



UNIVERSIDADE FEDERAL DO CEARÁ
DEPARTAMENTO DE ENGENHARIA DA COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E DE
COMPUTAÇÃO
MESTRADO ACADÊMICO EM ENGENHARIA DA COMPUTAÇÃO

MICHELE FARIA DE OLIVEIRA

REDES NEURAS PROFUNDAS PARA CLASSIFICAR IMAGENS DE
RADIOGRAFIAS ODONTOLÓGICA PANORÂMICA NA PERSPECTIVA DE GRUPOS
ETÁRIOS: UM ESTUDO EXPLORATÓRIO

SOBRAL/CE

2025

MICHELE FARIA DE OLIVEIRA

REDES NEURAIIS PROFUNDAS PARA CLASSIFICAR IMAGENS DE RADIOGRAFIAS
ODONTOLÓGICA PANORÂMICA NA PERSPECTIVA DE GRUPOS ETÁRIOS: UM
ESTUDO EXPLORATÓRIO

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

Orientador: Prof. Dr. Iális Cavalcante de Paula Júnior

SOBRAL/CE

2025

MICHELE FARIA DE OLIVEIRA

REDES NEURAIS PROFUNDAS PARA CLASSIFICAR IMAGENS DE RADIOGRAFIAS
ODONTOLÓGICA PANORÂMICA NA PERSPECTIVA DE GRUPOS ETÁRIOS: UM
ESTUDO EXPLORATÓRIO

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

Aprovada em:

BANCA EXAMINADORA

Prof. Dr. Iális Cavalcante de Paula
Júnior (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Flávio Henrique Duarte de Araújo
Universidade Federal do Piauí (UFPI)

Prof. Dr. Carlos Alexandre Rolim Fernandes
Universidade Federal do Ceará (UFC)

À minha mãe, alicerce da minha vida, em memória das minhas avós, e ao meu esposo, com todo o amor que me deram e que hoje me sustenta, e com a gratidão que aprendi com vocês a cultivar.

AGRADECIMENTOS

Agradeço a Deus por todas as coisas que foram feitas por Ele, pois, sem Ele, nada do que foi feito se fez. Ele é, foi e sempre será a vida e a luz que guia os meus caminhos.

Ao meu esposo, Fischer Jônatas, meu maior amor, agradeço pela serenidade com que caminha ao meu lado e pelo apoio incondicional em cada passo. Obrigada por acreditar em mim mais do que eu mesma. Obrigada por polir a minha alma todos os dias. Celebro o tempo que nos trouxe até aqui e a imensidão que é você.

À minha família, pelo cuidado, pelas orações e por acreditarem, desde o início, que tudo daria certo. Um agradecimento especial à minha mãe, Vera Lúcia, e à minha tia Ângela, que, mesmo sem compreenderem exatamente o que eu estava fazendo, estiveram presentes em todos os momentos, oferecendo força e coragem para continuar. Vocês me dão coragem pra viver, porque sou amada.

Ao Prof. Dr. Iális Cavalcante de Paula Júnior, minha gratidão pela orientação generosa e por ser um verdadeiro facilitador da vida, daqueles que chegam antes, destrancam portas, varrem os caminhos e lustram o que estava opaco. Um servidor público digno de honra, cuja atuação engrandece a nossa universidade.

Ao Prof. Dr. Márcio Amora, que sempre foi um grande incentivador e apoiador desde os primeiros momentos.

Aos meus amigos de vida, que tornam a jornada mais leve e alegre. Vocês habitam o lado esquerdo do meu peito e sabem quem são. Aos amigos que fiz no mestrado: Francisco Gean, Lucas Aguiar e João Artur. Minha sincera gratidão pelo apoio. Talvez vocês nem imaginem o quanto foram essenciais nesta caminhada. Em meio aos desafios, vocês me ofereceram presença, ensinamentos e generosidade.

Agradeço à Universidade Federal do Ceará, campus Sobral, pela formação e pela convivência com professores e técnicos administrativos competentes e humanos. Levo cada um de vocês comigo, para sempre.

Agradeço também à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo incentivo à pesquisa e ao desenvolvimento científico em nosso país.

Se o futuro é ancestral, então o melhor presente é lançar a pedra com coragem e perceber que aquilo que vislumbramos adiante já nos pertence, ou talvez sempre tenha nos pertencido. Seja como for, abraço, louvo e agradeço pelo nosso encontro ou reencontro, nesta vida. Obrigada e Obrigada!

Está escrito no grande livro da sabedoria popular: Que primeiro se deve viver, que é pra depois poetar. Se arme de amor e coragem, que a minh'arma eu num embainho. Se arme de amor e coragem, como a rosa se arma no espinho. Na tragédia, poesia; na prosa, tem picada feroz e carinho. Se arme de amor e coragem, que a minh'arma eu num embainho.

Ednardo, 1977.

RESUMO

A estimativa da idade cronológica é um processo essencial na odontologia forense, pois permite inferir a faixa etária de um indivíduo com base em características biométricas, especialmente as observadas em radiografias da arcada dentária. Embora métodos tradicionais da odontologia, como os de Nolla e Demirjian, sejam amplamente utilizados, eles frequentemente demandam avaliação especializada e podem apresentar limitações em termos de precisão e escalabilidade. Nesse contexto, as redes neurais profundas surgem como uma alternativa promissora, oferecendo uma abordagem não invasiva, automatizada e de baixo custo. Esta dissertação está estruturada em três partes principais. A primeira consiste em uma revisão sistemática da literatura, conduzida com o objetivo de identificar o estado da arte sobre o uso de aprendizado profundo na estimativa de idade a partir de radiografias panorâmicas odontológicas. Após um rigoroso processo de seleção, 17 estudos primários foram incluídos e analisados. Os principais achados apontam que as arquiteturas *ResNet* e *AlexNet* destacam-se entre as abordagens avaliadas, especialmente em faixas etárias mais jovens (3 a 14 anos), e que as métricas de acurácia e recall são as mais utilizadas para validação dos modelos. Na segunda etapa do trabalho, foi criado um banco de imagens composto por 1.545 radiografias panorâmicas, obtidas por meio de parcerias com clínicas odontológicas. As imagens foram anonimizadas e padronizadas, com aplicação de uma máscara para destacar a região anatômica de interesse. O banco foi estruturado para possibilitar a replicação e ampliação em estudos futuros. Por fim, a terceira etapa envolveu um estudo exploratório com sete arquiteturas de redes neurais profundas (*AlexNet*, *VGG16*, *ResNet50*, *DenseNet*, *MobileNetV2*, *EfficientNet-B3* e *InceptionV3*), com o objetivo de avaliar sua capacidade de estimar a idade a partir das imagens do banco construído. Os resultados obtidos reforçam os achados da literatura e indicam o potencial da abordagem em cenários reais. Esta pesquisa contribui com subsídios para a evolução de soluções inteligentes aplicadas à odontologia forense, podendo apoiar pesquisadores, desenvolvedores de ferramentas e profissionais da área da saúde.

Palavras-chave: Redes Neurais Profundas. Odontologia Forense. Banco de Imagens. Radiografias Panorâmicas.

ABSTRACT

Estimating chronological age is an essential process in forensic dentistry, as it allows inferring an individual's age group based on biometric characteristics, particularly those observable in dental radiographs. Although traditional methods, such as those proposed by Nolla and Demirjian, are widely adopted, they often require expert interpretation and may present limitations regarding accuracy and scalability. Deep neural networks have emerged as a promising alternative in this context, offering a non-invasive, automated, and low-cost approach. This dissertation comprises three main parts. The first presents a systematic literature review aimed at identifying the state of the art in using deep learning for age estimation based on panoramic dental radiographs. Following a rigorous selection process, 17 primary studies were included and analyzed. Key findings indicate that the ResNet and AlexNet architectures stand out among the evaluated approaches, particularly for younger age groups (3 to 14 years), and that accuracy and recall are the most frequently used performance metrics. The second stage involved the construction of a dataset composed of 1,545 panoramic radiographic images obtained through partnerships with dental clinics. All images were anonymized and standardized, and a mask was applied to highlight the anatomical region of interest. This dataset aims to facilitate replication and future expansion in similar studies. Finally, the third part of this work consisted of an exploratory study using seven convolutional neural network architectures (AlexNet, VGG16, ResNet50, DenseNet, MobileNetV2, EfficientNet-B3, and InceptionV3), aiming to evaluate their effectiveness in estimating age based on the constructed dataset. The results support previous findings from the literature and demonstrate the potential of deep learning approaches in real-world forensic scenarios. This research contributes valuable insights for developing intelligent solutions in forensic dentistry, supporting researchers, tool developers, and healthcare professionals.

Keywords: Deep Neural Networks. Forensic Odontology. Image Database. Panoramic Radiographs.

LISTA DE FIGURAS

Figura 1 – Planejamento do estudo	19
Figura 2 – Resultados do estudo	20
Figura 3 – Notação dentária da dentição permanente - Federação Dentária Internacional (FDI)	24
Figura 4 – Rede neural profunda	26
Figura 5 – Exemplo de arquitetura CNN para classificação de imagens. Adaptado de Vargas <i>et al.</i> (2016)	27
Figura 6 – Processo de busca e seleção dos trabalhos	48
Figura 7 – Distribuição de trabalhos por Qualis e países	50
Figura 8 – Técnicas de processamento	51
Figura 9 – Tipos de data augmentation	52
Figura 10 – Divisão dos dados e quantidade de imagens	53
Figura 11 – Regressão/classificação, <i>Transfer learning</i> e Base de dados	54
Figura 12 – Redes identificadas	54
Figura 13 – Comportamento das redes neurais	56
Figura 14 – Distribuição dos valores das métricas Acurácia e MAE	58
Figura 15 – Seleção de imagens para compor o banco de imagens de radiografias panorâmicas	62
Figura 16 – Aplicação de máscara para seleção de interesse da imagem	64
Figura 17 – Distribuição de imagens por idade e sexo	67
Figura 18 – Quantidade de imagens por faixa etária	68
Figura 19 – Agrupamento das imagens por faixa etária e sexo	69
Figura 20 – Distribuição de imagens	70
Figura 21 – Porcentagem do dentes parte superior da arcada dentária	73
Figura 22 – Porcentagem do dentes parte inferior da arcada dentária	74
Figura 23 – Distribuição de imagens por quantidade de dentes nas arcadas dentárias	75
Figura 24 – Matriz de confusão para predição do sexo no conjunto de teste	80
Figura 25 – Configuração do estudo	86
Figura 26 – Especificidade por faixa etária	96
Figura 27 – Matriz de confusão <i>Resnet</i>	97
Figura 28 – Matriz de confusão <i>EfficientNet</i>	98

Figura 29 – Matriz de confusão <i>MobileNet</i> e <i>InceptionV3</i>	99
Figura 30 – Matriz de confusão <i>VGG16</i> , <i>Dense201</i> e <i>AlexNet</i>	99
Figura 31 – Matriz de confusão <i>EfficientNetB5</i> , <i>InceptionV3</i> e <i>ResNet50</i>	100

LISTA DE TABELAS

Tabela 1 – Matriz de confusão multiclasse (referência: Classe A)	36
Tabela 2 – Comparativo entre banco de de imagens na área de Inteligência Artificial . .	42
Tabela 3 – Palavras-chaves selecionadas	45
Tabela 4 – Itens analisados nos trabalhos selecionados	46
Tabela 5 – Trabalhos selecionados no estudo	49
Tabela 6 – Identificação da imagem	71
Tabela 7 – Identificação dos dentes na imagem	71
Tabela 8 – Métricas de desempenho para classificação do sexo a partir de radiografias panorâmicas	80
Tabela 9 – Distribuição das imagens por faixa etária	87
Tabela 10 – Parâmetros de configuração utilizados no treinamento das redes neurais . . .	89
Tabela 11 – Técnicas para aumento de dados	90
Tabela 12 – Comparação de desempenho entre diferentes arquiteturas de redes neurais .	93
Tabela 13 – Resultados de desempenho dos modelos avaliados na avaliação externa . . .	100

LISTA DE ABREVIATURAS E SIGLAS

CLAHE	Equalização de Histograma Adaptativa com Limitação de Contraste
CNN	Redes Neurais Convolucionais
FDI	Federação Dentária Internacional
GPU	Unidade de Processamento Gráfico
JPEG	<i>Joint Photographic Experts Group</i>
LFW	<i>Faces in the Wild</i>
MAE	<i>Mean Absolute Error</i>
MLP	<i>Multilayer Perceptron</i>
MSE	<i>Mean Squared Error</i>
OMS	Organização Mundial da Saúde
ReLU	<i>Rectified Linear Unit</i>
RNP	Redes Neurais Profundas
ROI	Região de Interesse
RSL	Revisão Sistemática da Literatura
Tanh	Tangente Hiperbólica

