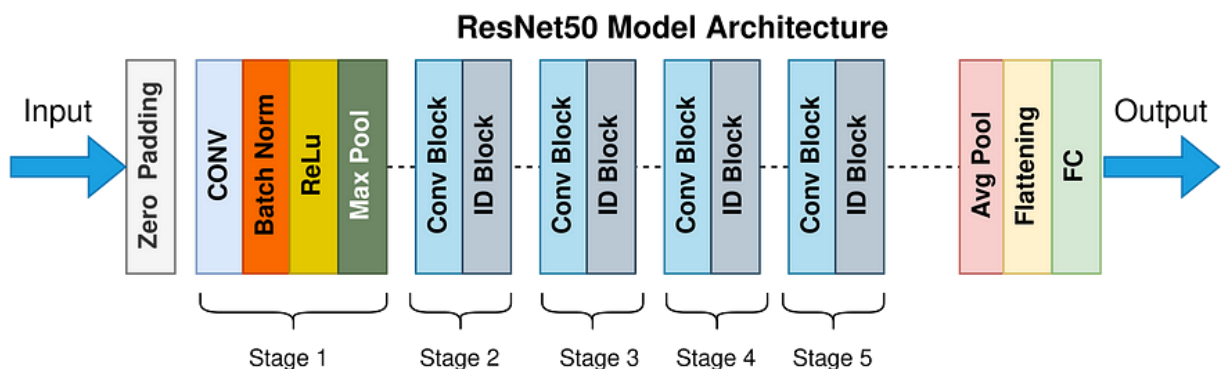


Figura 11. Estrutura do ResNet50.

Fonte: (Mukherjee, 2022).

A principal inovação da ResNet50 é a introdução de blocos residuais que permitem que as informações de entrada passem diretamente para as camadas posteriores, sem passar por várias camadas intermediárias, reduzindo assim o risco de informações "perdidas" durante o processo de treinamento. Essas conexões residuais também permitem que as redes mais profundas sejam treinadas com sucesso, o que era um desafio para as arquiteturas de CNNs

anteriores. A ResNet50 foi pré-treinada em grandes conjuntos de dados, como ImageNet, e tem sido amplamente utilizada em tarefas de visão computacional, como classificação de imagens, detecção de objetos e segmentação de imagens.

2.6.4. Modelo_1

Este modelo é uma arquitetura de rede neural convolucional (CNN) para classificação binária, foi configurado para ser usado em um problema de visão computacional com imagens em escala de cinza, que é a principal característica do dataset escolhido para este trabalho. Aqui estão algumas características importantes do modelo:

Camadas convolucionais: O modelo possui três camadas convolucionais (Conv2D) com 16, 32 e 64 filtros, respectivamente. Cada uma dessas camadas usa uma função de ativação ReLU para introduzir não-linearidade nas representações aprendidas pelas convoluções.

Camadas de pooling: Após cada camada convolucional, há uma camada de pooling (MaxPooling2D) com tamanho de pool (2, 2), que reduz a dimensão espacial das representações aprendidas, ajudando a reduzir o número de parâmetros e a evitar overfitting.

Flatten: Após as camadas de pooling, há uma camada Flatten que transforma as representações aprendidas em um vetor unidimensional, preparando para as camadas densas (fully connected).

Camadas densas: O modelo possui duas camadas densas (Dense) com 64 neurônios cada uma, usando a função de ativação ReLU. A última camada densa possui apenas um neurônio com função de ativação sigmoid, que é usada para a classificação binária (0 ou 1).

Dropout: Entre as duas últimas camadas densas, há uma camada Dropout com uma taxa de 0.5, o que ajuda a regularizar o modelo, reduzindo a chance de overfitting durante o treinamento.

Função de perda e otimizador: O modelo utiliza a função de perda 'binary_crossentropy' para treinamento e é otimizado usando o algoritmo Adam.

Métricas de avaliação: Além da acurácia, o modelo também monitora outras métricas de avaliação durante o treinamento, incluindo precisão (Precision), recall (Recall) e precisão em recall máximo (PrecisionAtRecall) com recall configurado para 1.0, o que é útil para

problemas de desequilíbrio de classes.

Em geral, este parece ser um modelo CNN razoavelmente simples, mas eficaz para classificação binária de imagens em escala de cinza, com a adição de técnicas de regularização (Dropout) e monitoramento de métricas de avaliação importantes para lidar com desequilíbrio de classes. No entanto, a eficácia real do modelo dependerá dos dados de treinamento específicos e do problema em questão. É importante ajustar hiperparâmetros, como taxa de aprendizado, tamanho do batch e número de épocas de treinamento, para obter o melhor desempenho do modelo.

2.7.Trabalhos relacionados

Várias pesquisas têm explorado o uso de técnicas de Deep Learning para a identificação de pneumonia em imagens de raio-x do tórax. A seguir, são apresentados alguns trabalhos relevantes que abordam esse tema.

Aqui está uma tabela com exemplos de trabalhos relacionados ao tema "Uso de Deep Learning para Identificação de Pneumonia em Imagens de Raio-X":

Tabela 1. Trabalhos relacionados.

Autor	Tema do Trabalho	Abordagens	Acurácia	Observações
Wang et al. (2017)	"ChestX-ray8: Hospital-Scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases"	Utilização de uma rede neural convolucional (CNN) para classificar 8 doenças torácicas, incluindo pneumonia, usando uma grande base de dados de raio-x.	Acurácia de 0,834 em identificação de pneumonia.	Trabalho pioneiro que propõe uma base de dados de referência e benchmarks para a tarefa de classificação de doenças torácicas em imagens de raio-x com aprendizado

				fraco.
Rajpurkar et al. (2017)	"CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning"	Desenvolvimento de uma CNN para identificação de pneumonia em imagens de raio-x, treinada em um grande conjunto de dados com anotações de radiologistas.	Acurácia de 0,821 em identificação de pneumonia.	Mostrou-se capaz de superar o desempenho médio de radiologistas em relação à identificação de pneumonia em imagens de raio-x.
Liang et al. (2019)	"Detecting Pneumonia in Chest X-Rays with Hybrid Feature Representations"	Proposição de uma abordagem que combina características de nível de pixel e de nível de região usando uma CNN, com o objetivo de melhorar a detecção de pneumonia em imagens de raio-x.	Acurácia de 0,841 em identificação de pneumonia.	Demonstra que a combinação de características de diferentes níveis pode melhorar a acurácia na identificação de pneumonia em comparação com abordagens que utilizam apenas características de um único nível.
Wang et	"Tienet: Text-Image	Proposição de uma	Acurácia de	Mostra que a

al. (2018)	Embedding Network for Common Thorax Disease Classification and Reporting in Chest X-rays"	arquitetura de rede neural que combina a representação de texto e imagem para classificação de doenças torácicas, incluindo pneumonia, em imagens de raio-x.	0,836 em identificação de pneumonia.	combinação de informações de texto e imagem pode melhorar o desempenho na classificação de doenças torácicas em comparação com abordagens que utilizam apenas informações de imagem.
------------	---	--	--------------------------------------	--

Em 2017, Wang et al. propuseram o "ChestX-ray8: Hospital-Scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases" [1]. Nesse trabalho, os autores criaram uma grande base de dados de imagens de raio-x torácicas, chamada de ChestX-ray8, contendo anotações de radiologistas para 8 doenças torácicas, incluindo pneumonia. Eles treinaram uma rede neural convolucional (CNN) em um ambiente de aprendizado fraco para a classificação e localização dessas doenças. A acurácia alcançada para a identificação de pneumonia foi de 0,834, demonstrando a viabilidade do uso de Deep Learning em um contexto de aprendizado fraco com grandes bases de dados.

Em outro trabalho em 2017, Rajpurkar et al. apresentaram o "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning" [2]. Nesse estudo, os autores desenvolveram uma CNN para a detecção de pneumonia em imagens de raio-x, treinada em um grande conjunto de dados com anotações de radiologistas. A CNN alcançou uma acurácia

de 0,821 na identificação de pneumonia, superando o desempenho médio de radiologistas na mesma tarefa. Isso destacou o potencial do uso de Deep Learning como uma ferramenta auxiliar na detecção de pneumonia em imagens de raio-x.

Em 2019, Liang et al. propuseram uma abordagem que combina características de nível de pixel e de nível de região usando uma CNN para a detecção de pneumonia em imagens de raio-x [3]. A ideia era aproveitar informações de diferentes níveis de detalhes nas imagens para melhorar o desempenho na identificação de pneumonia. A acurácia alcançada foi de 0,841, demonstrando que a combinação de características de diferentes níveis pode resultar em melhorias no desempenho do modelo.

Outra abordagem interessante foi apresentada por Wang et al. em 2018, no trabalho "Tienet: Text-Image Embedding Network for Common Thorax Disease Classification and Reporting in Chest X-rays" [4]. Nesse estudo, os autores propuseram uma arquitetura de rede neural que combina a representação de texto e imagem para a classificação de doenças torácicas, incluindo pneumonia, em imagens de raio-x. A abordagem utilizou informações de texto dos relatórios de radiologistas, juntamente com as imagens de raio-x, para melhorar a acurácia na classificação de doenças torácicas, incluindo pneumonia.

Em resumo, os trabalhos relacionados ao uso de Deep Learning para identificação de pneumonia em imagens de raio-x têm explorado diversas abordagens, incluindo o uso de grandes bases de dados, aprendizado fraco, combinação de características de diferentes níveis e integração de informações de texto e imagem. Esses estudos têm demonstrado resultados promissores

2.8.Considerações gerais

As abordagens baseadas em Deep Learning têm se mostrado promissoras para a identificação de pneumonia em imagens de raio-x do tórax. Os trabalhos relacionados revisados apresentam diferentes abordagens para melhorar a precisão na detecção de pneumonia, incluindo o uso de grandes bases de dados, aprendizado fraco, combinação de características de diferentes níveis e integração de informações de texto e imagem.

É importante observar que esses trabalhos têm algumas limitações, como a necessidade de bases de dados de treinamento grandes e bem anotadas, a possibilidade de viés e erro nas anotações de radiologistas utilizadas como rótulos, a variação na qualidade das imagens de

raio-x e a falta de generalização para diferentes populações e cenários clínicos.