SUMÁRIO

1	INTRODUÇÃO	15
1.1	Problema e Motivação	16
1.2	Objetivo	17
1.3	Construção da Proposta	18
1.4	Resultados e Contribuições	19
1.5	Resumos dos Capítulos da Dissertação	20
2	FUNDAMENTAÇÃO TEÓRICA E TRABALHOS RELACIONADOS	22
2.1	Técnicas Tradicionais de Identificação de Idade por meio da Arcada Dentária	22
2.2	Nomenclatura Dental pela Federação Dentária Internacional	24
2.3	Redes Neurais Profundas	25
2.4	Redes Neurais Convolucionais (CNN)	26
2.5	Arquiteturas de Redes Neurais Profundas	28
2.5.1	<i>EfficientNet</i>	28
2.5.2	<i>Resnet</i>	29
2.5.3	<i>AlexNet</i>	30
2.5.4	<i>VGG</i>	30
2.5.5	<i>MobileNet</i>	31
2.5.6	<i>DenseNet</i>	32
2.5.7	<i>Inception</i>	32
2.6	Aspectos Fundamentais da Análise com Redes Neurais Profundas	33
2.6.1	<i>Visão Geral sobre a Estratégia de Aumento de Dados</i>	33
2.6.2	<i>Validação com Holdout</i>	34
2.6.3	<i>Pré-processamento de Imagens Médicas</i>	35
2.6.4	<i>Visão Geral sobre as Métricas de Avaliação Essenciais</i>	35
2.7	Trabalhos Relacionados	37
2.7.1	<i>Estimativa Automatizada da Idade Dental Usando Aprendizado Profundo</i>	37
2.7.2	<i>Estado do Desenvolvimento Dental Usando Aprendizado Profundo</i>	38
2.7.3	<i>Comparação de Métodos Manuais e Automatizados de Estimativa da Idade Dental</i>	38

2.7.4	<i>Criação de Banco de Imagens</i>	39
3	REVISÃO SISTEMÁTICA DA LITERATURA	43
3.1	Planejamento	43
3.2	Execução	47
3.3	Análise	48
3.3.1	<i>Quais são as Características dos Trabalhos Feitos para Avaliar as Redes Neurais Profundas (RNP)?</i>	48
3.3.2	<i>Quais são as Principais RNP Utilizadas?</i>	54
3.3.3	<i>Quais são as Principais Características das RNP?</i>	55
3.4	Ameaças a Validade	58
3.5	Considerações Finais	59
4	BANCO DE IMAGENS	61
4.1	Aquisição das Imagens	62
4.2	Preparação das Imagens	63
4.3	Qualificação das Imagens	65
4.4	Banco de Imagens	65
4.4.1	<i>Identificação dos Dentes na Arcada Dentária</i>	70
4.5	Exemplo de Uso do Banco de Imagens	75
4.5.1	<i>Organização dos Dados e Pré-Processamento</i>	76
4.5.2	<i>Arquitetura do Modelo de Rede Neural Utilizada</i>	76
4.5.3	<i>Pseudo Código Utilizado no Experimento</i>	76
4.5.4	<i>Resultados e Análise</i>	79
4.5.5	<i>Código e Reprodutibilidade</i>	81
4.6	Ameaças a Validade do Conjunto de Imagens	81
4.7	Consideração Finais	82
5	ESTUDO EXPLORATÓRIO	84
5.1	Configuração do Estudo	84
5.1.1	<i>Base de Dados</i>	86
5.1.2	<i>Ambiente de Desenvolvimento</i>	87
5.1.3	<i>Classificação</i>	88
5.1.4	<i>Validação Cruzada</i>	88
5.1.5	<i>Pré-Processamento dos Dados</i>	89

5.1.6	<i>Aumento de Dados</i>	90
5.1.7	<i>Transfer Learning</i>	90
5.1.8	<i>Arquitetura de Redes Neurais Profundas</i>	91
5.1.9	<i>Validação Externa</i>	91
5.1.10	<i>Métricas de Avaliação</i>	91
5.2	Resultados e Discussão	92
5.2.1	<i>Utilização das Métricas Essenciais no Contexto de Estimativa de Idade</i> . .	92
5.2.2	<i>Resultados Encontrados para as Métricas Essenciais</i>	93
5.2.3	<i>Resultados Encontrados para Especificidade</i>	95
5.2.4	<i>Resultados Encontrados para Matriz de Confusão</i>	96
5.2.5	<i>Resultados para Validação Externa</i>	100
5.3	Implicações do Estudo	101
5.3.1	<i>Implicações para Pesquisadores</i>	101
5.3.2	<i>Implicações para Desenvolvedores de Ferramentas</i>	102
5.3.3	<i>Implicações para o Ensino e Formação Profissional</i>	103
5.4	Ameaças a Validade Relacionadas ao Estudo Exploratório	103
5.5	Considerações Finais do Estudo Exploratório	104
6	CONCLUSÃO E TRABALHOS FUTUROS	106
6.1	Resumo dos Principais Resultados	106
6.2	Principais Contribuições	107
6.3	Trabalhos Futuros	108
	REFERÊNCIAS	110

1 INTRODUÇÃO

A estimativa de idade é o processo de determinar a faixa etária de uma pessoa com base em características biométricas (MUALLA *et al.*, 2020). Hoje em dia, a pesquisa na área tem ganhado muito destaque, devido as necessidades impostas a ciência forense (KIM *et al.*, 2021; MUALLA *et al.*, 2020; ATAŞ *et al.*, 2022a; SATHYAVATHI, 2024). Infelizmente, devido a questões adversas tais como: acidentes aéreos, explosões, e catástrofes onde se faz necessários a identificação de pessoas, a estimativa de idade pode ser uma aliada importante para ajudar na identificação das pessoas falecidas (MUALLA *et al.*, 2020; GUO *et al.*, 2021; WANG *et al.*, 2022; YEOM *et al.*, 2023). Além das demandas de atividades forense, a classificação da faixa etária por meio da arcada dentária pode ajudar em outras aplicações, tais como: serviços de passaporte (SATHYAVATHI, 2024), identificação de maioridade (MOHAMED *et al.*, 2023), planejamento ortodôntico, arqueologia (GALIBOURG *et al.*, 2021) e odontologia pediátrica (DONG *et al.*, 2023).

Os métodos tradicionais de identificação da idade por meio da arcada dentária são amplamente utilizados na odontologia forense e incluem diversas técnicas baseadas na observação e análise do desenvolvimento dentário (MOHAMED *et al.*, 2023). Essas abordagens geralmente envolvem a avaliação de características morfológicas, como a erupção dentária, a formação das raízes e o grau de mineralização dos dentes, sendo aplicadas tanto em indivíduos vivos quanto em restos mortais (GALIBOURG *et al.*, 2021; SHI *et al.*, 2024). Um dos métodos mais populares e consolidados na literatura é o de Demirjian (SIVRI *et al.*, 2024), que utiliza critérios específicos para classificar o estágio de desenvolvimento de sete dentes do lado esquerdo da mandíbula, oferecendo uma pontuação que pode ser convertida em idade estimada (WANG *et al.*, 2022). Este método é amplamente reconhecido e aplicado em várias populações ao redor do mundo devido à sua simplicidade, reprodutibilidade e relativa precisão. No entanto, estudos recentes têm apontado limitações na sua generalização, especialmente quando aplicado a grupos populacionais distintos daqueles para os quais foi originalmente desenvolvido, o que pode comprometer sua acurácia (GUO *et al.*, 2021; WANG *et al.*, 2022; YEOM *et al.*, 2023; SATHYAVATHI, 2024; SIVRI *et al.*, 2024).

Diante disso, torna-se cada vez mais relevante a busca por formas menos custosas, mais objetivas e automatizadas para apoiar a estimativa de idade por meio de radiografias dentárias, como aquelas baseadas em inteligência artificial e redes neurais profundas (WANG *et al.*, 2022; ONG *et al.*, 2024). Tais abordagens oferecem potencial para reduzir o viés humano,

melhorar a precisão dos resultados e ampliar a aplicabilidade dos métodos em diferentes contextos forenses e populacionais.

Oportunamente, na literatura relacionada existem relatos de abordagens de aprendizado profundo utilizadas efetivamente para resolver muitos problemas em várias áreas de pesquisa, incluindo visão computacional (ATAŞ *et al.*, 2022b; ATAŞ, 2023; OZDEMIR, 2023), processo de linguagem natural (ATHANASIADOU *et al.*, 2018) e reconhecimento de objetos (COLLOBERT *et al.*, 2011; ISLAM *et al.*, 2024). Além disto, o uso de redes neurais profundas para identificação de idade por meio da arcada dentária também é discutido em vários trabalhos (MUALLA *et al.*, 2020; SIVRI *et al.*, 2024; MU; GANG, 2022), e ainda os trabalhos recentes têm mostrado resultados promissores do uso de rede neurais com essa finalidade (ATAŞ *et al.*, 2022a; MOHAMMAD *et al.*, 2024; OLIVEIRA *et al.*, 2024; SIVRI *et al.*, 2024). Assim, esta tem sido uma área de pesquisa ativa e constante, mas que ainda apresenta oportunidades de estudos (KIM *et al.*, 2023; MOHAMED *et al.*, 2023; SIVRI *et al.*, 2024).

Conhecer as redes neurais utilizadas no problema é fundamental para garantir a eficácia do modelo preditivo, especialmente em aplicações sensíveis como a estimativa de idade em contextos legais (MUALLA *et al.*, 2020; YEOM *et al.*, 2023; SIVRI *et al.*, 2024). Cada arquitetura apresenta características distintas em termos de profundidade, capacidade de generalização e tempo de processamento (MU; GANG, 2022). A escolha inadequada pode resultar em modelos com alto índice de falsos positivos, comprometendo a confiabilidade dos resultados. Para que o uso dessas redes seja efetivo, é indispensável que apresentem um bom nível de acerto, especialmente em diferentes faixas etárias e conjuntos de dados (WU *et al.*, 2022). Assim, compreender o comportamento de cada rede em relação ao problema proposto permite selecionar aquelas com melhor desempenho, garantindo maior precisão nas estimativas e reduzindo o risco de decisões equivocadas em contextos clínicos, forenses ou judiciais (WANG *et al.*, 2023; SIVRI *et al.*, 2024).

1.1 Problema e Motivação

Observa-se uma carência de estudos sistemáticos e replicáveis que consolidem as evidências disponíveis sobre o uso de redes neurais profundas na estimativa de idade por meio de radiografias panorâmicas odontológicas (KIM *et al.*, 2021; DONG *et al.*, 2023; SIVRI *et al.*, 2024; ONG *et al.*, 2024; OLIVEIRA *et al.*, 2024). Embora existam diversos estudos com diferentes arquiteturas de redes e propostas metodológicas, ainda é difícil identificar, de forma clara, quais

abordagens são mais eficazes, quais faixas etárias são mais bem classificadas, quais métricas são utilizadas e quais fatores interferem nos resultados, pelo fato da baixa replicabilidades destes estudos. Essa fragmentação do conhecimento e baixa replicabilidade dos estudos dificulta a consolidação dos resultados e a real eficácias do uso das redes neurais profundas para estimativa de idades (SIVRI *et al.*, 2024).

Além do mais, os estudos presentes na literatura, majoritariamente utilizam banco de imagens privados, não disponíveis, tais como (KIM *et al.*, 2021; DONG *et al.*, 2023; SIVRI *et al.*, 2024; ONG *et al.*, 2024). Essa restrição impede a replicação destes trabalhos e uma comparação das redes com as mesmas base de dados comuns. Logo a deficiências de comparação das redes disponíveis na literatura dificulta a ampla aplicação prática em contextos clínicos e forenses. Além, de prejudicar a comparação de novas redes emergentes com aquelas já propostas na literatura. Sem uma base de dados comum é um desafio para identificar a real eficácias das redes neurais profundas para estimativa de idade por meio de radiografias odontológicas panorâmicas.

Diante disso, este trabalho justifica-se pela necessidade de reunir, analisar criticamente e sintetizar os estudos existentes, de forma a oferecer uma visão geral sobre o estado da arte na área. Além disso, compreender as lacunas presentes na literatura poderá orientar futuras pesquisas e contribuir para implementações de novas redes neurais profundas com mais eficácia para a estimativa de idades por meio de radiografias. A motivação central, deste estudo é propor e executar um trabalho que possa que trazer uma visão geral sobre redes neurais profundas para estimativa de idades, e também que possa apoiar futuras pesquisas. Desta forma, beneficiar toda comunidade científica, os profissionais da área odontológica e forense, além de fomentar o uso da inteligência artificial como ferramenta complementar na estimativa de idade, contribuindo para a inovação e aprimoramento das práticas existentes.

1.2 Objetivo

Esta dissertação tem como principal objetivo *fornecer uma visão detalhada do uso de RNP para estimativa de idades por meio de imagens de radiografias odontológicas panorâmicas*. Além disto, este trabalho destina-se a criação de um banco de imagens público de radiografias panorâmicas para apoiar esta pesquisa. Este banco de imagens também poderá servir de comparação para futuros trabalhos na área, reduzindo o viés de cada estudo ao utilizar um banco de imagens particular, que não permite a comparação mais justas entre diferentes abordagens relatadas na literatura. Para atingir os objetivos gerais desta dissertação, os seguintes

objetivos específicos (OE) são definidos:

- *OE1* Investigar as estratégias de redes neurais profundas apresentadas em estudos anteriores, por meio de uma revisão sistemática da literatura.
- *OE2* Selecionar um grupo das melhores redes neurais profundas utilizadas para estimativa de idades.
- *OE3* Criar um repositório para disponibilização dos dados encontrados sobre as redes neurais identificadas.
- *OE4* Criar um banco de imagens com radiografias odontológicas panorâmicas.
- *OE5* Estabelecer um protocolo para padronização das expansões futuras do banco de imagens.
- *OE6* Criar um repositório para disponibilização do banco de imagens.
- *OE7* Realizar um estudo exploratório sobre as principais redes neurais profundas para estimativa de idades.
- *OE8* Criar um repositório para disponibilização do pacote de replicação do estudo, para apoiar pesquisas futuras.

1.3 Construção da Proposta

Considerando a falta de estudos que investigam a eficácia das redes neurais profundas para estimativa de idades em radiografias panorâmicas em um banco de imagens público, este trabalho se destina a preencher esta lacuna ao apresentar uma revisão sistemática da literatura sobre o uso de redes neurais profundas, um banco de imagens, disponível para toda comunidade, com a definição de seus protocolos para expansões futuras e um estudo exploratório sobre redes neurais profundas com uma abordagem para estimativa de idades com uso destas redes. Este estudo pode beneficiar toda comunidade acadêmica de pesquisadores de redes neurais profundas bem como a comunidade de odontologia forense e pesquisadores de saúde bucal.

A Figura 1 apresenta os passos do estudo, começando do Passo 1 onde foi conduzida uma revisão sistemática da literatura para encontrar estudos relevantes que mencionam o uso de redes neurais profundas para estimativa de idade. Este passo foi importante e necessário para identificação do estado da arte em relação ao assunto. Foram encontradas oportunidades de pesquisa e lacunas que motivou o restante desta pesquisa. Uma delas foi a falta de um banco de imagens público onde pesquisadores poderiam comparar suas abordagens. Um banco público é necessário para permitir comparações com menos vieses, por se tratar uma base única e

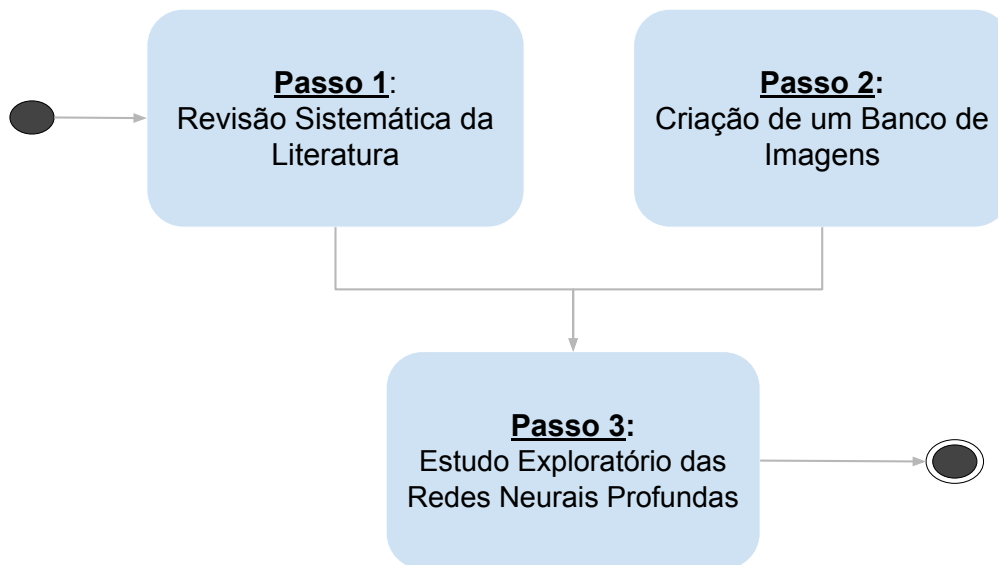


Figura 1 – Planejamento do estudo

padronizada. Desta forma, o Passo 2 é destinado a demonstração dos protocolos para criação do banco de imagens bem como apresentação das principais características do banco de imagens. Finalmente, o Passo 3, é constituído de um estudo exploratório com nove redes neurais profundas, identificadas no Passo 1 e que utiliza o banco de imagens provido no Passo 2.

1.4 Resultados e Contribuições

A Figura 2 ilustra as principais contribuições deste trabalho. Como pode ser observado, este trabalho apresenta resultados nos três passos do estudo. No Passo 1 foram levantados por meio de uma revisão sistemática da literatura trabalhos que demonstram o uso de redes neurais profundas para estimativa de idades em imagens em radiografias odontológicas panorâmicas. Como exemplo, neste passo foram listadas as principais características das redes neurais profundas no contexto da estimativa de idades.

No Passo 2 são apresentados os protocolos utilizados para criação de um banco de imagens, composto de 1.545 imagens. Este banco será disponibilizado para toda comunidade acadêmica. Ainda, é demonstrado um exemplo de uso do banco de imagens com redes neurais profundas.

Finalmente, no Passo 3, como principal resultado, é demonstrado uma abordagem para estimativa de imagens por meio de redes neurais profundas. Além disto, o estudo exploratório conduzido neste passo apresenta uma avaliação das eficácias das redes neurais profundas para estimativa de idade. Todos os passos deste estudo fornece um repositório específico para apoiar

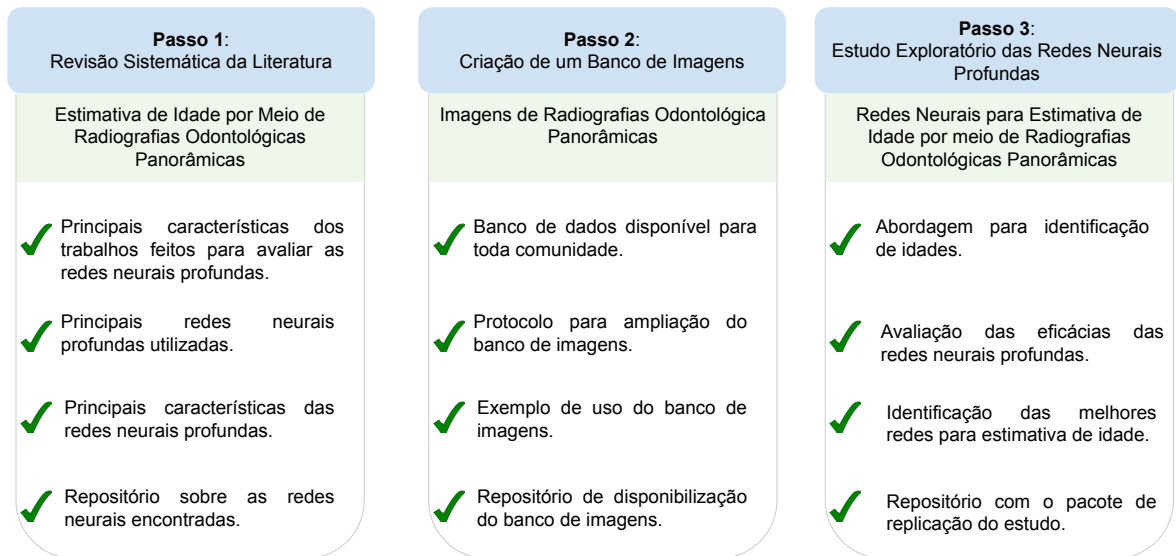


Figura 2 – Resultados do estudo

replicações futuras e novas pesquisa na área.

1.5 Resumos dos Capítulos da Dissertação

Capítulo 2: O Capítulo 2 fornece os conceitos essenciais para dar suporte a esta dissertação. Além de detalhes sobre as redes neurais profundas utilizadas neste estudo. O capítulo apresenta: técnicas tradicionais de identificação de idade por meio da arcada dentária, nomenclatura dental pela Federação Dentária Internacional, arquiteturas das redes neurais profundas utilizadas e aspectos fundamentais da análise com redes neurais profundas. Este capítulo também discute trabalhos relacionados.

Capítulo 3: Para o Capítulo 3 o principal objetivo foi identificar os estudos primários relatados na literatura sobre o uso de redes neurais profundas para classificar imagens de radiografias odontológica panorâmica na perspectiva de grupos etários. Para esse propósito, foi conduzida uma Revisão Sistemática da Literatura (RSL) em seis bases de dados e foram encontrados 85 estudos, destes, foram selecionados 17 trabalhos para leitura completa. Ainda, foram apontadas as principais características dessas redes bem como uma indicação da eficácia de seu uso considerando diferentes faixas etárias. Estes resultados podem apoiar pesquisadores, construtores de ferramentas baseadas em redes neurais profundas, além de profissionais da ciência forense.

Capítulo 4: O Capítulo 4 apresenta o banco de imagens criado e utilizado neste estudo. O principal objetivo deste capítulo foi criar um banco de imagens que possa contribuir

com esta e futuras pesquisa na área. Ainda, foi demonstrado todo o processo de construção do banco de imagens para permitir que novas imagens possam ser incorporadas no futuro. O banco de imagens é composto de 1.545 imagens de radiografias panorâmicas extraídas de três clínicas diferentes em duas unidades da Federação.

Capítulo 5: O Capítulo 5 apresenta um estudo exploratório com nove redes neurais profundas. Este estudo foi projetado e conduzido para avaliar o desempenhos das principais redes neurais profundas encontradas no Capítulo 3 e utilizar o banco de imagens construído no Capítulo 4. As nove arquiteturas utilizadas no estudo foram avaliadas por meio das métricas: acurácia, *F1-score*, precisão, *Recall*, especificidade e tempo de treino. Além disso, foi feita uma etapa de validação externa para as três arquiteturas que apresentaram o melhor desempenho.

Capítulo 6: O Capítulo 6 apresenta a conclusão desta dissertação, revisando os resultados referentes aos objetivos específicos. Ainda, são discutidas as contribuições e implicações deste trabalho. Também são discutidos alguns possíveis trabalhos futuros como consequência deste estudo.

2 FUNDAMENTAÇÃO TEÓRICA E TRABALHOS RELACIONADOS

Esta seção apresenta a fundamentação teórica para apoiar o presente trabalho e uma discussão sobre alguns trabalhos relacionados. Assim, a seção está organizada da seguinte forma: a Seção 2.1 apresenta as técnicas tradicionais de identificação de idade por meio da arcada dentária. A Seção 2.2 apresenta a nomenclatura dental pela Federação Dentária Internacional utilizada neste trabalho. As Seções 2.3 e 2.4 são discutidos alguns pontos importantes sobre redes neurais profundas e redes neurais convolucionais, respectivamente. Ainda, as Seções 2.5 e 2.6 apresentam uma visão geral sobre as arquiteturas das redes neurais profundas e aspectos fundamentais da análise com redes neurais profundas. Finalmente, alguns trabalhos relacionados a esta dissertação são discutidos na Seção 2.7.

2.1 Técnicas Tradicionais de Identificação de Idade por meio da Arcada Dentária

A estimativa da idade cronológica a partir de estruturas dentárias tem sido amplamente empregada na odontologia legal, sobretudo por meio da análise de radiografias panorâmicas. A dentição é considerada uma das estruturas biológicas mais resistentes do corpo humano, frequentemente preservando-se mesmo em condições extremas, o que a possibilita a identificação de indivíduos de forma menos custosa e eficaz (WANG *et al.*, 2022; ATAŞ *et al.*, 2022a; MOHAMMAD *et al.*, 2024).

Os métodos tradicionais de estimativa de idade baseiam-se, em sua maioria, na avaliação dos estágios de desenvolvimento e mineralização dos dentes permanentes. Esses métodos comparam a imagem radiográfica dos dentes com estágios de referência previamente definidos em tabelas ou escalas específicas (ODZHAKOV; APOSTOLOV, 2019). No Brasil, um dos métodos mais conhecidos é o proposto Nicodemo *et al.* (1974), que adaptou os estágios de mineralização definidos por Nolla *et al.* (1952) às características da população brasileira.

O método de Nicodemo utiliza oito estágios de desenvolvimento dentário, que vão desde o início da formação da coroa até o fechamento completo do ápice radicular. A aplicação envolve a identificação do estágio de mineralização de cada dente na imagem radiográfica, e a posterior consulta a uma tabela que fornece intervalos etários estimados para cada estágio. A média dessas estimativas compõe o intervalo de idade atribuído ao indivíduo. Embora seja amplamente utilizado no Brasil, estudos demonstram que sua acurácia é maior em faixas etárias mais jovens. Gobbo *et al.* (2021), por exemplo, encontrou melhores resultados entre 7 e 11 anos

de idade, enquanto Moreno *et al.* (2014) observaram desempenho superior na faixa de 10 a 15 anos ao considerar apenas os segundos e terceiros molares.

Outros métodos internacionalmente reconhecidos também são empregados, como o de Moorrees *et al.* (1963), que define múltiplos estágios de formação para cada dente, e o método de Demirjian *et al.* (1973), que se baseia em um sistema de pontuação relacionado ao desenvolvimento de sete dentes mandibulares esquerdos. Este último tem sido amplamente utilizado em diferentes contextos e adaptado para populações específicas, como demonstrado no estudo de Willems *et al.* (2001), que revisou a técnica de Demirjian para crianças belgas, corrigindo a tendência de superestimação da idade observada com a aplicação direta do método original.

Além disso, para indivíduos adultos, destaca-se o método de Kvaal *et al.* (1995), que utiliza medidas radiográficas da câmara pulpar como indicativo da idade, considerando a redução de seu tamanho ao longo do tempo em decorrência da deposição de dentina secundária. Esse método se mostra útil principalmente quando os processos de formação dentária já foram concluídos, o que limita a aplicabilidade das técnicas baseadas em estágios de desenvolvimento.

É importante ressaltar que a precisão de qualquer um desses métodos está sujeita à variabilidade populacional. Características étnicas, geográficas e socioeconômicas influenciam tanto o desenvolvimento dentário quanto a maturação óssea (BULGARELI *et al.*, 2018; YEOM *et al.*, 2023), e, por isso, métodos desenvolvidos em uma população nem sempre oferecem resultados confiáveis em outras (MUALLA *et al.*, 2020). Adicionalmente, a faixa etária analisada é um fator determinante, sendo comum uma maior precisão nas estimativas de idade em crianças e adolescentes, quando os dentes ainda estão em processo de desenvolvimento (SATHYAVATHI, 2024).