Além disso, é fundamental respeitar as diretrizes éticas e regulatórias no uso de imagens de raio-x e dados de pacientes, garantindo a privacidade e a confidencialidade das informações.

Apesar dessas limitações, os trabalhos revisados evidenciam o potencial das abordagens de Deep Learning para auxiliar na identificação de pneumonia em imagens de raio-x do tórax, podendo ser uma ferramenta complementar aos radiologistas na prática clínica, auxiliando no diagnóstico precoce e no tratamento adequado dos pacientes com pneumonia. Futuras pesquisas podem explorar ainda mais essas abordagens, aprimorando os modelos e validando-os em diferentes configurações clínicas, com o objetivo de melhorar a precisão e a eficácia do diagnóstico de pneumonia em imagens de raio-x.

CAPÍTULO III – METODOLOGIA

No presente capítulo, são apresentados os materiais e métodos utilizados na pesquisa, com o objetivo de alcançar os objetivos propostos. Serão detalhados o tipo de pesquisa realizado, as características do dataset utilizado, o pré-processamento dos dados, o treinamento das Redes Neurais Convolucionais (Convolutional Neural Networks - CNNs) e a análise do desempenho dos modelos implementados.

3.1. Tipo de pesquisa

Segundo (Marques, 2015, p. 47) Pesquisa é um procedimento reflexivo sistemático, controlado e crítico, que permite descobrir novos factos ou dados, relações ou leis.

3.1.1. Quanto a abordagem

Quanto à abordagem é uma pesquisa quantitativa. A pesquisa quantitativa é uma abordagem sistemática para coletar e analisar dados numéricos e estatísticos. Ela é amplamente utilizada em diversas áreas do conhecimento, como ciências sociais, psicologia, educação e saúde. Segundo Creswell (2014), a pesquisa quantitativa é caracterizada pela utilização de métodos estatísticos para analisar os dados coletados e testar hipóteses de pesquisa.

3.1.2. Quanto a natureza

Quanto à natureza esta é uma pesquisa aplicada. A pesquisa aplicada é uma abordagem de pesquisa que tem como objetivo aplicar o conhecimento científico para resolver problemas práticos e atender às necessidades da sociedade e do mercado. Segundo Yin (2015), a pesquisa aplicada é uma abordagem de pesquisa que busca desenvolver soluções concretas para problemas existentes, e que muitas vezes é utilizada em contextos empresariais e governamentais.

3.2. Técnicas e instrumentação de recolha de dados

Quanto à técnica e instrumentação de recolha de dados foi usada a entrevista não estruturada. De acordo com Fontes (2016), a entrevista pode ser estruturada, semi-estruturada ou não estruturada, dependendo do grau de flexibilidade das perguntas. Na entrevista estruturada, as perguntas são pré-determinadas e padronizadas para todos os participantes, enquanto na entrevista não estruturada, as perguntas são mais abertas e não há um roteiro fixo. Já na entrevista semi-estruturada, há um equilíbrio entre perguntas pré-determinadas e perguntas

abertas.

3.3. População

A população da é de 10 pessoas envolvendo, 2 Médicos de clinica geral, 1 Internistas, 3 Técnicos de medicina, 3 Técnicos de radiologia, 1 Pediatra.

3.4. Amostra

Para este trabalho foram usadas imagens de Raio-x baixadas do site *kaggle* para o treinamento das CNN's. Baixou-se as imagens pois o Hospital Provincial de Pemba não possuía nem a quantidade mínima para este estudo.

3.5. Como os dados recolhidos serão tratados

Esta secção detalha o método proposto para encontrar um modelo classificatório para o diagnóstico de pneumonias baseado em imagens de radiografias do tórax. A Figura abaixo descreve os passos da proposta.

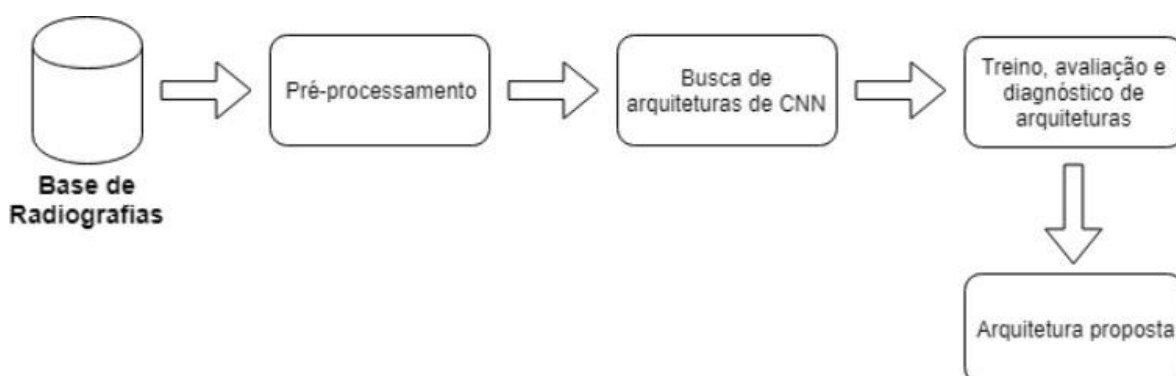


Figura 12. Métodos propostos.

Fonte: Autor

4. Metodologia

O presente trabalho foi feito em 4 fases nomeadamente construção do dataset, pré-processamento do dataset, treinamento das Redes Neurais Convolucionais, e análise do desempenho das Redes Neurais Convolucionais implementadas.

- Construção do dataset:** nesta fase foi feita a construção do dataset, foram colhidos dados da plataforma do Kaggle.
- Pré-processamento do dataset:** nesta fase foi feito o pré-processamento das imagens

que foram utilizadas, consistiu em utilizar técnicas de redimensionamento, e aumento de dados, de modo a obter-se um conjunto de dados suficientes para o treinamento dos modelos.

- c) **Treinamento das CNN's:** nesta fase foram feitas as implementações de cada rede utilizando linguagem de programação Python e a biblioteca Keras que permite a implementação em alto nível de Redes Neurais.
- d) **Análise de Desempenho das Redes Neurais Convolucionais implementadas:** nesta fase após o treinamento das redes escolhidas, serão analisadas as métricas recall, acurácia, precisão e F1_score de cada rede.

3.1. Construção do conjunto de dados (ou dataset)

Nesta fase devido a restrições de acesso e disponibilidade de dados, foram realizados esforços para obter imagens de raio-x de tórax no Hospital Provincial de Pemba, porém, infelizmente, não foi possível obter acesso a esse recurso para inclusão no dataset utilizado neste estudo. É importante destacar que a ausência dessas imagens pode ter limitações potenciais na representatividade do dataset em relação à população estudada e pode afetar a generalização dos resultados obtidos. Esforços alternativos foram tomados para garantir a validade e confiabilidade dos dados utilizados, incluindo a obtenção de imagens de outras fontes confiáveis e a adoção de procedimentos rigorosos de anotação e pré-processamento para mitigar possíveis vieses.

No decorrer da pesquisa, foi identificada a plataforma Kaggle ([kaggle.com](https://www.kaggle.com)), uma comunidade online composta por cientistas de dados e profissionais de Machine Learning. A plataforma Kaggle permite aos usuários encontrar e publicar conjuntos de dados relevantes para suas pesquisas (Kaggle, 2010). A utilização dessa plataforma foi explorada como uma fonte alternativa para obtenção de dados de raio-x de tórax para construção do dataset utilizado neste estudo. A plataforma Kaggle é conhecida por sua ampla variedade de conjuntos de dados de diferentes áreas e pela comunidade ativa de cientistas de dados que compartilham seus dados e conhecimentos. A utilização da plataforma Kaggle pode ter oferecido uma oportunidade valiosa para acessar dados de qualidade para o desenvolvimento do dataset utilizado nesta pesquisa.

Foram obtidas 5856 imagens de raio-x de tórax rotuladas de uma plataforma Kaggle,

contendo 4281 imagens com diagnóstico de pneumonia e 1575 imagens de tórax normais e as restantes 3 para a validação do modelo.

4.1.1. Pré-processamento do dataset

O pré-processamento de imagens é uma etapa crucial no campo de processamento de imagens, e é amplamente abordado em obras de autores renomados como Rafael C. Gonzalez e Richard E. Woods em seu livro clássico "Digital Image Processing". O livro abrange diversas técnicas de pré-processamento de imagens, como filtragem, transformações de intensidade e correção de cores, entre outros tópicos relevantes. Outra referência importante é o livro "Fundamentals of Digital Image Processing" de A.K. Jain, que também explora os fundamentos do pré-processamento de imagens, incluindo técnicas de realce de imagem, filtragem e correção de ruído.

Várias técnicas de pré-processamento de imagens foram utilizadas neste trabalho, a seguir são mencionadas estas técnicas:

Normalização (rescale): A técnica de rescale foi aplicada para normalizar os valores dos pixels das imagens, dividindo-os por 255. Isso é comum em tarefas de processamento de imagens para garantir que todos os valores dos pixels estejam na faixa de 0 a 1, o que pode ajudar no treinamento do modelo.

Rotação (rotation_range): A técnica de rotação foi aplicada com uma faixa de rotação de 10 graus. Essa técnica é usada para aumentar a diversidade do conjunto de treinamento, gerando imagens rotacionadas em diferentes ângulos, o que pode ajudar o modelo a aprender a lidar com imagens em diferentes orientações.

Zoom (zoom_range): A técnica de zoom foi aplicada com uma faixa de zoom de 0.1. Essa técnica permite a geração de imagens com diferentes níveis de zoom, o que pode ajudar o modelo a aprender a lidar com variações de escala nas imagens de entrada.

Deslocamento horizontal e vertical (width_shift_range e height_shift_range): As técnicas de deslocamento horizontal e vertical foram aplicadas com uma faixa de deslocamento de 0.1. Essas técnicas permitem a geração de imagens com deslocamento horizontal e vertical aleatório, o que pode ajudar o modelo a aprender a lidar com variações na posição dos objetos nas imagens.

Espelhamento horizontal (horizontal_flip): A técnica de espelhamento horizontal foi aplicada, o que pode ajudar o modelo a aprender a lidar com imagens espelhadas horizontalmente, como imagens de objetos em diferentes orientações.