Por essas razões, é recomendável que a estimativa de idade dentária, sobretudo em contextos legais, seja utilizada de forma complementar a outros métodos de avaliação (GUO *et al.*, 2021; WANG *et al.*, 2022). Ainda, a combinação de dados radiográficos, análises ósseas, documentação clínica e histórico pessoal tende a oferecer uma estimativa mais confiável e robusta da idade cronológica de um indivíduo. Desta forma, o uso de diferentes técnicas para estimativa de idade pode aprimorar a acurácia dos resultados, minimizar vieses decorrentes de variações individuais ou populacionais e fornecer uma base mais sólida para decisões em contextos jurídicos, forenses ou clínicos (MUALLA *et al.*, 2020; GUO *et al.*, 2021; KIM *et al.*, 2021).

2.2 Nomenclatura Dental pela Federação Dentária Internacional

A padronização da nomenclatura dental é essencial para a comunicação clara entre profissionais de odontologia, bem como para o registro, intercâmbio e análise de informações clínicas, radiográficas e forenses (BEAINI *et al.*, 2016; FORTES *et al.*, 2020). Um dos sistemas de identificação dental mais amplamente adotados no mundo é o sistema de nomenclatura proposto pela FDI, também conhecido como Sistema de Dois Dígitos da FDI.

Nesse sistema, cada dente da dentição humana é representado por um código numérico de dois dígitos. O primeiro dígito indica o quadrante e o tipo de dentição (permanente ou decídua), enquanto o segundo dígito corresponde à posição do dente dentro do quadrante, contada a partir da linha média (incisivos centrais) em direção aos molares (SARE, 1997).

Para a dentição permanente, os quadrantes são numerados de 1 a 4, no sentido horário, a partir do quadrante superior direito:

- 1: Superior direito
- 2: Superior esquerdo
- 3: Inferior esquerdo
- 4: Inferior direito

A Figura 3 apresenta a identificação dos dentes segundo a nomenclatura FDI.

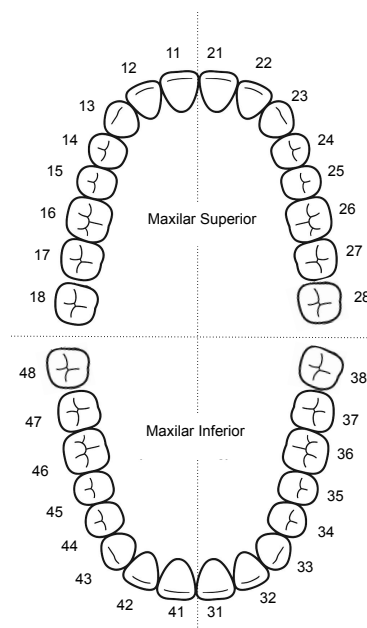


Figura 3 – Notação dentária da dentição permanente - FDI

Essa nomenclatura é útil em contextos clínicos e científicos, pois permite uma

identificação rápida e precisa dos dentes em exames radiográficos e prontuários odontológicos. Além disso, o sistema FDI é compatível com *softwares* odontológicos e protocolos internacionais, sendo recomendado pela Organização Mundial da Saúde Organização Mundial da Saúde (OMS).

Em estudos de estimativa de idade a partir de radiografias panorâmicas, como os descritos neste trabalho, o uso de uma nomenclatura padronizada é fundamental para garantir a rastreabilidade dos dentes analisados, a reprodutibilidade dos métodos aplicados e a comparabilidade dos resultados com outras pesquisas nacionais e internacionais (BEAINI *et al.*, 2016).

2.3 Redes Neurais Profundas

As RNP (*Deep Neural Networks*) representam uma evolução dentro da área de Inteligência Artificial. Elas são caracterizadas pela presença de múltiplas camadas ocultas entre a entrada e a saída. Essa estrutura permite a modelagem de relações não lineares nos dados. Sua popularização, sobretudo na última década, está diretamente relacionada ao avanço das capacidades computacionais e à disponibilidade de grandes conjuntos de dados rotulados, além do desenvolvimento de técnicas como normalização por lotes (*batch normalization*), funções de ativação não lineares mais eficientes, e algoritmos de otimização mais robustos (LECUN *et al.*, 2015; SCHMIDHUBER, 2015).

A Figura 4 apresenta um diagrama esquemático de uma rede neural profunda, composta por múltiplas camadas de processamento interligadas. À esquerda, temos a camada de entrada, responsável por receber os dados brutos (por exemplo, uma imagem, vetor de características ou sinais biomédicos). Cada nó dessa camada representa uma variável ou pixel da entrada. Em seguida, observa-se uma série de camadas ocultas, também chamadas de camadas intermediárias. Cada neurônio artificial dessas camadas realiza uma combinação ponderada das entradas recebidas, seguida da aplicação de uma função de ativação não linear. Por fim, a informação é propagada até a camada de saída, que produz o resultado final da rede, como uma classificação (ex: “masculino” ou “feminino”) ou uma regressão (ex: estimativa de idade cronológica).

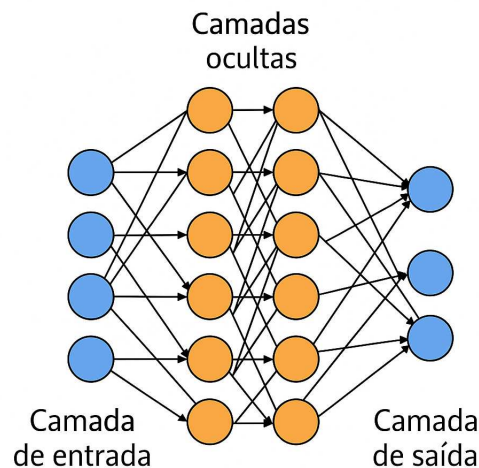


Figura 4 – Rede neural profunda

A principal distinção entre uma rede neural rasa (*Shallow Neural Network*) como *Perceptron* simples (*Single-Layer Perceptron*) e as redes neurais profundas reside na quantidade de camadas ocultas presentes na arquitetura (GOODFELLOW *et al.*, 2016). As Múltiplas camadas ocultas, permitem a aprendizagem de representações hierárquicas e cada vez mais abstratas dos dados. Essa característica torna as redes profundas adequadas para tarefas como reconhecimento de padrões em imagens médicas, sinais biométricos e linguagem natural, onde a complexidade dos dados exige maior capacidade de modelagem (OLVERES *et al.*, 2021; HAYKIN, 2001; SZE *et al.*, 2017).

É importante destacar que as *Redes Neurais Convolucionais (CNN)* são uma sub-classe das redes profundas, desenvolvidas para lidar com dados estruturados espacialmente, como imagens. Redes Neurais profundas baseadas em camadas densas, como o *Multilayer Perceptron (MLP)*, possuem capacidade limitada para lidar com dados complexos, sendo mais indicadas para dados tabulares ou sequenciais (GOODFELLOW *et al.*, 2016).

2.4 Redes Neurais Convolucionais (CNN)

Dentro das RNP, destaca-se uma arquitetura especialmente adequada à análise de imagens: *Convolutional Neural Networks (CNN)*. A CNN foi proposta por LeCun *et al.* (1998), e são inspiradas no processamento do olho humano. Dessa forma, elas foram projetadas para processar dados com estrutura em grade, como imagens bidimensionais. A Figura 5 apresenta um diagrama esquemático de uma CNN compostas por três tipos de camadas com funções diferentes,

são elas: *camadas convolucionais, camadas de pooling e camadas totalmente conectadas.*

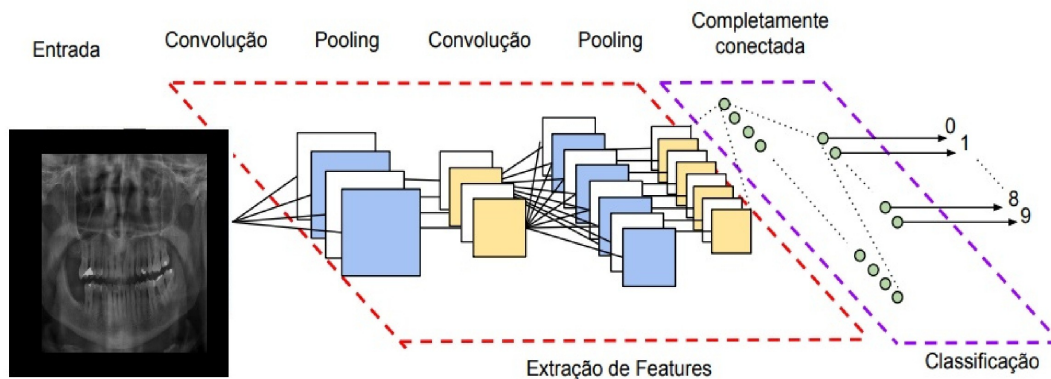


Figura 5 – Exemplo de arquitetura CNN para classificação de imagens. Adaptado de Vargas *et al.* (2016)

As camadas convolucionais são responsáveis pela extração automática de características locais da imagem, como contornos, texturas e padrões espaciais. Ela é feita por meio da aplicação de filtros ou *kernels* que percorrem a imagem de entrada, realiza as operações de convolução e gera os mapas de ativação. Os mapas de ativação destacam as regiões onde os padrões (borda horizontais ou verticais, textura e etc.) ocorrem. Cada filtro aprende, ao longo do treinamento, a detectar um tipo específico de característica visual. Após a convolução, é comum aplicar uma função de ativação não linear, como a *Rectified Linear Unit* (ReLU) ou Tangente Hiperbólica (Tanh), que introduz não linearidade no modelo. Dessa forma, a rede constrói uma representação hierárquica dos dados à medida que novas camadas convolucionais são adicionadas (GOODFELLOW *et al.*, 2016).

As camadas de *pooling*, também chamadas de camadas de subamostragem, têm como principal função reduzir a dimensionalidade dos mapas de ativação, mantendo as informações mais relevantes (KIM *et al.*, 2021). Na maioria dos casos, essa redução é realizada por meio da operação de *max pooling*, que seleciona o valor máximo dentro de uma região definida para aplicar a um tensor de saída. Essa camada também contribui para a eficiência computacional, uma vez que, ao reduzir o número de parâmetros a serem aprendidos pela rede, aumenta-se a agilidade do treinamento. Além disso, ela torna o modelo mais robusto, pois introduz uma invariância espacial, o que reforça sua capacidade de reconhecer padrões mesmo eles estão levemente deslocados. (UKWANDU *et al.*, 2022; GOODFELLOW *et al.*, 2016)

Por fim, as camadas totalmente conectadas (ou densas) realizam a etapa final de processamento, na qual todas as ativações anteriores são achatadas em um vetor unidimensional

e conectadas a todos os neurônios da camada seguinte. É nessa fase que a rede realiza a combinação global das informações extraídas nas etapas anteriores, permitindo gerar a predição final. Essa camada associa os padrões detectados nas camadas convolucionais às classes-alvo em tarefas de classificação ou a valores contínuos em problemas de regressão. Assim como nas camadas convolucionais, as camadas densas também utilizam funções de ativação, sendo a *Softmax* ou sigmoide frequentemente utilizadas na camada de saída, a depender do tipo de tarefa, classificação multiclasse, binária ou regressão (VAZ; BALAJI, 2021; GOODFELLOW *et al.*, 2016).

Portanto, a CNN é composta por camadas especializadas na extração e interpretação de padrões visuais e têm se mostrado particularmente eficazes em aplicações médicas como, classificação de doenças, e segmentação de estruturas anatômicas (UKWANDU *et al.*, 2022; MIENYE *et al.*, 2025). A sua capacidade de identificar automaticamente padrões sutis e complexos é particularmente valiosa neste contexto onde variações visuais podem indicar alterações clínicas significativas (SARVAMANGALA; KULKARNI, 2022).

2.5 Arquiteturas de Redes Neurais Profundas

Existe uma ampla variedade de arquiteturas de redes neurais convolucionais, desenvolvidas para responder a diferentes demandas, como profundidade da rede, número de categorias a serem classificadas e volume de dados disponíveis. Essas variações estruturais possibilitam que os modelos sejam ajustados de forma mais precisa às características específicas de cada aplicação, o que resulta em um aprendizado mais eficiente e adaptado ao contexto da tarefa.

2.5.1 *EfficientNet*

A *EfficientNet* é uma família de arquiteturas de redes neurais convolucionais proposta por Tan e Le (2019). A principal inovação apresentada pelos autores está no método de escalonamento composto (*compound scaling*), que permite expandir simultaneamente três dimensões da rede de maneira equilibrada: a profundidade (número de camadas), a largura (quantidade de canais por camada) e a resolução da imagem de entrada. Esse escalonamento é adotado com o objetivo de que, à medida que se aumenta a resolução da imagem de entrada, a arquitetura da rede também possa ser expandida de forma proporcional. Isso envolve a adição

de mais camadas, para ampliar o campo receptivo, e mais canais, para permitir a detecção de padrões visuais mais refinados (TAN; LE, 2019).

A versão base da *EfficientNet*, foi obtida por meio de uma técnica de otimização automatizada de arquiteturas, o *Neural Architecture Search* (NAS). A partir desse modelo, variantes como EfficientNet-B1 até B7 foram derivadas por meio aplicação progressiva do escalonamento composto. Cada versão apresenta um número maior de camadas, canais e resolução, conforme os coeficientes de escala estabelecidos.

A literatura tem explorado o uso da *EfficientNet* em diversas aplicações de saúde, incluindo classificação de lesões cutâneas, identificação de doenças oftalmológicas e, mais recentemente, em análises odontológicas, como a estimativa de idade a partir de radiografias panorâmicas (ATAŞ *et al.*, 2022a; GUO *et al.*, 2021; SATHYAVATHI, 2024). Além disso, estudos comparativos demonstram que a *EfficientNet* apresenta desempenho superior ou equivalente a arquiteturas mais complexas, como *ResNet152*, *InceptionV2* ou *ResNet*, com menor número de parâmetros e tempo de inferência reduzido, tornando-a uma escolha recorrente em contextos que exigem acurácia com eficiência (YADAV; JADHAV, 2019; TAN; LE, 2019; MU; GANG, 2022; ATAŞ *et al.*, 2022a).

2.5.2 Resnet

A arquitetura *ResNet* (*Residual Network*) foi proposta por He *et al.* (2016) como uma solução ao problema de degradação de desempenho em redes neurais profundas. Esse problema se refere à observação de que, à medida que se aumenta a profundidade de uma rede, ela pode apresentar acurácia inferior mesmo em dados de treinamento, devido a dificuldades de otimização. Dessa forma, a principal inovação das *ResNets* foi a introdução dos *blocos residuais*, que incluem conexões de atalho (*skip connections*) entre a entrada e a saída de um bloco.

As conexões de atalho, permitem que os gradientes fluam de forma mais eficiente durante o processo de retropropagação. Em vez de aprender diretamente uma função mapeadora $H(x)$, o bloco residual aprende a função $F(x) = H(x) - x$, ou seja, a diferença residual entre a entrada e a saída. Além disso, a rede utiliza uma camada chamada *Global Average Pooling* (GAP) para reduzir a dimensionalidade dos dados antes da camada final de classificação (HE *et al.*, 2016).

Entre as variantes mais conhecidas, destacam-se a *ResNet18*, *ResNet34*, *ResNet50*, *ResNet101* e *ResNet152*, diferenciando-se principalmente pelo número de camadas e blocos

residuais. A *ResNet50*, em particular, tem sido amplamente empregada em tarefas de classificação de imagens médicas. Pesquisas como as de Yadav e Jadhav (2019) e (MUALLA *et al.*, 2020) demonstraram o potencial da *ResNet50* na estimativa da idade dentária a partir de radiografias panorâmicas.

2.5.3 *AlexNet*

A *AlexNet* foi uma das primeiras arquiteturas de redes neurais convolucionais profundas a alcançar resultados expressivos em tarefas de classificação de imagens. Proposta por Krizhevsky *et al.* (2012), a rede conquistou o primeiro lugar na competição *ImageNet Large Scale Visual Recognition Challenge (ILSVRC)* de 2012, com uma redução de erro de mais de 10% em relação ao segundo colocado. Esse resultado marcou o início da era moderna do aprendizado profundo em visão computacional.

A arquitetura *AlexNet* introduziu diversos avanços práticos que se tornaram referência para modelos posteriores, como o uso de unidades de ativação não lineares do tipo ReLU, aplicação de técnicas de regularização como *dropout*, e treinamento distribuído em múltiplas GPUs. Embora seja considerada uma arquitetura clássica e tenha sido superada em desempenho por redes mais recentes, como *VGG*, *ResNet* e *EfficientNet*, a *AlexNet* continua sendo utilizada no processamento de imagens médicas (YADAV; JADHAV, 2019; TANG *et al.*, 2023).

2.5.4 *VGG*

A arquitetura *VGG* foi proposta por Simonyan e Zisserman (2014) no *Visual Geometry Group (VGG)* da Universidade de Oxford e se destacou na competição *ImageNet Large Scale Visual Recognition Challenge (ILSVRC)* de 2014. Embora não tenha vencido a competição, a *VGG* atraiu ampla atenção na comunidade científica por sua simplicidade de arquitetura e desempenho robusto. A principal inovação da *VGG* está no uso sistemático de camadas convolucionais com pequenos filtros de tamanho 3×3 , empilhadas de forma profunda para capturar padrões complexos de forma hierárquica. Ao contrário de modelos anteriores, como a *AlexNet*, que utilizavam filtros maiores (11×11 , 5×5), a *VGG* mostrou que o empilhamento de múltiplos filtros pequenos pode alcançar a mesma recepção de campo com menos parâmetros e maior capacidade de generalização.

As versões mais utilizadas da *VGG* são a *VGG16* e a *VGG19*, que possuem, respectivamente, 16 e 19 camadas. Ambas utilizam as mesmas ideias básicas, diferenciando-se

na quantidade de camadas convolucionais intermediárias. A rede *VGG* ainda é amplamente utilizada como arquitetura de base para *transfer learning*, principalmente em aplicações que envolvem conjuntos de dados limitados, como imagens médicas. Diversos estudos empregaram a *VGG16* como extrator de características para tarefas como detecção de tumores, classificação de tecidos e reconhecimento de padrões em radiografias (YADAV; JADHAV, 2019; HUANG *et al.*, 2022).

2.5.5 *MobileNet*

A arquitetura *MobileNet* foi desenvolvida pelo *Google* com o objetivo de oferecer modelos leves e eficientes, voltados para dispositivos móveis e embarcados com recursos computacionais limitados. A primeira versão da *MobileNet* foi apresentada por Howard *et al.* (2017), e evoluções posteriores, como *MobileNetV2* e *MobileNetV3* trouxeram melhorias significativas em desempenho e capacidade de generalização, mantendo baixo custo computacional.

O diferencial da *MobileNet* está no uso de convoluções separáveis em profundidade (*depthwise separable convolutions*), que reduzem drasticamente o número de parâmetros e operações sem comprometer de forma significativa a acurácia. Essa técnica consiste em dividir a operação de convolução tradicional em duas etapas: uma convolução *depthwise*, que aplica um filtro por canal de entrada, e uma convolução *pointwise* (1×1), que combina os resultados de forma eficiente. A versão *MobileNetV2*, proposta Sandler *et al.* (2018) introduziu dois diferenciais:

- **Blocos com conexões residuais:** Consistem em atalhos que ligam diretamente a entrada de um bloco à sua saída, sem passar por todas as transformações intermediárias. Isso permite que a rede aprenda uma função residual, ou seja, apenas a diferença entre a entrada e a saída esperada. Esse mecanismo é semelhante aos da *ResNet*, facilita a propagação de gradientes durante o treinamento e melhora a estabilidade em redes profundas.
- **Camadas com expansões lineares e estreitamento:** A estratégia amplia temporariamente a dimensionalidade para enriquecer a representação de características e, em seguida, reduz a dimensão para manter a leveza do modelo. Esse mecanismo melhora a capacidade de generalização sem comprometer a eficiência.

2.5.6 *DenseNet*

A *DenseNet* (Densely Connected Convolutional Network), proposta por Huang *et al.* (2017), surgiu como uma evolução das redes residuais, com o intuito de aprimorar a propagação de gradientes e promover o reaproveitamento eficiente de características entre as camadas. Diferentemente das arquiteturas tradicionais, em que cada camada transmite sua saída apenas para a seguinte, a *DenseNet* estabelece conexões diretas e cumulativas entre todas as camadas de um mesmo bloco. Assim, cada camada recebe como entrada a concatenação das saídas de todas as camadas anteriores, o que favorece a reutilização das representações extraídas ao longo da rede.

Essa abordagem densa reduz a redundância, melhora a eficiência do treinamento e contribui para uma representação mais robusta. Isso ocorre porque a rede evita recalcular mapas de características já aprendidos, o que resulta em modelos mais compactos, sem prejuízo de desempenho (HUANG *et al.*, 2017).

Para organizar essa estrutura de conexões, a *DenseNet* é dividida em blocos modulares, conhecidos como *dense blocks*, intercalados por camadas de transição. Esses blocos agrupam múltiplas camadas convolucionais com conexões cumulativas, enquanto as camadas de transição atuam na redução da dimensionalidade e no controle da complexidade da rede, utilizando operações de convolução e *pooling*. Essa organização facilita a escalabilidade da arquitetura e contribui para seu bom desempenho, mesmo em redes muito profundas.

No que diz respeito às variantes da arquitetura, destacam-se a *DenseNet121*, *DenseNet169* e *DenseNet201*, que se diferenciam principalmente pelo número de camadas. A *DenseNet121* é considerada uma versão mais enxuta, ideal para cenários com limitações de memória ou processamento. Já as versões mais profundas, como a *DenseNet169* e a *DenseNet201*, são capazes de capturar representações mais complexas e detalhadas dos dados, sendo recomendadas para aplicações que demandam alta capacidade de generalização e maior volume de informações.

2.5.7 *Inception*

A família de redes *Inception* surgiu como uma proposta para o aumento da profundidade e largura das redes neurais convolucionais, sem elevar excessivamente o custo computacional. A primeira versão, conhecida como *GoogLeNet* ou *InceptionV1*, foi apresentada

por Szegedy *et al.* (2015) no contexto do *ImageNet Large Scale Visual Recognition Challenge* (ILSVRC) de 2014. A principal inovação dessa arquitetura foi o uso do módulo *Inception*, que realiza, em paralelo, convoluções de diferentes tamanhos (1×1 , 3×3 , 5×5) e operações de *pooling*, permitindo que a rede aprenda representações em múltiplas escalas e reduzindo o número de parâmetros por meio de convoluções 1×1 como pré-processamento (SZEGEDY *et al.*, 2015).

Novas versões da arquitetura foram desenvolvidas, como a *InceptionV2* e a *InceptionV3*, que incorporaram técnicas mais sofisticadas para melhorar a eficiência e o desempenho da rede. A *InceptionV3*, proposta por Szegedy *et al.* (2016), introduziu a fatoração de convoluções maiores em operações menores (por exemplo, transformar uma convolução 5×5 em duas 3×3), além de usar convoluções assimétricas (como 1×7 e 7×1) e camadas auxiliares para regularização e auxílio no treinamento. Essas estratégias permitiram o aumento da profundidade da rede com controle do custo computacional, mantendo elevado desempenho em tarefas de classificação (SZEGEDY *et al.*, 2016). A arquitetura altamente modular e eficiente da rede *Inception* tornou-a amplamente utilizada em uma variedade de aplicações, incluindo reconhecimento de objetos e análise de imagens médicas (PAN *et al.*, 2023; AKTER *et al.*, 2022; MUJAHID *et al.*, 2022).

2.6 Aspectos Fundamentais da Análise com Redes Neurais Profundas

Esta seção está organizada da seguinte forma: Visão Geral sobre a Estratégia de Aumento de Dados (Seção 2.6.1), Validação com Holdout (Seção 2.6.2), Pré-processamento de Imagens Médicas (Seção 2.6.3) e Visão Geral sobre as Métricas de Avaliação Essenciais (Seção 2.6.4).

2.6.1 Visão Geral sobre a Estratégia de Aumento de Dados

O aumento de dados (*data augmentation*) é uma estratégia fundamental para contornar a escassez de imagens rotuladas e consequentemente melhorar a capacidade de generalização dos modelos de aprendizado profundo. Essa técnica consiste na criação de amostras artificiais a partir de transformações aplicadas sobre imagens reais, permitindo expandir o conjunto de treinamento sem a necessidade de novas coletas ou rotulações manualmente custosas (KEBAILI *et al.*, 2023).

Essas transformações podem ser classificadas em geométricas e fotométricas. As transformações geométricas alteram a estrutura espacial da imagem por meio do reposiciona-

mento dos pixels, como ocorre em operações de rotação, espelhamento, translação e redimensionamento. Já as transformações fotométricas modificam atributos visuais da imagem, como brilho, contraste, saturação e ruído, sem alterar sua estrutura geométrica. No caso específico de radiografias panorâmicas odontológicas, o uso de técnicas de aumento de dados é particularmente relevante, pois a qualidade visual das imagens pode variar significativamente conforme o modelo do equipamento de raios-X, as condições de instalação do aparelho (como iluminação, blindagem e calibração) e o posicionamento do paciente durante a realização do exame. Essas alterações podem afetar diretamente o brilho, o contraste e a orientação da imagem e representam ruídos comuns nos dados de entrada. Dessa forma, o uso das transformações sintéticas permite simular essas variações, contribuindo para a robustez do modelo frente a diferentes condições de aquisição (MUMUNI; MUMUNI, 2022; FRID-ADAR *et al.*, 2018; XU *et al.*, 2023).

2.6.2 Validação com Holdout

A validação cruzada é um método de amostragem utilizado para avaliar o desempenho preditivo de modelos estatísticos e de aprendizado de máquina. Consiste em particionar o conjunto total de dados em subconjuntos de treinamento, validação e teste. O objetivo central desse método é garantir que os resultados obtidos durante a fase de treinamento não sejam consequência de sobreajuste (*overfitting*) (WILIMITIS; WALSH, 2023).

Entre as técnicas de validação cruzada mais comuns estão o método o *K-Fold* e o *Holdout*. A técnica *K-Fold* subdivide os dados em *K* subconjuntos, onde cada subconjunto é utilizado alternadamente como validação, enquanto os demais são empregados para treinamento, proporcionando uma estimativa do desempenho médio do modelo. Já validação cruzada do tipo *Holdout* consiste em particionar o conjunto total de dados em subconjuntos independentes de treinamento, validação e teste, utilizando o treinamento para construir o modelo, a validação para ajustes intermediários e otimização de hiperparâmetros, e o teste para avaliação final e imparcial do desempenho preditivo. (BISHOP; NASRABADI, 2006; KOHAVI *et al.*, 1995).

Neste estudo específico, optou-se por uma abordagem utilizando a técnica *Holdout*. As etapas de treinamento e validação foram realizadas por meio de particionamento estratificado. Esse método mantém a distribuição original das faixas etárias presente no banco de dados, preservando a representatividade dos dados e garantindo consistência nas avaliações intermediárias do modelo. Por fim, a avaliação em um conjunto de teste independente, não utilizado nas etapas anteriores de ajuste de hiperparâmetros, serviu para estimar com precisão o desempenho real do

modelo em dados inéditos, simulando, assim, condições próximas às aplicações reais.