Modo de cor (color_mode): O modo de cor foi definido como 'grayscale', o que significa que as imagens de entrada são convertidas para escala de cinza. Isso pode ser útil em tarefas onde a cor não é uma informação relevante que é o caso deste estudo, e também pode ajudar a reduzir a quantidade de dados necessários para treinar o modelo.

Depois de aplicadas todas essas técnicas obteve-se o seguinte resultado:

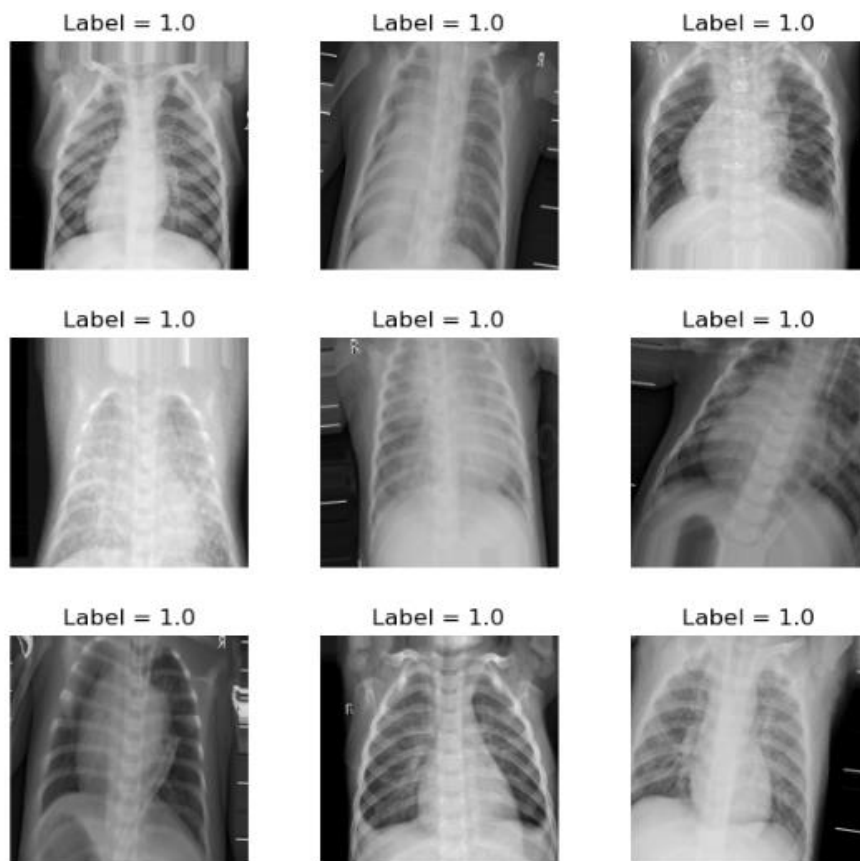


Figura 13. Imagens de Raio-X de tórax depois de aplicadas as técnicas de pré – processamento.

Fonte: Autor.

Essas técnicas de pré-processamento e aumento de dados foram aplicadas para preparar os dados de treinamento e validação antes de alimentá-los em um modelo de aprendizado de

máquina, com o objetivo de melhorar a capacidade do modelo em generalizar para novos dados e melhorar o desempenho do modelo.

4.2.Implementação

Nesta secção serão abordadas as plataformas, bibliotecas e as métricas de validação de modelos que foram usadas para a implementação dos modelos descritos na secção 2.6. O trabalho faz teste de algumas arquitecturas bem conhecidas para o processamento de imagens, as arquitecturas implementadas usando a linguagem Python, o framework TensorFlow e a biblioteca Keras. A tabela 2 faz uma descrição do número de camadas por cada um dos algoritmos implementados nessa pesquisa, onde o modelo mais profundo é o modelo Modelo_1 que é um modelo genérico criado pelo autor e o menos profundo é o ResNet50.

Tabela 2. Descrição dos modelos implementados.

Modelo	Número de Camadas
Inception_v3	48
DenseNet121	121
ResNet50	50
Modelo_1	10

4.2.1.1.Materiais

As bibliotecas usadas para implementação dos modelos de inteligência artificial foram as seguintes:

- a) O TensorFlow é uma biblioteca de código aberto amplamente utilizada para aprendizado de máquina e desenvolvimento de modelos de aprendizado profundo. Ele oferece uma ampla gama de ferramentas e recursos para construir, treinar e implantar modelos de aprendizado de máquina em diversas aplicações, incluindo visão computacional, processamento de linguagem natural, reconhecimento de fala, robótica e muito mais. O TensorFlow é conhecido por sua flexibilidade, escalabilidade e suporte a várias plataformas, o que o torna uma escolha popular entre os pesquisadores e desenvolvedores de aprendizado de máquina (Abadi et al., 2016).
- b) OpenCV (Open Source Computer Vision Library) é uma biblioteca popular de código

aberto utilizada para processamento de imagens e visão computacional. Ela fornece uma ampla gama de funções e algoritmos para realizar tarefas como leitura e gravação de imagens e vídeos, manipulação de pixels, detecção de objetos, rastreamento de movimento, entre outras aplicações. O OpenCV é amplamente utilizado em projetos de pesquisa e aplicativos de visão computacional em diversas áreas, como robótica, automação industrial, medicina, realidade aumentada e veículos autônomos (OpenCV, 2021).

- c) Keras é uma biblioteca de alto nível para aprendizado de máquina e deep learning em Python, que oferece uma interface simplificada e amigável para a construção, treinamento e avaliação de modelos de redes neurais. Ela é amplamente utilizada pela comunidade de ciência de dados e aprendizado de máquina devido à sua flexibilidade, modularidade e facilidade de uso, tornando o processo de criação e treinamento de modelos de aprendizado de máquina mais acessível para desenvolvedores e pesquisadores (Chollet, 2015).
- d) Scikit-learn é uma biblioteca de aprendizado de máquina em Python que oferece uma ampla gama de algoritmos de classificação, regressão, clusterização e outras técnicas de mineração de dados. Essa biblioteca é amplamente utilizada em projetos de aprendizado de máquina devido à sua facilidade de uso, documentação abrangente e eficiente implementação de algoritmos populares. É uma ferramenta poderosa para a análise de dados e construção de modelos de aprendizado de máquina em uma ampla variedade de aplicações. (Pedregosa et al., 2011).
- e) Python é uma linguagem de uso gratuito e de código-fonte aberto, conta com uma vasta biblioteca padrão e documentação que possibilitam que muitas coisas sejam feitas sem dependências adicionais. É uma linguagem de alto nível interpretada, orientada a objecto dinâmica e flexível.
- f) Kaggle é um site comunitário conhecido para cientistas de dados competirem em aprendizado de máquina.
- g) Matplotlib e panda – Biblioteca para criação de gráficos e visualizações de dados.
- h) Numpy – Biblioteca especializada para arrays e matrizes em Python.

4.2.1.2. Métodos de Validação de classificação

A validação de modelos é uma etapa crucial no processo de aprendizado de máquina para avaliar a capacidade do modelo em generalizar e sua performance. Uma das ferramentas comumente utilizadas nesse processo é a construção de uma matriz de confusão, que é uma tabela resumida que exhibe as previsões do modelo para cada uma das classes. A matriz de confusão fornece informações importantes sobre a capacidade do modelo em fazer previsões corretas, identificando os verdadeiros positivos, falsos positivos, falsos negativos e verdadeiros negativos. Essa análise ajuda a compreender os erros de previsão do modelo e a identificar áreas que podem ser aprimoradas para melhorar a qualidade das previsões.

A tabela 3 é um exemplo:

Tabela 3. Matriz de confusão.

	Classe Prevista: Sim	Classe Prevista: Não
Classe Real: Sim	Verdadeiro Positivo (VP)	Falso Positivo (FP)
Classe Real: Não	Falso Negativo (FN)	Verdadeiro Negativo (VN)

A matriz de confusão é uma tabela que resume as previsões feitas pelo modelo para cada uma das classes em um problema de classificação. Ela possui quatro elementos principais:

1. **Verdadeiro Positivo (VP):** Número de casos onde o modelo previu corretamente a classe positiva e o valor real era também positivo.
2. **Falso Positivo (FP):** Número de casos onde o modelo previu incorretamente a classe positiva, mas o valor real era na verdade negativo.
3. **Falso Negativo (FN):** Número de casos onde o modelo previu incorretamente a classe negativa, mas o valor real era na verdade positivo.
4. **Verdadeiro Negativo (VN):** Número de casos onde o modelo previu corretamente a classe negativa e o valor real era também negativo.

A partir dos valores na matriz de confusão, é possível calcular várias métricas de avaliação,

como acurácia, precisão, recall, F1-score, entre outras, para avaliar a capacidade de generalização e performance do modelo de aprendizado de máquina. A matriz de confusão é uma ferramenta valiosa para entender os erros de previsão do modelo e identificar áreas de melhoria.

a) Precisão

A precisão é uma métrica de avaliação de modelos de aprendizado de máquina que mede a proporção de previsões corretas feitas pelo modelo em relação ao total de previsões realizadas. Ela é calculada como o quociente entre o número de verdadeiros positivos (VP) e a soma dos verdadeiros positivos e falsos positivos (VP + FP). Em outras palavras, a precisão é a capacidade do modelo em fazer previsões corretas para uma classe específica, em relação a todas as previsões feitas para essa classe, sejam elas corretas ou incorretas.

$$\text{Precisão} = \frac{VP}{(VP + FP)}$$

Uma alta precisão indica que o modelo está fazendo poucas previsões incorretas para uma determinada classe, enquanto uma baixa precisão indica que o modelo está fazendo muitas previsões incorretas para essa classe.

b) Acurácia

A acurácia é uma métrica de avaliação de modelos de aprendizado de máquina que mede a proporção de previsões corretas feitas pelo modelo em relação ao total de previsões realizadas, considerando todas as classes. Ela é calculada como o quociente entre o número total de previsões corretas (Verdadeiro Positivo + Verdadeiro Negativo) e o número total de previsões (Verdadeiro Positivo + Falso Positivo + Verdadeiro Negativo + Falso Negativo).