2.6.3 Pré-processamento de Imagens Médicas

O pré-processamento de imagens é uma etapa fundamental, especialmente em contextos médicos, como imagens de radiografias panorâmicas. Seu principal objetivo é aprimorar a qualidade visual e eliminar artefatos ou ruídos que possam comprometer a interpretação ou o desempenho de modelos computacionais (SATHYAVATHI, 2024). Ao transformar a imagem brutas em representações mais limpas e consistentes, o pré-processamento prepara as imagens para classificações mais precisas, reduzindo variabilidades técnicas e realçando as informações relevantes (SHI *et al.*, 2024).

Dentre as técnicas empregadas nesse processo, o redimensionamento é uma das mais básicas. Ele assegura que todas as imagens possuam dimensões compatíveis com os requisitos das redes neurais, o que é crucial para o treinamento de modelos. Isso é feito por meio do mapeamento proporcional dos pixels da imagem original para uma nova resolução, geralmente utilizando interpolação para preservar a integridade visual (TU *et al.*, 2023).

A conversão para escala de cinza simplifica os dados visuais ao representar a intensidade luminosa de cada pixel em tons que variam do preto ao branco. Esse processo reduz a complexidade da imagem, mantendo as informações estruturais essenciais, o que favorece tanto a análise humana quanto computacional. Em geral, essa representação utiliza 8 bits, permitindo 256 níveis distintos de cinza (TASYUREK; GUL, 2023).

Outra técnica amplamente adotada Equalização de Histograma Adaptativa com Limitação de Contraste (CLAHE). O CLAHE ajusta o contraste de pequenas regiões da imagem individualmente e de forma controlada, evitando o excesso de realce em áreas homogêneas. Isso permite uma melhoria local no contraste sem intensificar ruídos ou gerar distorções, especialmente nas bordas e cantos da imagem (SARVAMANGALA; KULKARNI, 2022).

A seleção criteriosa das etapas de pré-processamento impacta diretamente na capacidade do modelo de extrair representações discriminativas, sendo um fator decisivo no sucesso de aplicações com dados clínicos.

2.6.4 Visão Geral sobre as Métricas de Avaliação Essenciais

A matriz de confusão é uma importante ferramenta para avaliação do desempenho dos modelos deste trabalho. Consiste em uma tabela cruzada que compara os rótulos verdadeiros

com os previstos pelo classificador, permitindo uma visualização clara das classificações corretas e incorretas (SATHYANARAYANAN; TANTRI, 2024). Seu objetivo é facilitar a análise da distribuição dos acertos e erros do modelo. É partir dela, são obtidos valores fundamentais para o cálculo das métricas de desempenho como: Verdadeiro Positivo (VP), Verdadeiro Negativo (VN), Falso Positivo (FP) e Falso Negativo (FN), conforme apresentado na Tabela 1.

- **Verdadeiro Positivo (VP):** Exemplo corretamente identificado como pertencente a uma determinada classe;
- **Verdadeiro Negativo (VN):** Exemplo corretamente identificado como não pertencente à classe analisada;
- **Falso Positivo (FP):** Exemplo incorretamente classificado como pertencente à classe analisada;
- **Falso Negativo (FN):** Exemplo incorretamente classificado como não pertencente à classe analisada.

Tabela 1 – Matriz de confusão multiclasse (referência: Classe A)

	Classe Predita A	Classe Predita B	Classe Predita C
Classe Verdadeira A	VP	FN	FN
Classe Verdadeira B	FP	VN	VN
Classe Verdadeira C	FP	VN	VN

A partir da matriz de confusão, são derivadas as seguintes métricas de desempenho:

- **Acurácia:** Mede o desempenho geral do modelo, calculando a proporção entre previsões corretas e o total de observações:

$$\text{Acurácia} = \frac{VP + VN}{VP + VN + FP + FN} \quad (2.1)$$

- **Precisão:** Avalia a proporção das observações positivas corretamente classificadas em relação ao total classificado como positivo:

$$\text{Precisão} = \frac{VP}{VP + FP} \quad (2.2)$$

- **Recall (Sensibilidade):** Mede a capacidade do modelo em identificar corretamente os casos positivos:

$$\text{Recall} = \frac{VP}{VP + FN} \quad (2.3)$$

- **F1-Score:** Representa a média harmônica entre precisão e recall:

$$\text{F1-Score} = \frac{2 \times \text{Precisão} \times \text{Recall}}{\text{Precisão} + \text{Recall}} \quad (2.4)$$

- **Especificidade:** Avalia a capacidade do modelo em identificar corretamente casos negativos:

$$\text{Especificidade} = \frac{VN}{VN + FP} \quad (2.5)$$

2.7 Trabalhos Relacionados

Esta seção está organizada da seguinte forma: Estimativa Automatizada da Idade Dental Usando Aprendizado Profundo (Seção 2.7.1), Estado do Desenvolvimento Dental Usando Aprendizado Profundo (Seção 2.7.2), Comparação de Métodos Manuais e Automatizados de Estimativa da Idade Dental (Seção 2.7.3) e Criação de Banco de Imagens (Seção 2.7.4).

2.7.1 Estimativa Automatizada da Idade Dental Usando Aprendizado Profundo

O estudo de Mualla *et al.* (2020) apresenta uma abordagem automatizada para a estimativa da idade dental a partir de imagens de raio-X. A proposta é estruturada em três etapas principais: pré-processamento da imagem, extração de características e classificação. Na fase de extração, foram utilizadas duas redes convolucionais profundas pré-treinadas, *AlexNet* e *ResNet-101*. Os classificadores aplicados incluíram algoritmos como árvore de decisão, k-vizinhos mais próximos (K-NN), discriminante linear e máquina de vetores de suporte. O uso do aprendizado por transferência teve como objetivo reduzir a necessidade de treinamento do zero. Os resultados indicaram melhor desempenho com as características extraídas pela *AlexNet*, sendo o K-NN o classificador com melhores métricas de avaliação.

Complementarmente, o trabalho de Ataş *et al.* (2022a) propõe uma rede neural modificada com base no modelo *InceptionV3* para a estimativa da idade dental forense, utilizando imagens de radiografias panorâmicas. O estudo avaliou diferentes modelos de aprendizado por transferência, como *EfficientNetB4*, *ResNet50V2*, *DenseNet201* e *MobileNetV2*, tendo o modelo *InceptionV3* apresentado os melhores resultados. Os mapas de calor gerados sugerem que o modelo concentra-se em estruturas relevantes para a tarefa, como dentes, tecido gengival e maxila. O erro absoluto médio (MAE) reportado foi de 3,13 anos.

No contexto brasileiro, o estudo de Oliveira *et al.* (2024) explora a estimativa de idade com aprendizado profundo a partir de um conjunto de dados nacional. Utilizando uma adaptação da arquitetura *InceptionV4* e aplicando técnicas de aumento de dados, o modelo

alcançou um MAE de validação de $3,83 \pm 0,224$ anos. O estudo também discute desafios como o desbalanceamento das classes etárias e a influência da faixa etária sobre o desempenho do modelo, ressaltando a importância de se considerar características populacionais específicas em modelos de predição de idade.

Por fim, Wang *et al.* (2023) avalia o desempenho de duas redes convolucionais: *VGG16* e *ResNet101* em uma população do leste da China. A *VGG16* apresentou resultados superiores na maioria das faixas etárias, especialmente nas idades mais jovens. No entanto, a diferença de desempenho entre os modelos diminuiu em faixas etárias mais avançadas.

2.7.2 Estado do Desenvolvimento Dental Usando Aprendizado Profundo

O trabalho de Ong *et al.* (2024) propõe uma abordagem totalmente automatizada para a classificação do estágio de desenvolvimento dental segundo o método de *Demirjian*, utilizando radiografias panorâmicas. O sistema é composto por três etapas: detecção com *YOLOv5*, segmentação com *U-Net* e classificação com *EfficientNet*. O desempenho do sistema foi avaliado etapa por etapa, com bons resultados em termos de precisão e segmentação, especialmente para molares.

De forma semelhante, Shi *et al.* (2024) desenvolve um *framework* em três etapas para estimar a idade dental em crianças. Inicialmente, é utilizada a *YOLOv3* para a localização e numeração dos dentes. Em seguida, o estágio de desenvolvimento dentário é estimado por meio da rede *SOS-Net*, baseada em uma adaptação do método de *SOS-Net*. Por fim, uma meta-análise de grupo único é aplicada para a estimativa da idade com base em dados estatísticos do conjunto de referência.

2.7.3 Comparação de Métodos Manuais e Automatizados de Estimativa da Idade Dental

O estudo de Guo *et al.* (2021) compara o desempenho de um método manual baseado em critérios legais com um modelo automatizado de aprendizado profundo para estimativa da idade com base em radiografias panorâmicas. Foram avaliados três limiares etários (14, 16 e 18 anos), utilizando modelos convolucionais e regressão logística manual. Os resultados sugerem que os modelos automáticos alcançaram desempenhos ligeiramente superiores em todos os limiares avaliados.

Na mesma linha, Sivri *et al.* (2024) realiza uma análise comparativa entre o método de *Demirjian* e modelos de aprendizado profundo com diferentes arquiteturas pré-treinadas. O

estudo destaca limitações dos métodos manuais, como o tempo de execução e a subjetividade, e propõe o uso de modelos como *DenseNet201* e *InceptionV3* como alternativas viáveis. Embora não afirme superioridade absoluta, os resultados indicam uma boa correspondência.

2.7.4 Criação de Banco de Imagens

O banco de imagens *CheXpert* (IRVIN *et al.*, 2019) apresenta um extenso conjunto de dados com 224.316 radiografias torácicas de 65.240 pacientes, anotadas para 14 condições clínicas com base em relatórios radiológicos. Os rótulos foram extraídos por meio de um anotador automático baseado em processamento de linguagem natural, que também é capaz de capturar as incertezas presentes nos laudos. O trabalho investiga diferentes estratégias para lidar com esses rótulos incertos no treinamento de redes neurais convolucionais e avalia o desempenho de um modelo resultante na detecção de cinco patologias. Nos experimentos realizados, o modelo apresentou desempenho comparável ou superior ao de radiologistas em três condições específicas: Cardiomegalia, Edema e Derrame Pleural. O *CheXpert* é disponibilizado como um *benchmark* público, amplamente utilizado para o desenvolvimento e a avaliação de modelos de interpretação automatizada de radiografias torácicas.

O banco de imagens *DeepFashion* (LIU *et al.*, 2016) apresenta um conjunto de dados em larga escala com mais de 800.000 imagens de roupas, cada uma anotada com 50 categorias de vestuário em nível fino e mais de 1.000 atributos visuais. Além disso, muitas das imagens incluem pontos de referência anatômicos (*landmarks*) para análise de estrutura. No mesmo trabalho, os autores propõem a *FashionNet*, uma rede neural projetada para aprender de forma conjunta a localização dos pontos de referência, a categoria e os atributos do vestuário. Os resultados experimentais indicam que essa abordagem de aprendizagem conjunta oferece melhor desempenho em tarefas como classificação de categorias e previsão de atributos, quando comparada a métodos tradicionais. O *DeepFashion* foi desenvolvido com o objetivo de estimular pesquisas em reconhecimento e recuperação visual de roupas, especialmente em contextos de busca por similaridade e recomendação de moda.

O trabalho sobre o banco de imagens *ImageNet* (DENG *et al.*, 2009) descreve a criação de uma ampla base de dados de imagens organizada segundo a hierarquia semântica do *WordNet*, contendo milhões de imagens anotadas manualmente e distribuídas em milhares de categorias visuais. A construção do banco de dados envolveu a coleta de imagens da internet, seguidas por um processo de validação colaborativa usando o *Amazon Mechanical Turk*, no qual

os participantes confirmavam a correspondência entre as imagens e os conceitos definidos. O *ImageNet* tornou-se um recurso central para o desenvolvimento do reconhecimento visual em larga escala, sendo amplamente utilizado em tarefas como classificação de objetos, detecção e localização. Os autores também destacam como objetivo a expansão contínua da base, tanto em volume quanto em diversidade de categorias, visando ampliar ainda mais sua utilidade para a comunidade de pesquisa em visão computacional.

O banco de imagens *Labeled* (HUANG *et al.*, 2008) *Faces in the Wild* (LFW) é uma base de dados composta por 13.233 imagens de rostos de 5.749 indivíduos distintos, capturadas em condições não controladas, com variações significativas de iluminação, pose, expressão facial e resolução. Foi desenvolvida com o objetivo de avaliar algoritmos de reconhecimento facial em cenários realistas, nos quais as imagens não são padronizadas. O trabalho estabelece divisões padronizadas dos dados e define métricas específicas para tarefas de verificação de identidade, com foco na correspondência de pares de rostos, a fim de permitir comparações consistentes entre diferentes métodos. Além da versão original, o *Faces in the Wild* (LFW) disponibiliza variações alternativas, como imagens alinhadas e versões com representações por superpixels, ampliando suas possibilidades de uso em diferentes abordagens experimentais.

O banco de imagens *Microsoft COCO* (*Common Objects in Context*) (LIN *et al.*, 2014) é um conjunto de dados em larga escala voltado para tarefas de detecção e segmentação de objetos em cenas complexas, reunindo mais de 2,5 milhões de instâncias de objetos pertencentes a 91 categorias, distribuídas em aproximadamente 330.000 imagens. O diferencial do *MS COCO* está no foco em imagens não icônicas, ou seja, aquelas em que os objetos aparecem em contextos naturais e frequentemente ocorrem múltiplos objetos por imagem. As anotações foram realizadas por meio do *Amazon Mechanical Turk* e incluem segmentações de instância detalhadas, além de legendas descritivas. O objetivo principal do *dataset* é fornecer um recurso abrangente e desafiador para o avanço de algoritmos de visão computacional, especialmente aqueles voltados à detecção, segmentação semântica e compreensão de cenas em ambientes realistas.

Diferentemente de banco de imagens como o *COCO*, voltado para múltiplas tarefas visuais em contexto geral, ou o *DeepFashion*, cuja ênfase está em atributos visuais da moda, o banco de imagens apresentado neste trabalho foi direcionado a um problema específico: a estimativa de idade a partir de radiografias panorâmicas. Isso o aproxima de propostas como o *CheXpert*, que também trabalha com imagens médicas e aplica técnicas de anotação sistemática baseadas em conhecimento especializado, ainda que em um domínio e com uma complexidade

diferentes.

Ao comparar o banco de imagens a ser apresentado com os trabalhos consolidados na área de inteligência artificial com imagens, observa-se que, embora em menor escala, ele compartilha princípios metodológicos relevantes com projetos como o *ImageNet*, *COCO* e *CheXpert*. Tal como esses conjuntos de dados, a proposta desenvolvida envolveu um processo de curadoria estruturado, com definição prévia da região de interesse e padronização das imagens por meio da aplicação de máscaras. Essa abordagem buscou garantir a consistência da base e facilitar a posterior aplicação de redes neurais profundas.

No que diz respeito ao uso, o conjunto foi elaborado com o intuito de possibilitar a avaliação de redes neurais para estudos de classificação ou regressão. Ainda que restrito a um escopo mais delimitado, o presente banco de imagens permite investigações que dialogam com práticas atuais da área, como o uso de *transfer learning* e a comparação entre arquiteturas distintas. Além disso, abre espaço para análises complementares envolvendo métodos tradicionais de estimativa de idade, comuns na odontologia forense.

A contribuição para a comunidade acadêmica está centrada na disponibilização de uma base organizada, com foco em uma aplicação concreta, e no incentivo a estudos que cruzem abordagens computacionais e biomédicas. Ainda que não tenha a abrangência de grandes *benchmarks*, o presente banco de imagens se propõe a colaborar com a produção de conhecimento na interseção entre inteligência artificial e odontologia, especialmente em contextos nos quais há carência de bases públicas representativas da realidade brasileira.

Tabela 2 – Comparativo entre banco de de imagens na área de Inteligência Artificial

Nome do Dataset	Área de Aplicação	Qtd. de Imagens	Principais Usos	Extratégias de construção	Fonte de Dados	Formato das Anotações	Método de Anotação	Formato de Arquivo	Licença de Uso
ImageNet	Classificação e detecção de objetos	14 milhões+	Treinamento de redes profundas, benchmark em classificações visuais	Anotações manuais extensas, uso de WordNet, escalabilidade	Web (WordNet + busca por imagens)	Labels hierárquicos (WordNet), múltiplas classes	Crowdsourcing (Mechanical Turk)	JPEG	Acadêmica (restrita para pesquisa)
COCO (Common Objects in Context)	Detecção e segmentação de objetos	330 mil+	Análise de objetos em contexto real, segmentação semântica	Contexto realista, múltiplas tarefas por imagem	Fotografias reais com anotações humanas	Bounding boxes, segmentação, legendas	Crowdsourcing e equipe interna	JPEG	Permissiva para pesquisa acadêmica
LFW (Labeled Faces in the Wild)	Reconhecimento facial	13 mil+	Verificação e identificação facial	Ambiente não controlado, imagens da web, anotação de identidade	Imagens da web (celebridades e figuras públicas)	Nome da pessoa, imagem de rosto	Verificação cruzada com nome e imagem	JPEG	Gratuito para uso acadêmico
CheXpert	Diagnóstico médico por imagem (radiografia)	224 mil+ radiografias	Classificação de patologias torácicas	Etiquetas clínicas com incerteza, validação por especialistas	Radiografias clínicas de hospitais universitários	Etiquetas clínicas (presença/ausência/indefinido)	Anotação por NLP em laudos + validação médica	DICOM convertido para PNG	Gratuito mediante aceite de termo
Open Images V6	Visão computacional geral	9 milhões+	Detecção, segmentação e classificação de objetos diversos	Diversidade de categorias, múltiplas anotações por imagem	Imagens públicas com anotações do Google	Bounding boxes, segmentação, relacionamentos	Crowdsourcing com validação automatizada	JPG, PNG	Creative Commons BY 4.0
DeepFashion	Moda e vestuário	800 mil+	Reconhecimento de roupas, busca por similaridade visual	Anotações ricas com landmarks, atributos e categorias	Sites de compras online e catálogos de moda	Landmarks corporais, categorias, atributos	Anotação semiautomática e validação manual	JPG	Permissiva para pesquisa
Presente trabalho	Radiografias panorâmicas odontológicas	1545	Auxílio à estimativa de idade em contextos forenses e odontológicos	Seleção manual de imagens	Clínicas odontológicas especializadas	arquivo CSV	Máscara aplicada automaticamente por regra pré-definida	JPG	Permissiva para pesquisa

3 REVISÃO SISTEMÁTICA DA LITERATURA

Os métodos tradicionais de identificação da idade por meio da arcada dentária são amplamente utilizados na odontologia forense e incluem diversas técnicas baseadas na observação e análise do desenvolvimento dentário (MOHAMED *et al.*, 2023). Um dos métodos mais populares é o de *Demirjian* (SIVRI *et al.*, 2024), amplamente reconhecido e aplicado em várias populações ao redor do mundo. Porém, se faz necessárias formas menos custosas e mais eficazes para apoiar a estimativa de idade por meio de radiografias dentárias (WANG *et al.*, 2022; ONG *et al.*, 2024). Neste sentido, apesar dos avanços significativos e quantidade de trabalhos em áreas de pesquisa relacionadas utilizando redes neurais profundas para estimativa de idade por meio de imagens de radiografias odontológicas panorâmicas, se faz necessário uma sumarização da literatura para identificar as abordagens propostas, seus resultados alcançados, desafios, vantagens e desvantagens, lacunas e oportunidades de trabalhos ainda não explorados.

Assim, este capítulo relata o planejamento, execução e os resultados de uma RSL e está organizado da seguinte forma: a Seção 3.1 apresenta o planejamento da RSL demonstrando o protocolo adotado para o estudo. A Seção 3.2 descreve a execução da revisão de acordo com o protocolo e mostra o processo de seleção dos artigos relevantes. A Seção 3.3 apresenta e discute os resultados da RSL. Ainda, a Seção 3.4 discute as principais ameaças à validade do estudo e a Seção 3.5 discute as conclusões e trabalhos futuros.

3.1 Planejamento

Na fase de planejamento, foi definido o protocolo para conduzir a RSL. As atividades realizadas nesta fase foram: (i) definir o objetivo da RSL, (ii) especificar as questões de pesquisa; (iii) selecionar os bancos de dados para pesquisar os trabalhos; (iv) construir a *string* de busca e (v) aplicar os critérios de inclusão e exclusão.

Objetivo: Identificar e analisar RNP para *Estimativa de Idade por meio de Radiografias Odontológicas Panorâmicas*. Foi definido este objetivo devido o constante relato de uso de RNP na prática forense para identificação de idades de pessoas por meio de radiográficas. É esperado que pesquisadores e usuários de RNP possam se beneficiar de tal sumarização dos trabalhos relatados neste estudo.

Questões de pesquisa: As questões de pesquisa (QP) visam investigar e sumarizar o estado da arte do uso de redes neurais profundas para estimativa de idade. Para atingir esse

objetivo, foi estabelecido três questões de pesquisa de propósito geral.

QP 1 - *Quais são as características dos trabalhos feitos para avaliar as RNP?* Por meio dessa questão de pesquisa serão sumarizadas as principais características dos estudos escolhidos nessa RSL. Foram buscadas informações, tais como: disponibilidades das bases de dados utilizadas, técnicas de processamento de imagem entre outras características.

QP 2 - *Quais são as principais RNP utilizadas?* Essa QP foi formulada com intuito de identificar quais são as principais redes neurais profundas mencionadas na literatura para estimativa de idade por meio de radiografias odontológicas panorâmicas. Por meio dessa questão de pesquisa, é esperado a sumarização das principais redes neurais profundas da literatura.

QP 3 - *Quais são as principais características das RNP?* Essa QP tem com intuito fazer uma sumarização das principais características das RNP. É esperado que seja apresentada uma sumarização dessas características para auxiliar os usuários dessas redes na escolha daquelas que melhor se adaptem às suas necessidades.

Bases de dados eletrônicas: As bases de dados usadas para a busca dos estudos primários foram: ACM Digital Library¹, PubMed², Science Direct³, Scopus⁴, Springer Link⁵ e Web of Science⁶. Elas foram escolhidas porque apresentam uma grande coleção de trabalhos. Além disso, abrigam trabalhos de conferências e periódicos de grande importância para a comunidade acadêmica. Por meio de uma revisão *ad-hoc* foram selecionados alguns trabalhos para compor o estudo. Assim, foi observado quais bases de dados indexavam os trabalhos encontrados, o que também motivou a escolha delas.

String de busca: Para identificar os trabalhos relevantes para estimativa de idade com uso de RNP, foi formulada uma *string* de busca para encontrar estudos primários. Com base em trabalhos relacionados na revisão *ad hoc*, foram selecionadas algumas palavras-chave para compor a *string* de busca. A Tabela 3 apresenta as palavras-chaves escolhidas para compor as *string* de busca. Foram selecionados três tópicos de interesse (primeira coluna) e, em seguida, foram selecionadas as palavras-chave (segunda coluna). Por meio dessas palavras-chave selecionadas foi performada a seguinte *string* de busca utilizada.

¹ <http://dl.acm.org/>

² <https://pubmed.ncbi.nlm.nih.gov/>

³ <https://www.sciencedirect.com/>

⁴ <https://www.scopus.com/>

⁵ <http://link.springer.com/>

⁶ <https://shre.ink/bn01>

Tabela 3 – Palavras-chaves selecionadas

Tópicos	Palavras-chaves selecionadas
Redes neurais	Artificial Neural Networks (ANN), Deep Learning, Machine Learning, Neural Network Architectures, Supervised Learning, Unsupervised Learning, Recurrent Neural Networks (RNN) Convolutional Neural Networks (CNN), Backpropagation, Neural Network Training, Neural Network Applications, Neural Network Optimization, Feedforward Neural Networks, Neural Network Modeling, Neural Network Algorithms, Neural Network Performance
Imagens radiográficas	Panoramic Dental Radiographs, Panoramic Radiography Dental Panoramic X-rays Orthopantomogram, Panoramic Oral Radiography
Faixa etária	Age Groups, Age Classification, Age-Specific Analysis, Age Categories

(Artificial Neural Networks OR ANN OR Deep Learning OR Neural Networks OR Machine Learning OR Convolutional Neural Networks OR CNN) AND (Panoramic Dental Radiographs OR Panoramic Radiography OR Dental Panoramic X-rays OR Orthopantomogram OR Panoramic Oral Radiography) AND (Age Groups OR Age Classification OR Age-Specific Analysis OR Age Categories)

Para avaliar se a *string* de busca estava retornando trabalhos relevantes foi verificado se os trabalhos inicialmente selecionados por meio da revisão *ad hoc* poderiam ser encontrados. Foram identificados na revisão *ad hoc* seis trabalhos de interesse, sendo que quatro trabalhos também puderam ser retornados pela *string* de busca.

Crítérios de Inclusão e Exclusão: Os critérios de inclusão e exclusão permitem classificar cada estudo na revisão sistemática da literatura como um candidato a ser incluído ou excluído (KEELE *et al.*, 2007). Como uma RSL pode envolver um grande número de estudos, foi limitado o escopo deste trabalho para selecionar apenas artigos em inglês publicados e revisados por pares, disponíveis em formato eletrônico, publicados em conferências e periódicos e relacionados aos termos da *string* de busca. Aqueles trabalhos considerados como literatura cinza foram removidos.

Foram incluídos estudos que: **1 - Estimativa de idade por meio de radiografias odontológicas panorâmicas** (que discutem a estimativa de idade com implementação de técnicas de RNP); **2 - Faixa etária** (consideram a perspectiva de grupos etários para estimativa de idade); **3 - Uso de redes neurais** (relatam o uso de RNP, redes convolucionais CNN, redes recorrentes, e outras arquiteturas relevantes para para estimativa de idade); **4 - Estudos empíricos** (realizam análise comparativa entre diferentes modelos de RNP para estimativa de idade); **5 - Métricas** (reportam métricas de desempenho (ex: acurácia, precisão, recall, F1-score) das RNP para estimativa de idade); **6 - Publicação** (foram publicados nos últimos 5 anos).

Foram excluídos estudos que: **1 - Métodos que não usam RNP** (não focam explicitamente em métodos de RNP para estimativa de idade, mas apenas se referem à RNP como um trabalho paralelo); **2 - Estudos que não apresentam uma descrição do comportamento**

das RNP utilizadas (não apresentam uma validação e caracterização das RNPs utilizadas no estudo e que não apresentam resultados empíricos sobre o uso de RNPs para estimativa de idade); **3 - Uso de RNP que não são voltados para estimativa de idade** (que utilizam RNP para outras aplicações odontológicas, como segmentação ou detecção de cáries, e que não estejam focados em estimativa de idade); **4 - Estudos que não identificam as idades dos pacientes** (que não consideram a perspectiva de grupos etários); **5 - Prática odontológica** (que não aplicam os modelos de RNP no contexto clínico ou que falham em relacionar os resultados à prática odontológica); **6 - Artigos curtos** (em formato de resumo sem acesso ao texto completo, ou trabalhos curtos de uma ou duas páginas); **7 - Literatura cinza** (teses, dissertações, e relatórios técnicos que não foram publicados em revistas revisadas por pares ou conferências relevantes).