A fórmula para calcular a acurácia é:

$$\text{Acurácia} = \frac{(VP + VN)}{(VP + FP + VN + FN)}$$

c) Recall

O Recall, também conhecido como Sensibilidade ou Taxa de Verdadeiro Positivo, é uma métrica de avaliação de modelos de aprendizado de máquina que mede a proporção de casos

positivos corretamente identificados pelo modelo em relação ao total de casos positivos reais. Em outras palavras, o Recall indica a capacidade do modelo em identificar corretamente os casos de uma determinada classe.

$$\text{Recall} = \frac{VP}{(VP + FN)}$$

d) F1_Score

O F1-Score é uma métrica de avaliação de modelos de aprendizado de máquina que combina a precisão e o recall em uma única medida, sendo útil quando se busca um equilíbrio entre essas duas métricas. É a média harmônica entre a precisão e o recall, e é calculado pela seguinte fórmula:

$$\text{F1_Score} = \frac{2 * (\text{Precisão} * \text{Recall})}{(\text{Precisão} + \text{Recall})}$$

4.3.Treinamento das CNN's

Ao treinar redes neurais convolucionais (CNNs) para tarefas de processamento de imagens, é fundamental ajustar hiperparâmetros como tamanho do lote, número de épocas e taxas de aprendizado para otimizar o desempenho do modelo (Goodfellow, I., Bengio, Y., & Courville, A., 2016)

Os modelos foram treinados na plataforma Kaggle, que é baseada em Jupyter. Essa plataforma foi escolhida por ser gratuita e por oferecer aceleradores que permitem um treinamento rápido de modelos de Aprendizado de Máquina, foram treinados também utilizando GPUs como aceleradores, o que permite um treinamento mais rápido e eficiente em comparação com o uso de apenas a CPU. O treinamento dos modelos foi realizado nos kernels da plataforma Kaggle, utilizando uma GPU para acelerar o processamento, e utilizando uma memória RAM de 16 GB.

4.4.Avaliação dos Classificadores

A técnica de avaliação dos classificadores usada neste trabalho é a avaliação em um conjunto de teste separado.

De acordo com Hastie, Tibshirani e Friedman (2009) em "The Elements of Statistical

Learning: Data Mining, Inference, and Prediction' publicado pela Springer, a avaliação em um conjunto de teste separado é uma técnica comum para estimar a capacidade de generalização de um modelo de aprendizado de máquina. Um conjunto de teste independente, que não foi usado durante o treinamento do modelo, é reservado para avaliar o desempenho do modelo em dados não vistos. Isso permite uma estimativa mais realista do desempenho do modelo em dados futuros e pode ajudar a evitar problemas de superestimação de desempenho associados à avaliação em conjuntos de dados de treinamento.

Após treinar o modelo, ele é avaliado no conjunto de teste usando as métricas de perda (loss), acurácia (accuracy), precisão (precision), recall (recall) e F1-score (F1 Score). Essas métricas são comumente utilizadas na avaliação de modelos de classificação para medir o desempenho do modelo em termos de sua capacidade de fazer previsões corretas, minimizar erros de classificação e equilibrar a precisão e recall. A utilização de um conjunto de teste separado permite avaliar a capacidade de generalização do modelo, ou seja, sua capacidade de fazer previsões precisas em dados não vistos durante o treinamento.

CAPÍTULO IV: ANÁLISE E DISCUSSÃO DOS RESULTADOS

Neste capítulo, foram apresentados e discutidos os resultados obtidos ao utilizar técnicas de Deep CNN's para identificar pneumonia em imagens de Raio-x. A utilização de GPUs disponíveis na plataforma Kaggle possibilitou o treinamento dos modelos em um tempo reduzido, levando em média 3 horas. Antes do treinamento, as imagens foram processadas, sendo reduzidas para o tamanho de 256x256 e normalizadas para valores entre 0 e 1. Para avaliação dos modelos, foi aplicada a técnica de a avaliação em um conjunto de teste separado. Em seguida, foi calculada a média das métricas utilizadas para validar os modelos. Em resumo, o uso de GPUs e técnicas de processamento de imagens possibilitou treinar os modelos de forma eficiente e a técnica de avaliação em um conjunto de teste separado foi aplicada para avaliar a performance dos modelos.

Tabela 4.Resultados classificação geral.

Modelo	Accuracy	F1 Score	Precision	Recall
Inception_v3	0.87	0.89	0.88	0.91
DenseNet121	0.90	0.92	0.94	0.90
ResNet50	0.89	0.92	0.87	0.97
Modelo_1	0.95	0.85	0.96	0.98

Quando se utiliza o Graphic Processing Unit (GPU) para treinar modelos de aprendizado de máquina, é importante escolher um tamanho de lote (batch_size) que seja um divisor exato do número total de imagens no conjunto de treinamento, de forma que não haja restos na divisão. No caso específico mencionado, o conjunto de treinamento é composto por 5216 imagens.

Para evitar esgotar a memória disponível durante o treinamento, foram escolhidos tamanhos de lote de 32 imagens e um número de épocas de treinamento de 25. Esses valores foram selecionados levando em consideração a capacidade de processamento dos modelos.

De acordo com a literatura de Goodfellow et al. (2017), a técnica de normalização em lote (batch normalization) atua na padronização da média e variância de cada unidade de uma rede neural, com o objetivo de estabilizar o processo de aprendizado. Nas arquiteturas de modelos estudadas aqui, a técnica de normalização em lote foi aplicada após todas as camadas de

convolução, resultando em uma melhora no desempenho e na redução do overfitting nos modelos.

A aplicação da técnica de normalização em lote torna o modelo mais estável, permitindo que ele alcance altas precisões nos dados de validação, chegando a até 97%. Esses resultados são consistentes com a literatura e destacam os benefícios da normalização em lote para melhorar o desempenho de modelos de aprendizado de máquina.

A seguir são apresentados os resultados de cada arquitetura e por fim é feita a análise e discussão dos resultados.

4.1.Desempenho do Inception_V3

Os gráficos ilustrados nas figuras 13 e 14 mostram o comportamento da acurácia de treinamento e a perda do modelo Inception_V3 obtido pela execução da do treinamento em 25 épocas, e implementando a técnica de aumento de dados para as imagens de treinamento. De acordo com Géron (2019), a técnica de aumento de dados é amplamente utilizada em deep learning para aumentar o tamanho do conjunto de dados de treinamento e melhorar o desempenho do modelo.

A ideia é gerar novas amostras de dados a partir das amostras existentes, aplicando transformações aleatórias como rotação, translação, corte, inversão, ampliação, entre outras. Essas transformações permitem que o modelo veja a mesma imagem de várias maneiras diferentes e aprenda recursos mais robustos e invariantes à variações na posição, tamanho e orientação dos objetos presentes nas imagens. Como afirmou Géron (2019), a técnica de aumento de dados ajuda a reduzir o overfitting, que ocorre quando o modelo aprende a representação dos dados de treinamento tão bem que não consegue generalizar para dados não vistos durante o treinamento.

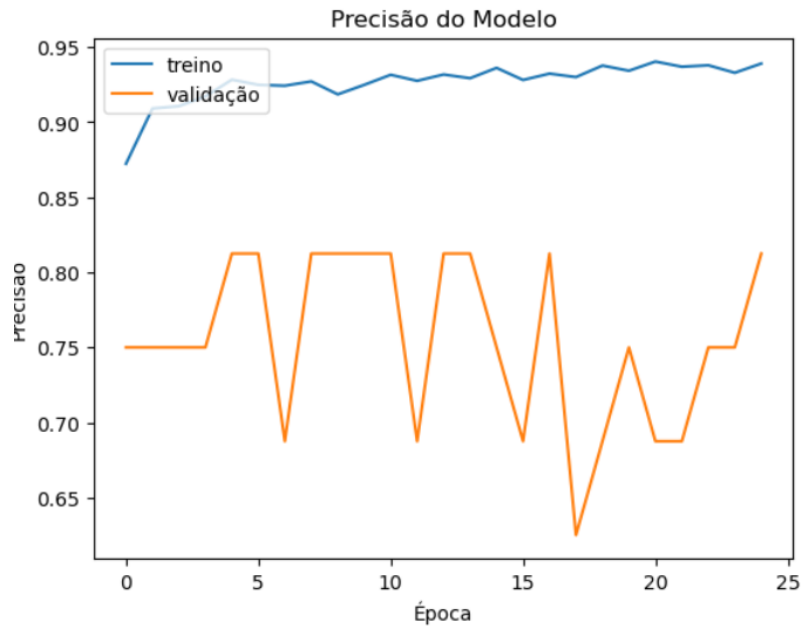


Figura 14. Precisão do modelo Inception_V3.

Fonte: Autor.

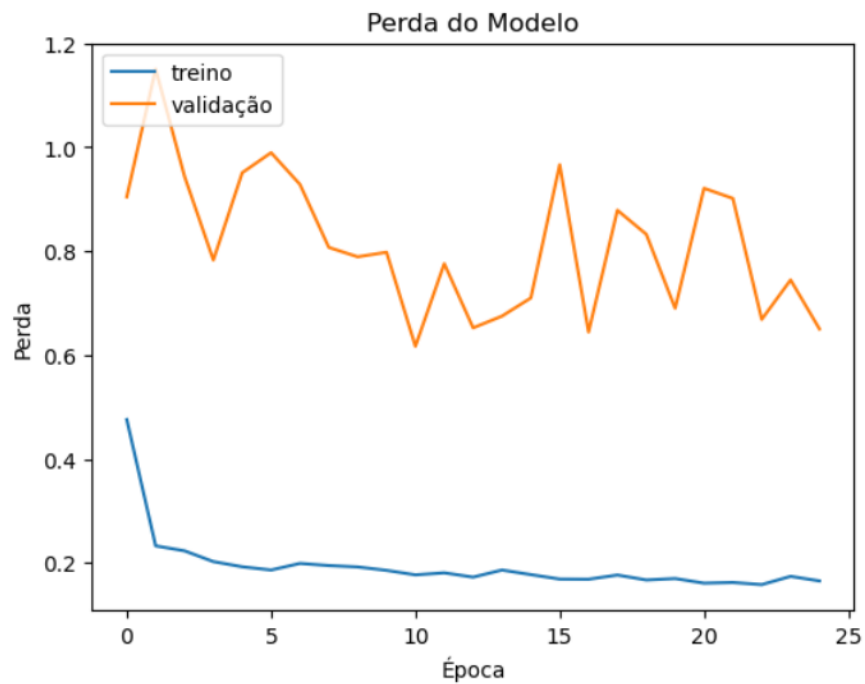


Figura 15. Perda do modelo Inception_V3.

Fonte: Autor.

Nesta fase mesmo usando todas essas técnicas o modelo não apresentou bons resultados como pode se ver no gráfico 13 onde é possível perceber que a partir da época 15 até aproximadamente a época 17 o modelo sofre uma descida acentuada, na linha de tendência de validação em relação a linha que representa o treino, essa linha de

validação tende a subir quando se aproxima a época 25.

Após o treinamento fez-se o teste do modelo treinado com dados não vistos e este obteve uma acurácia de 85.7 % o que não seria um mau resultado se a Matriz de confusão não mostrasse o inverso.

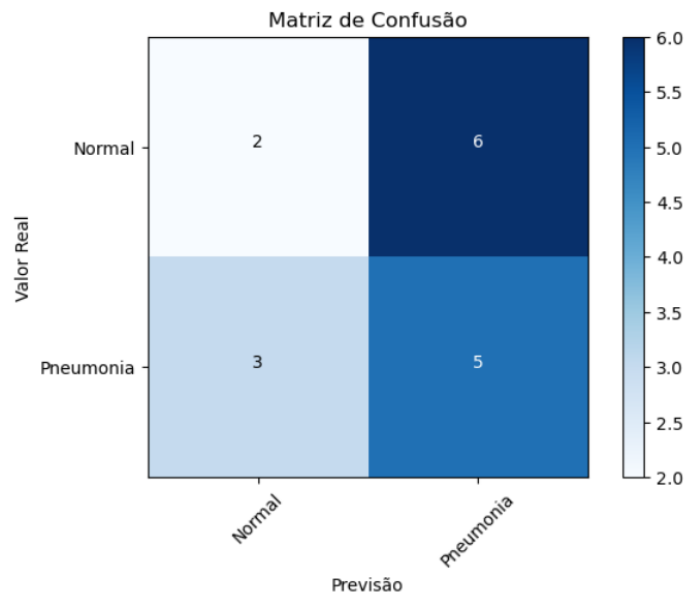


Figura 16. Matriz de confusão Inception_V3.

Fonte: Autor.

Na Matriz de confusão o Verdadeiro Positivo e previsto apenas 3 vezes, Falso Negativo e previsto 5 vezes, Verdadeiro Negativo e previsto 6 vezes e por fim têm o Falso Positivo que acontece apenas 2x, a seguir segue a tabela contendo a métricas da avaliação final do modelo.

Tabela 5. Resultado do teste do modelo

Modelo	Accuracy	F1 Score	Precision	Recall
Inception_v3	0.85	0.53	0.45	0.62

Com base no apresentado na tabela se teve a certeza de que algo não estava bem com o modelo então decidiu-se aumentar o número de épocas de treinamento assim a seguir segue os resultados dos testes depois de aumentar o número de épocas.

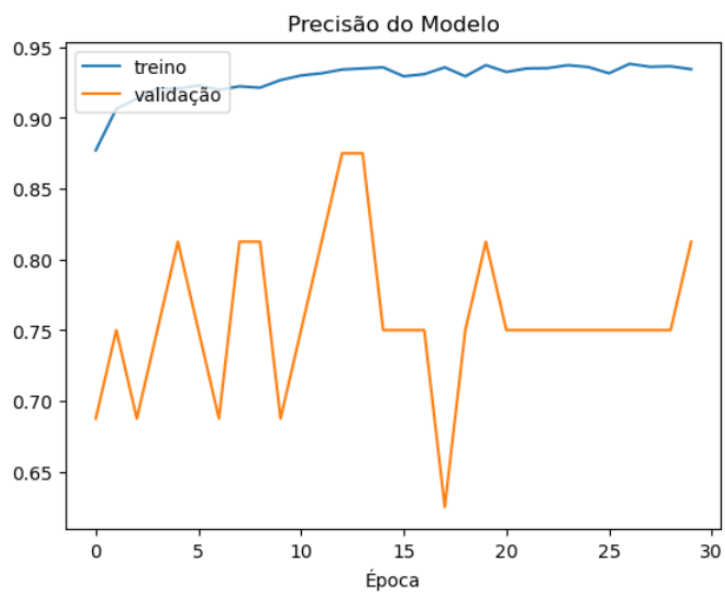


Figura 17. Gráfico de precisão do Inception_V3.

Fonte: Autor.

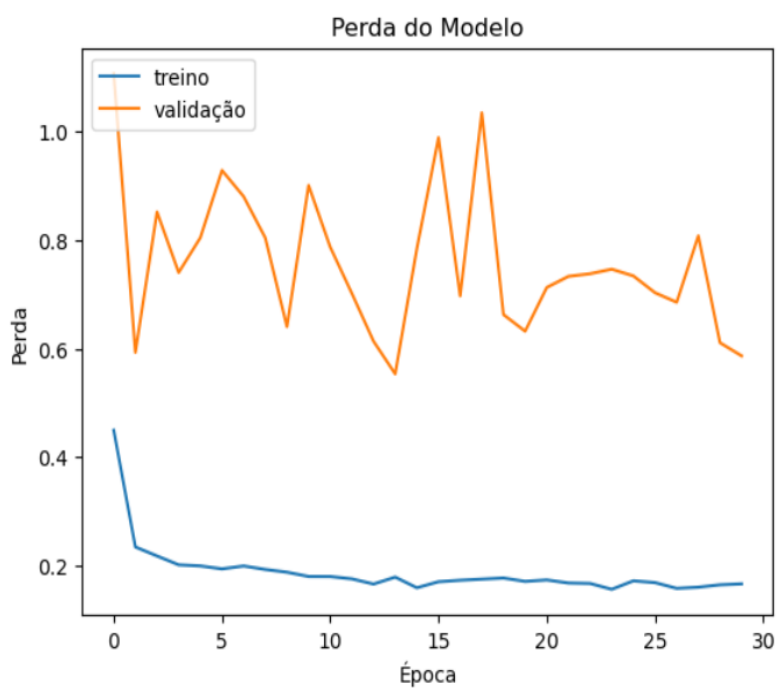


Figura 18. Gráfico de perda do Inception_V3.

Fonte: Autor.

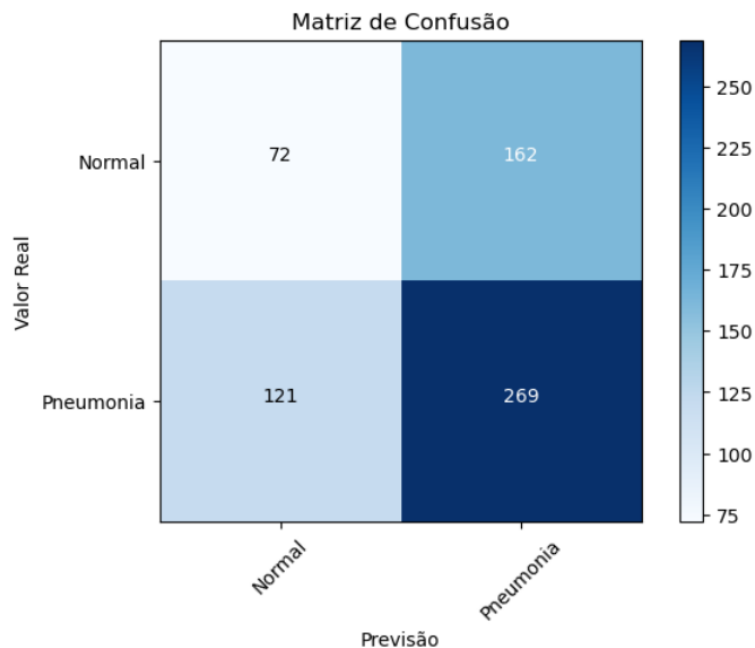


Figura 19. Matriz de confusão Inception_V3.

Fonte: Autor.

Depois de aumentado o número de épocas obteve-se os seguintes resultados, representados na tabela 6:

Tabela 6. Resultados do teste do modelo Inception_V3.

Modelo	Accuracy	F1 Score	Precision	Recall
Inception_v3	0.87	0.89	0.88	0.91

Foi possível constatar que houve uma melhoria significativa.

4.2.Desempenho do DenseNet121

A tabela 7 apresenta os resultados do desempenho de treinamento da arquitectura DenseNet121, que possui 121 camadas como o próprio nome sugere.

Após o treinamento, fez-se o teste do modelo treinado com dados não vistos e obteve-se uma acurácia de 90%. Na implementação do modelo foi aplicado as mesmas técnicas de treinamento usadas no modelo anterior o Inception_V3, e testou-se o treinamento com 25 épocas. A seguir é apresentado os resultados obtidos do treinamento da arquitectura DenseNet121, a matriz de confusão e os gráficos de precisão do modelo e de perda do modelo.

Tabela 7. Resultados do teste do modelo Inception_V3.

Modelo	Accuracy	F1 Score	Precision	Recall
DenseNet121	0.90	0.92	0.94	0.90

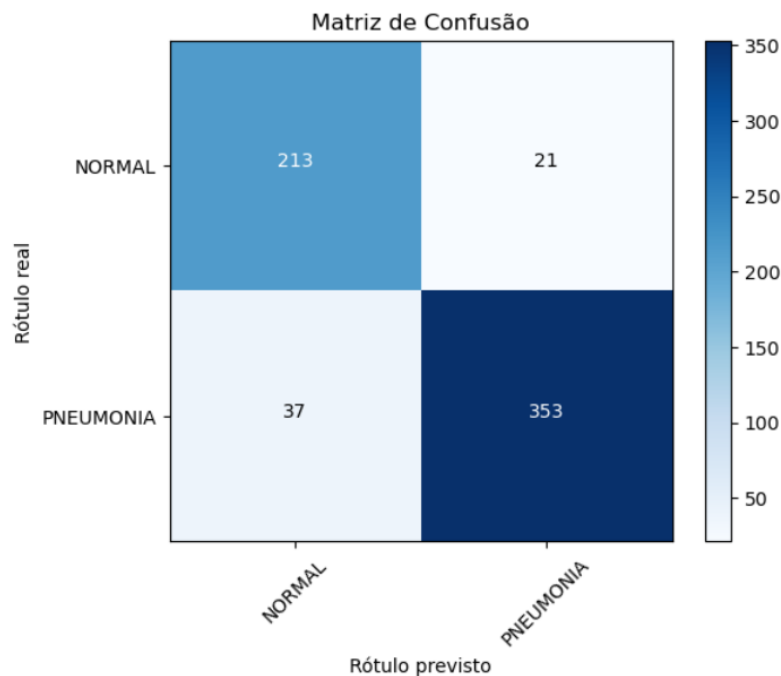


Figura 20. Matriz de confusão DenseNet121.

Fonte: Autor.

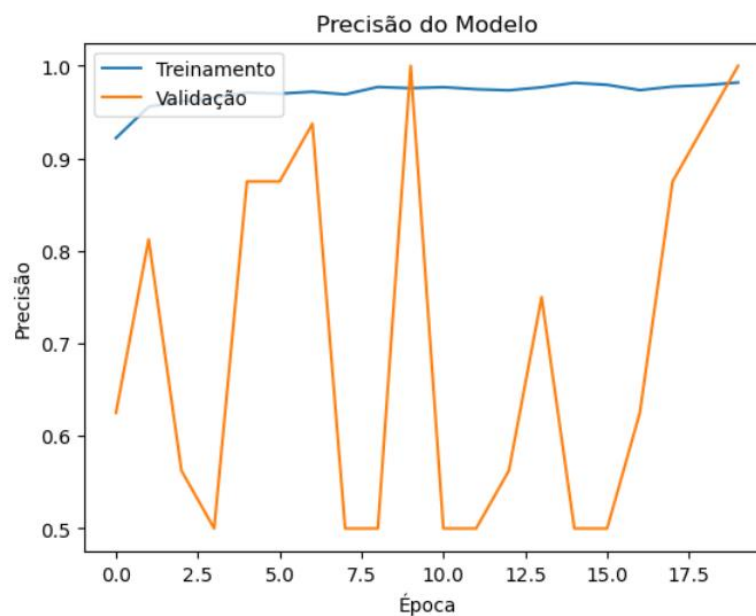


Figura 21. Gráfico de precisão DenseNet121.

Fonte: Autor.

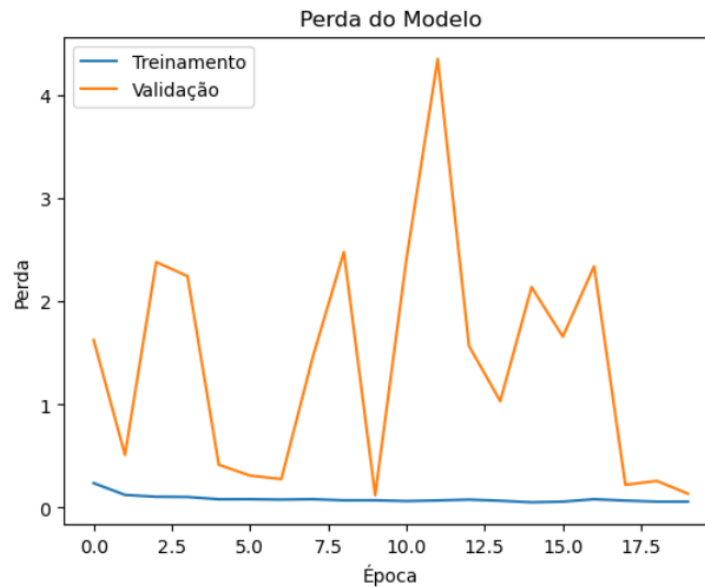


Figura 22. Gráfico de perda DenseNet121.

Fonte: Autor.

4.3.Desempenho do Resnet50

De acordo com Kaiming He et al. (2016), a ResNet50 é uma arquitetura de rede neural convolucional profunda que é capaz de superar as limitações de profundidade das redes convolucionais tradicionais. A ResNet50 utiliza conexões residuais para permitir que a informação flua diretamente através da rede, sem ser afetada pela camada de ativação, e dessa forma, permite o treinamento de modelos mais profundos e precisos. Essas conexões residuais ajudam a resolver o problema de desaparecimento de gradiente que ocorre quando a rede se torna muito profunda.

Após o treinamento, fez-se o teste do modelo treinado com dados não vistos e este obteve uma acurácia de 89%. As técnicas mencionadas na subsecção 4.1. foram aplicadas neste modelo.

A seguir são ilustrados os resultados de treinamento da arquitectura ResNet50 e as figuras 22, 23 representando a matriz de confusão e gráfico de acurácia.

Tabela 8. Resultados teste do modelo ResNet50.

Modelo	Accuracy	F1 Score	Precision	Recall
ResNet50	0.89	0.92	0.87	0.97

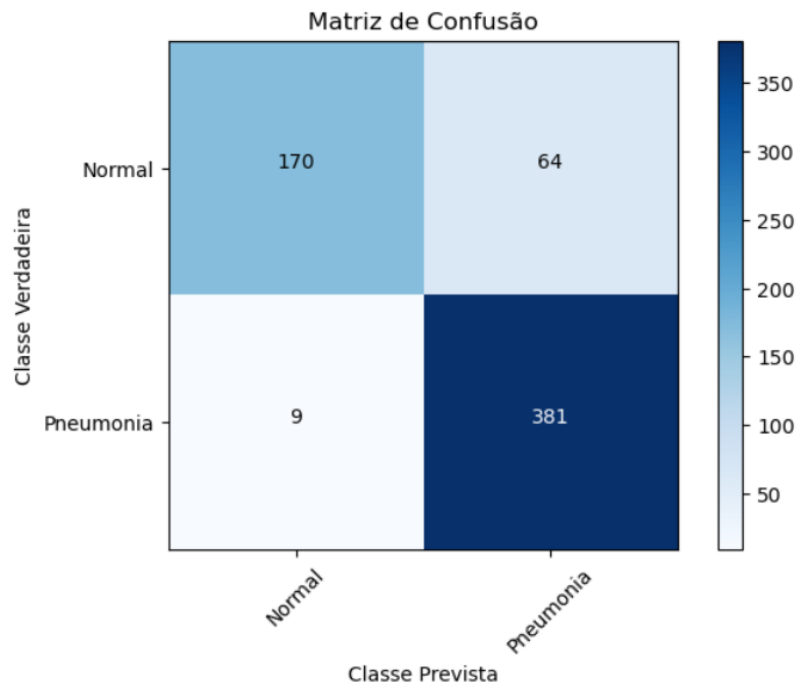


Figura 23. Matriz de confusão ResNet50.

Fonte: Autor.

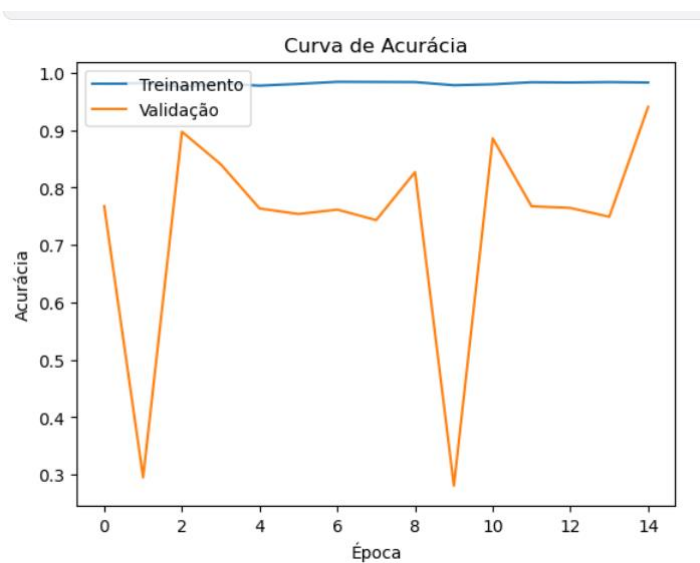


Figura 24. Gráfico de acurácia ResNet50.

Fonte: Autor.

4.4.Desempenho do Modelo_1

Este é um modelo de rede neural convolucional (CNN) para classificação binária de imagens. Ele tem uma arquitetura bastante comum para problemas de classificação de imagem, que consiste em camadas convolucionais e camadas de pooling, seguidas por camadas densas para

a classificação final.

O modelo começa com uma camada convolucional com 16 filtros, seguida por uma camada de pooling para reduzir a resolução da imagem pela metade. Depois, há mais duas camadas convolucionais e de pooling, com o número de filtros dobrando em cada camada (32 e 64).

Foi adicionada uma camada de "flatten" para transformar os mapas de características 3D em um vetor 1D, seguida por duas camadas densas, com a segunda camada tendo apenas um neurônio com função de ativação sigmoid para a classificação binária. Este modelo também usa regularização de dropout com uma taxa de 50% para evitar overfitting durante o treinamento. O modelo é compilado com a função de perda "binary_crossentropy" e o otimizador "adam".

O modelo é treinado com um conjunto de treinamento e validação, usando o gerador de fluxo de dados para alimentar as imagens de forma eficiente para o modelo. O treinamento é feito por 25 épocas, com o número de etapas por época sendo definido pelo número total de imagens no conjunto de treinamento dividido pelo tamanho do lote de treinamento. Além disso, um objeto de histórico é passado como um callback para coletar as métricas de treinamento e validação em cada época.

Uma característica principal deste modelo é que ele é diferente dos anteriores não é pré-treinado. Ele é definido do zero e treinado do início ao fim com o conjunto de dados de treinamento fornecido. Se fosse um modelo pré-treinado, significaria que ele já havia sido treinado em um grande conjunto de dados de imagem e seria usado para transferir o conhecimento aprendido para uma nova tarefa, como a classificação de uma imagem específica. No entanto, neste caso, o modelo foi treinado do zero para a tarefa de classificação de imagens específicas.

Após o treinamento, fez-se o teste do modelo treinado com dados não vistos e este alcançou uma acurácia de 95%. A seguir são ilustrados os resultados nas figuras 24, 25 e 26.

Tabela 9. Resultados teste do modelo Modelo_1.

Modelo	Accuracy	F1 Score	Precision	Recall
Modelo_1	0.95	0.85	0.96	0.98

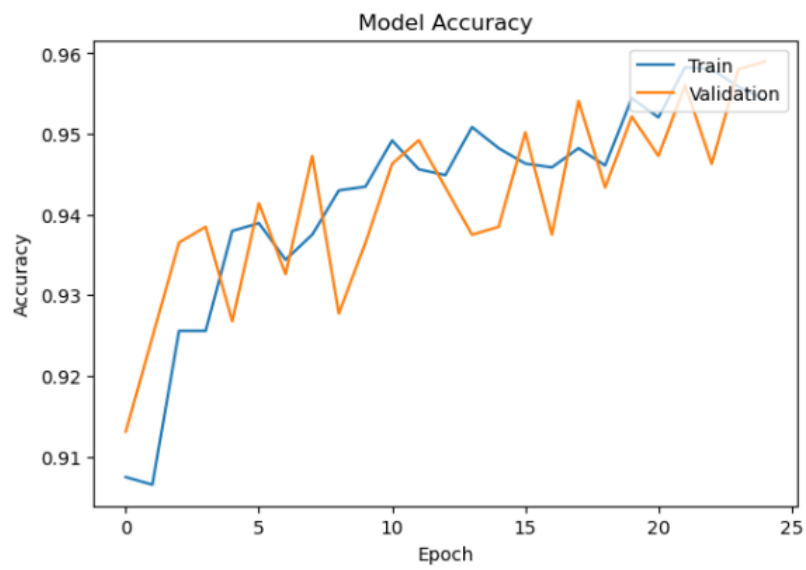


Figura 25. Gráfico de acurácia Modelo_1.

Fonte: Autor.

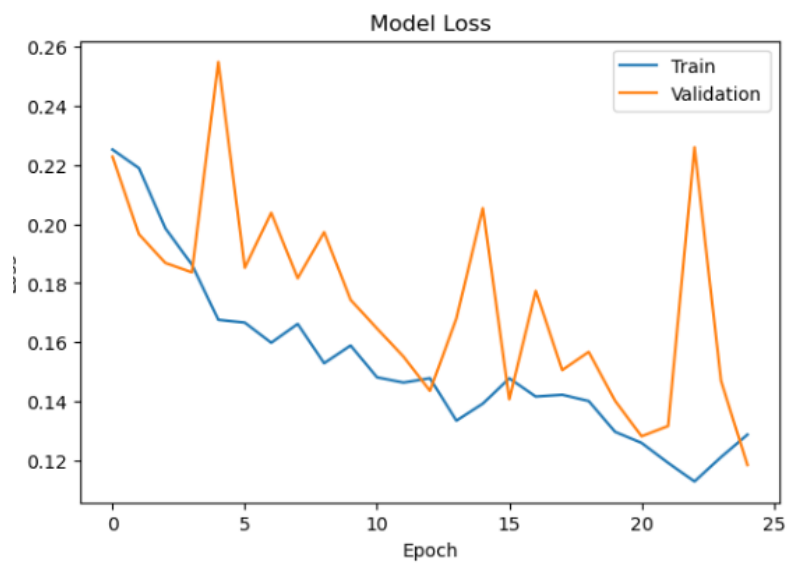


Figura 26. Gráfico de perda Modelo_1.

Fonte: Autor.

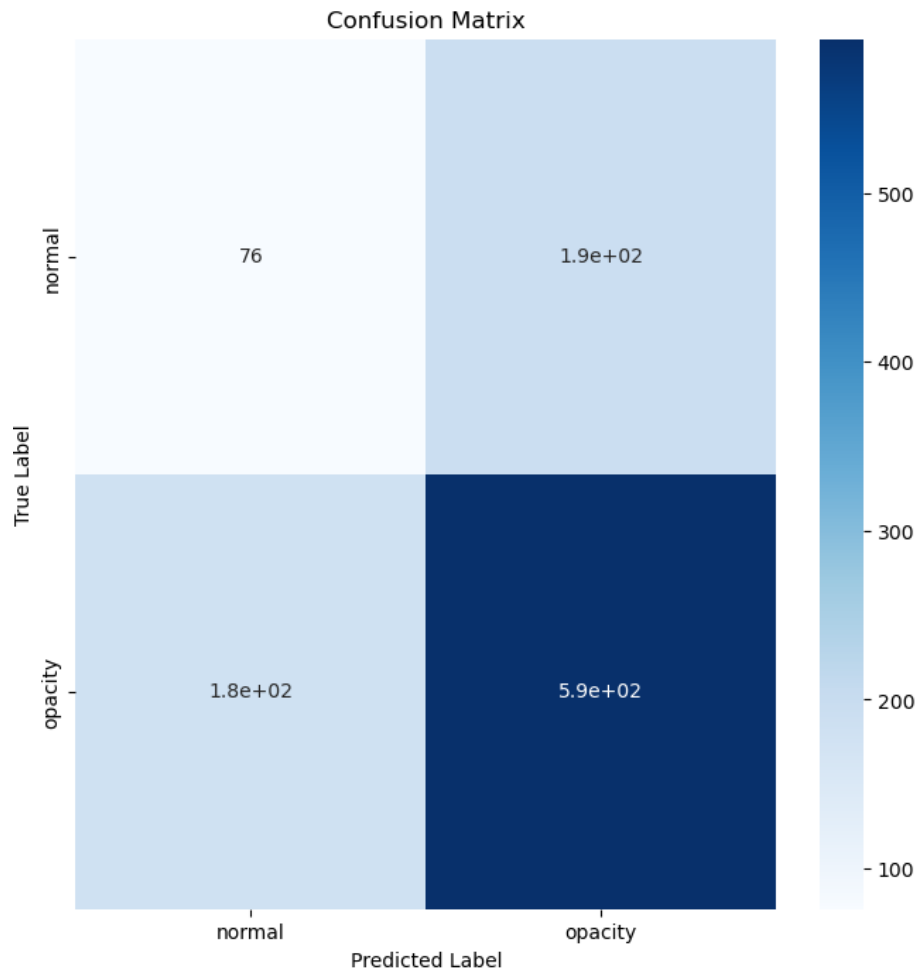


Figura 27. Matriz de confusão Modelo_1.

Fonte: Autor.

4.5.Discussão

No experimento, foram seleccionadas quatro arquitecturas **de Deep Convolutional Neural Networks** para a classificação de Pneumonia, sendo que três foram pré-treinadas e uma foi construída do zero. As arquitecturas escolhidas foram Inception_V3, DenseNet121, ResNet50 e o Modelo_1. Para avaliar o desempenho dos modelos, foram utilizadas métricas importantes como a acurácia, recall, precisão e F1 score.

Após o treinamento e teste, foi observado que o Modelo_1 teve o melhor desempenho, com 95% de acurácia e uma taxa de erro menor, indicando que este modelo teve uma maior precisão em classificar correctamente a Pneumonia. Por outro lado, o modelo Inception_V3 apresentou a menor acurácia, com 87%. Vale ressaltar que algumas imagens de radiografia de tórax contêm Pneumonia em estágios iniciais, o que pode causar confusão no processo de

treinamento. No entanto, a medição global mostrou que o modelo não apresentou overfit ou underfit.

É importante destacar que o desempenho dos modelos pode variar dependendo da qualidade dos dados e do tamanho do conjunto de treinamento. Além disso, as métricas utilizadas para avaliar o desempenho podem ter limitações e é necessário ter cuidado ao interpretá-las. Portanto, é sempre importante realizar experimentos rigorosos e avaliar diferentes métricas para obter uma avaliação mais precisa do desempenho dos modelos.

No geral, o experimento apresentou resultados promissores e sugere que o Modelo_1 pode ser uma boa opção para a classificação de Pneumonia em imagens de radiografia de tórax. No entanto, mais pesquisas são necessárias para validar esses resultados e explorar outras arquiteturas de rede neural que podem ser mais adequadas para esse tipo de tarefa.

CAPÍTULO V: CONCLUSÃO

Considerando os resultados apresentados neste estudo, é possível concluir que o uso de modelos de Deep Convolutional Neural Networks pode ser uma ferramenta valiosa na identificação de Pneumonia em imagens de radiografia de tórax. O *modelo_1* obteve a melhor acurácia, indicando uma maior precisão em classificar corretamente os casos de Pneumonia.

É importante destacar que a identificação precoce de Pneumonia é fundamental para um tratamento eficaz, e a utilização de modelos de Deep Learning pode auxiliar os profissionais de saúde na triagem e diagnóstico preciso de casos suspeitos. No entanto, é necessário levar em conta as limitações do modelo, como a qualidade dos dados e a diversidade das imagens de radiografia de tórax.

Em conclusão, o uso de modelos de Deep Convolutional Neural Networks é uma abordagem promissora para a identificação de Pneumonia em imagens de radiografia de tórax, e pode ser uma ferramenta valiosa para auxiliar os profissionais de saúde no diagnóstico e tratamento precoce da doença.

Para o desenvolvimento deste trabalho quatro objectivos específicos foram propostos, o primeiro deles foi a construção do conjunto de dados que foi alcançado usando dados de sites colaborativos com destaque para o *Kaggle.com* e aplicando técnicas para multiplicação artificial do número de imagens para ter o conjunto de dados mais robusto. O segundo objectivo de implementar algoritmos de deep learning também foi alcançado usando o framework TensorFlow, a biblioteca de alto nível Keras e a linguagem de programação Python. O terceiro objectivo que consistia em realizar o treinamento foi alcançado usando a soluções grátis baseadas em Cloud como o Kaggle. O último objectivo de realizar avaliação de desempenho do modelo foi alcançado com ajuda da biblioteca scikit-learn. Os resultados aqui apresentados mostram que é possível fazer o uso de modelos de deep learning para a identificação de doenças pulmonares com destaque para a Pneumonia com base em imagens de radiografia de tórax.

5.1.Limitações da Pesquisa

Treinamento: na solução baseada em Cloud escolhida para o treinamento dos modelos que e grátis como o Kaggle encontrou-se algumas limitações tais como: O Kaggle possui uma limitação de tempo para execução do treinamento, o uso de aceleradores como GPU ou TPU

só podem ser usados por 30 horas por semana e uma sessão pode ficar aberta por 3 horas para TPU e 8 horas para o GPU. O Google Colab é uma ótima escolha para ultrapassar as limitações de tempo do Kaggle, mas esse serviço requer que a sessão criada fique activa até o fim do treinamento em caso de oscilação da conexão à Internet a sessão é interrompida e em alguns casos não pode ser retomada. O treinamento usando aceleradores do Google Colab é muito lento em relação ao uso dos aceleradores do Kaggle.

5.2.Sugestões

- Realizar Optimização dos modelos aqui implementados para alcançar uma alta precisão (acima dos 95 %).
- Realizar o estudo para implementação da técnica de transfer learning para melhorar a performance dos modelos.
- Expandir o conjunto de dados utilizado no estudo para incluir imagens de pacientes de diferentes faixas etárias, gêneros e etnias, a fim de avaliar a robustez dos modelos em diferentes cenários.
- Investigar a possibilidade de combinar informações clínicas e radiológicas, como dados de exames laboratoriais e informações sobre a sintomatologia do paciente, para melhorar a performance dos modelos de Deep Learning na identificação de Pneumonia.
- Explorar o uso de técnicas de interpretabilidade de modelos de Deep Learning, como o Grad-CAM, para visualizar as regiões das imagens que são mais relevantes para a tomada de decisão do modelo, auxiliando assim a compreender as bases para a classificação da doença.
- Investigar o desempenho do modelo em diferentes conjuntos de dados, com variações nas características das imagens de radiografia de tórax, como resolução, níveis de iluminação e qualidade da imagem.
- Avaliar a capacidade do modelo em identificar diferentes estágios da Pneumonia, como estágios iniciais ou avançados da doença, para entender melhor sua capacidade em diagnósticos precisos.

- Comparar o desempenho do modelo com outros métodos de diagnóstico, como a avaliação feita por profissionais de saúde, para entender melhor as vantagens e desvantagens de cada método.
- Investigar o uso de técnicas de transferência de aprendizagem para melhorar o desempenho do modelo, utilizando modelos pré-treinados em conjuntos de dados semelhantes.
- Implementar o modelo em um sistema de diagnóstico em tempo real, para avaliar sua eficácia em um ambiente clínico.
- Essas são algumas sugestões para trabalhos futuros que podem aprimorar a aplicação do modelo em questão e expandir o conhecimento na área de diagnóstico de Pneumonia através de imagens de radiografia de tórax.

6. Referências Bibliográficas

Abadi, M., et al. (2016). TensorFlow: A System for Large-Scale Machine Learning. In Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16) (pp. 265-283). USENIX Association.

Alpaydin, E. (2010). Introduction to machine learning (2nd ed.). Cambridge, MA: MIT Press.

Alpaydin, E. (2010). Introduction to machine learning (2nd ed.). Cambridge, MA: MIT Press.

American Thoracic Society. (2021). Pneumonia. Obtido em 11 de abril de 2023, de <https://www.thoracic.org/patients/patient-resources/resources/pneumonia.pdf>

American Thoracic Society; Infectious Diseases Society of America. (2019). Guidelines for the management of adults with community-acquired pneumonia. Clinical Infectious Diseases, 68(4), e1-e47. <https://doi.org/10.1093/cid/ciy812>

Bishop, C. M. (2006). Pattern recognition and machine learning (1st ed.). New York: Springer.

Chest X-Ray Images (Pneumonia). Kaggle. <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>. Acesso em 18 de abril de 2023.

Chollet, F. (2015). Keras: Deep learning library for Theano and TensorFlow. <https://github.com/fchollet/keras>

Chollet, F. (2018). Deep learning with Python. Manning Publications.

Costa, A. (24 de Março de 2022). *pneumonia*. Obtido de saudebemestar: <https://www.saudebemestar.pt/pt/medicina/pneumologia/pneumonia/>

Creswell, J. W. (2014). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Sage publications.

Electrodermal Activity Meter. (2006). Obtido em Janeiro de 2023, de

http://instruct1.cit.cornell.edu/courses/ee476/FinalProjects/s2006/hmm32_pjw32/index.html

Ferreira, A. M., et al. (2020). Pneumonia: revisão de literatura. Brazilian Journal of Health Review, 3(3), 8293-8303. <https://doi.org/10.34119/bjhrv3n3-298>

Fontes, A. M. G. Técnicas de coleta de dados em pesquisa qualitativa: entrevista. In: Conselho Federal de Enfermagem. O trabalho de enfermagem no contexto da saúde brasileira. Brasília: Cofen, 2016. p. 225-238.

Géron, A. (2019). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. O'Reilly Media, Inc.

Gonzalez, R.C., & Woods, R.E. (2018). Digital Image Processing. (4^a ed.). Pearson Education.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction (2nd ed.). New York: Springer.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction (2nd ed.). New York: Springer.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR) (pp. 770-778).

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

- Huang, G., Liu, Z., van der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR) (pp. 4700-4708).
- Jain, A.K. (1988). Fundamentals of Digital Image Processing. Prentice Hall.
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255-260.
- Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4, 237-285.
- Kaggle. (2010). Kaggle: Your Home for Data Science. Recuperado de <https://www.kaggle.com/>
- Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). ImageNet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems (pp. 1097-1105).
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
- Liang, S., Zhang, J., Zhang, L., Huang, H., & Ji, S. (2019). Detecting pneumonia in chest x-rays with hybrid feature representations. *IEEE Transactions on Medical Imaging*, 38(5), 1227-1236.
- Long, B., Koyfman, A., & Gottlieb, M. (2020). Evaluation and management of community-acquired pneumonia in the emergency department. *The Western Journal of Emergency Medicine*, 21(2), 342–354. <https://doi.org/10.5811/westjem.2019.11.44902>
- Mandell, L. A., Niederman, M. S., & Wunderink, R. G. (2019). Infectious Diseases Society of America/American Thoracic Society consensus guidelines on the management of community-acquired pneumonia in adults. *Clinical Infectious Diseases*, 71(6), e1–e27. <https://doi.org/10.1093/cid/ciz1057>
- Marieb, E. N., & Hoehn, K. (2018). Anatomia e fisiologia. São Paulo: Pearson.
- Mayo Clinic. (2021). Pneumonia. <https://www.mayoclinic.org/diseases->

[conditions/pneumonia/symptoms-causes/syc-20354204](https://doi.org/10.1007/978-94-011-3532-0_2)

Meier, R. (2012). Professional Android 4 application development. John Wiley & Sons.

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013). Playing Atari with deep reinforcement learning. arXiv preprint arXiv:1312.5602.

Moore, K. L., Dalley, A. F., & Agur, A. M. R. (2018). Anatomia orientada para a clínica. Rio de Janeiro: Guanabara Koogan.

Murphy, K. (2012). Machine learning: a probabilistic perspective. Chance Encounters: Probability in Education. London: The MIT Press. https://doi.org/10.1007/978-94-011-3532-0_2

Nunes, P. (30 de Outubro de 2015). *projecto*. Obtido de know: <https://knoow.net/cienceconempr/gestao/projecto/>

OpenCV. (2021). OpenCV: Open Source Computer Vision Library. Disponível em: <http://opencv.org/>. Acesso em: 24 de abril de 2023.

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., ... & Lerer, A. (2019). PyTorch: An imperative style, high-performance deep learning library. Advances in Neural Information Processing Systems, 32, 8024-8035.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12(Oct), 2825-2830.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. Journal of machine learning research, 12(Oct), 2825-2830.

Pereira, L. C. O., & da Silva, M. L. (2009). *Android para desenvolvedores*. Brasport.

Raghavendran, K., Mylotte, J. M., & Scannapieco, F. A. (2007). Nursing home-associated pneumonia, hospital-acquired pneumonia and ventilator-associated pneumonia: the contribution of dental biofilms and

- periodontalinflammation. *Periodontology* 2000, 44, 164.
- Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., ... & Ng, A. Y. (2017). CheXNet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225.
- Santos, M. T. (2 de Fevereiro de 2021). *pneumonia-o-que-e*. Obtido de saude.abril: <https://saude.abril.com.br/medicina/pneumonia-o-que-e/>
- Santos, S. E. (12 de Novembro de 2018). *pneumonia-tratamento-e-eficaz-quanto-mais-precoce-o-diagnostico*. Obtido de atlasdaude: <https://www.atlasdaude.pt/publico/content/pneumonia-tratamento-e-eficaz-quanto-mais-precoce-o-diagnostico>
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85-117.
- Sethi, S. (2020, December 4). *Considerações gerais sobre pneumonia*. Manual MSD Versão Saúde Para a Família; Manuais MSD. <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>
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
- Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction (2nd ed.). Cambridge, MA: MIT Press.
- Szegedy, C., Vanhoucke, V., Ioffe, S., & Shlens, J. (2016). Rethinking the Inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 2818-2826).
- The Belmont Report. (Julho de 2000). Obtido em Janeiro de 2008, de <http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.htm>
- Torres, A., et al. (2021). Diagnosis and treatment of community-acquired pneumonia. *The Lancet Infectious Diseases*, 21(7), e107-e120. <https://doi.org/10.1016/S1473->

Tortora, G. J., & Derrickson, B. (2017). *Princípios de anatomia e fisiologia*. Porto Alegre: Artmed.

Trabalho em equipe: o significado desse conceito e como ele realmente pode funcionar. (n.d.). Blog Da Samba. <https://sambatech.com/blog/insights/trabalho-em-equipe/>

Vagas, R. V.(2005). *Gerenciamento de Projetos: estabelecendo diferenciais Competitivos*; prefacio de Reeve Harold R. 6. Ed. Rio de Janeiro: Brasport.

Vargas, R. V.(2009). *Manual Prático de Projeto: utilizando o PMBOK Guide*. 4. Ed. Rio de Janeiro: Brasport.

Wang, D., Khosla, A., Gargeya, R., Irshad, H., & Beck, A. H. (2018). Tienet: Text-image embedding network for common thorax disease classification and reporting in chest x-rays. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 9049-9058).

Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. M. (2017). ChestX-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2097-2106).

World Health Organization. (2021, November 11). *Pneumonia*. Who.int; World Health Organization: WHO. <https://www.who.int/news-room/fact-sheets/detail/pneumonia>

Yin, R. K. (2015). *Case study research: Design and methods* (5th ed.). Thousand Oaks, CA: Sage Publications.