Seleção inicial do estudo: Os pesquisadores autores desse trabalho leram o título, resumo e palavras-chaves dos estudos recuperados e os classificaram como incluídos ou excluídos, aplicando os critérios de inclusão e exclusão. Para este trabalho foi adotado o seguinte critério: um estudo foi incluído se ele satisfizesse todos os critérios de inclusão e excluído se satisfizesse ao menos um critério de exclusão.

Seleção final: Os autores desse trabalho fizeram a leitura completa dos artigos selecionados para identificar os tópicos de interesse a serem sumarizados para compor o corpo de conhecimento deste trabalho. A Tabela 4 apresenta quais foram os itens analisados para sumarizar os artigos selecionados. Os itens foram separados em três grupos: sendo eles classificação do trabalho, qualificação do estudo e métricas utilizadas. Para auxiliar todo o processo de planejamento, execução, e análise dos dados reportados na RSL foi utilizada a ferramenta START (FABBRI *et al.*, 2016). Essa ferramenta auxilia todas as etapas de criação de uma RSL.

Tabela 4 – Itens analisados nos trabalhos selecionados

Grupos	Item analisado
Classificação do trabalho	Ano de publicação, Tipo de publicação, Evento publicado, Classificação Capes e Número de Citações
Qualificação do estudo	1) De qual população de qual país foram levantados os dados? 2) O estudo apresenta uma rede neural? 3) O estudo apresenta o número de imagens analisadas? Qual foi? 4) Qual é o grupo etário que o estudo trabalha? 5) A base de dados trabalhada no artigo é pública? 6) Foi feita uma técnica de processamento nas imagens? 7) Foi utilizado um classificador? 8) O estudo apresenta classificação de imagens de radiografias odontológica panorâmica?
Métricas utilizadas	1) Quais foram as métricas utilizadas? 2) Caso tenha, qual foi a acurácia relatada? 3) O estudo reporta métricas de desempenho?

3.2 Execução

Esta seção apresenta a execução dos protocolos descritos na Seção 3.1. Assim, foi aplicado a *string* de busca nas seis bases de dados para buscar estudos primários. Com a finalidade de encontrar abordagens mais atuais e consolidadas na literatura foram incluídos apenas trabalhos publicados entre os anos de 2020 a 2024. O processo de busca foi realizado de 11 agosto a 01 de setembro de 2024.

Com as informações dos trabalhos encontrados extraídas no formato *BibTex* para ferramenta StArt, os critérios de inclusão e exclusão foram aplicados à filtragem dos estudos. As informações dos trabalhos tais como título, resumo, e palavras-chave dos trabalhos foram analisadas diretamente na ferramenta StArt. Foi dada uma classificação de prioridade de leitura para cada trabalho como: muito alta, alta, baixa e muito baixa. Além de classificar o trabalho como aceito ou rejeitado.

A Figura 6 apresenta as etapas e buscas dos estudos selecionados nesta revisão sistemática da literatura. Inicialmente, foi feita a busca dos trabalhos nas bases de dados, sendo que nessa etapa foram encontrados 85 publicações. A *Science Direct* foi a base que mais retornou trabalhos, no total 25 publicações e a ACM foi a que menos retornou com apenas quatro trabalhos. Além disso, dois estudos identificados na Revisão *ad hoc* não estavam presentes em nenhuma das bases consultadas, mas foram incluídos manualmente para compor o conjunto de estudos analisados. Outros quatro estudos provenientes da *ad hoc* foram encontrados nas bases de dados, sendo dois na *Spring link*, um na *Web of Science* e um na *Scopus*.

Posteriormente, o processo de seleção dos trabalhos, conforme indicado na Figura 6 foi feito em três etapas: remoção de duplicados (A) sendo que 20 artigos foram classificados como duplicados, esta etapa foi feita por meio da ferramenta StArt, assim, passaram para a próxima etapa (B) 65 trabalhos. Após a leitura do título, resumo e palavras-chave dos trabalhos 22 artigos foram selecionados e 42 artigos foram rejeitados. Finalmente, foram lidos os textos completos dos 22 artigos (C), dos quais 5 foram excluídos por não atenderem aos critérios de inclusão e exclusão. Dessa forma, esta RSL apresenta a sumarização de 17 trabalhos considerados relevantes para responder as questões de pesquisa levantadas na Seção 3.1.

A Tabela 5 apresenta os trabalhos selecionados para compor a presente revisão sistemática da literatura. Os trabalhos estão ordenados por o ano de publicação. Como pode ser

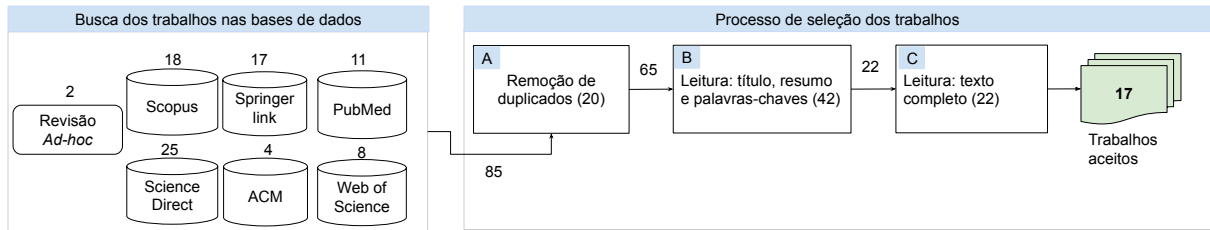


Figura 6 – Processo de busca e seleção dos trabalhos

observado na Tabela 5 foram selecionados seis trabalhos do ano de 2024, para os anos de 2023, 2022 e 2021 foram aceitos segundo os critérios de inclusão e exclusão estabelecidos: cinco, um e quatro trabalhos respectivamente. Finalmente, para o ano de 2020 foi selecionados apenas um trabalho.

3.3 Análise

Esta seção apresenta as respostas das questões de pesquisa (QP) realizadas na Seção 3.1. Dessa forma, Seção 3.3.1 responde a QP sobre as principais características dos estudos selecionados apresentando uma visão geral desses estudos. A Seção 3.3.2 apresenta as redes neurais que os estudos selecionados utilizam. Finalmente, a Seção 3.3.3 sumariza as principais características das redes neurais identificadas no estudo.

3.3.1 Quais são as Características dos Trabalhos Feitos para Avaliar as RNP?

Em resposta, RQ1 formulada na Seção 3.1, esta seção apresenta uma visão geral dos estudos primários selecionados na RSL conduzida neste trabalho. Com a finalidade de qualificar os trabalhos selecionados no estudo, a Figura 7 apresenta o quantitativo de trabalhos em relação ao Qualis (Figura 7(a)), população do estudo (Figura 7(b)) ano de publicação (Figura 7(c)) e evento de publicação (Figura 7(d)).

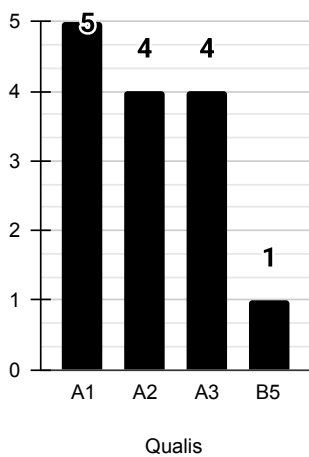
Como pode ser observado a relevância dos trabalhos a maioria dos estudos (cerca de 76%) apresenta o Qualis superior (A1, A2 e A3). Porém, três estudos não foi possível identificar ou eles não possuem Qualis. Além disso, foi identificado para a população de quais países os estudos pertencem. A China e Coreia do Sul apresentam seis e quatro trabalhos, respectivamente. Foi selecionado apenas um estudo referente à população brasileira.

Ainda, no ano de 2024 foram publicados seis trabalhos, seguido do ano de 2023 com

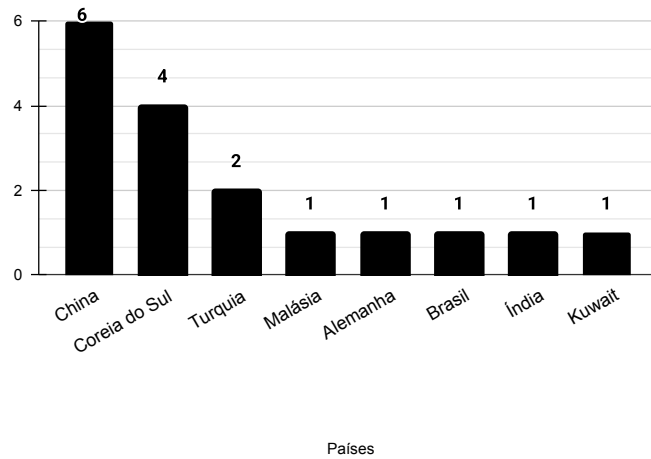
Tabela 5 – Trabalhos selecionados no estudo

Título do trabalho	Referência	Ano
<i>Human Age Estimation Using Deep Convolutional Neural Network based on Dental Images (Orthopantomogram)</i>	(SATHYAVATHI, 2024)	2024
<i>Accuracy of automated forensic dental age estimation lab (F-DentEst Lab) on large Malaysian</i>	(MOHAMMAD <i>et al.</i> , 2024)	2024
<i>Fully automated deep learning approach to dental development assessment in panoramic radiographs</i>	(ONG <i>et al.</i> , 2024)	2024
<i>Estimation of human age using machine learning on panoramic radiographs for Brazilian patients</i>	(OLIVEIRA <i>et al.</i> , 2024)	2024
<i>Dental age estimation: A comparative study of convolutional neural network and Demirjian's method.</i>	(SIVRI <i>et al.</i> , 2024)	2024
<i>Deep learning methods for fully automated dental age estimation on orthopantomograms</i>	(SHI <i>et al.</i> , 2024)	2024
<i>Estimating chronological age through learning local and global features of panoramic radiographs in the Korean population</i>	(YEOM <i>et al.</i> , 2023)	2023
<i>Forensic Dental Age Estimation Using Modified Deep Learning Neural Network</i>	(ATAŞ <i>et al.</i> , 2022a)	2023
<i>A population-based study to assess two convolutional neural networks for dental age estimation</i>	(WANG <i>et al.</i> , 2023)	2023
<i>An automatic methodology for full dentition maturity staging from OPG images using deep learning</i>	(DONG <i>et al.</i> , 2023)	2023
<i>Age Group Classification of Dental Radiography without Precise Age Information Using Convolutional Neural Networks</i>	(KIM <i>et al.</i> , 2023)	2023
<i>Age Estimation using Panoramic Radiographs by Transfer Learning</i>	(MU; GANG, 2022)	2022
<i>Age-group determination of living individuals using first molar images based on artificial intelligence</i>	(KIM <i>et al.</i> , 2021)	2021
<i>DENSEN: a convolutional neuralnetwork for estimating chronological agesfrom panoramic radiographs</i>	(WANG <i>et al.</i> , 2022)	2021
<i>Age estimation on panoramic dental X-ray images using deep learning</i>	(WALLRAFF <i>et al.</i> , 2021)	2021
<i>Accurate age classification using manual method and deep convolutional neural network based on orthopantomogram images</i>	(GUO <i>et al.</i> , 2021)	2021
<i>Dental Age Estimation Based on X-ray Images</i>	(MUALLA <i>et al.</i> , 2020)	2020

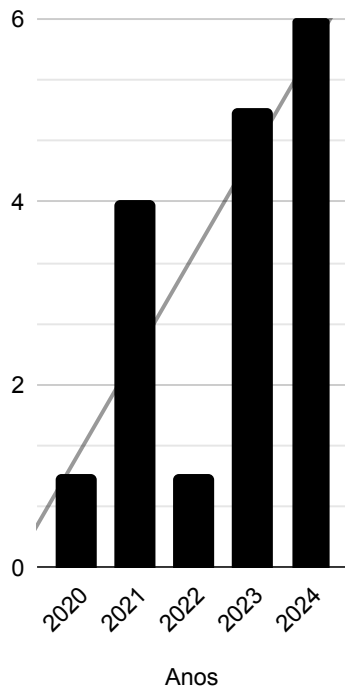
cinco trabalhos, como pode ser observado na Figura 7(c) a pesquisa de redes neurais profundas para estimativa de idade está em ascensão. Finalmente, dos 17 trabalhos selecionados 16 (cerca de 94%) trabalhos são estudos de periódicos e um trabalho foi publicado em *workshop*. Os eventos onde os trabalhos selecionados foram publicados tem como principais destaques, os periódicos *Scientific Reports* e *BMC oral health* com três publicações cada.



(a) Qualis



(b) Países



(c) Anos



(d) Publicação

Figura 7 – Distribuição de trabalhos por Qualis e países

Foi analisado se nos estudos seleccionados utilizaram-se técnicas de pré-processamento nas imagens. A Figura 8 apresenta uma sumarização das técnicas de processamento das imagens como preparação para redes neurais. Como pode ser observado, o processamento mais utilizado foi o *Data Augmentation*, cerca de 52% dos trabalhos seleccionados utilizam essa técnica para preparar suas imagens. O segundo tratamento que mais ocorreu nos estudos analisados foi o Redimensionamento das imagens com cerca de 21% dos trabalhos fizeram menção ao uso dessa técnica. A fim de encontrar as técnicas de *data augmentation* utilizadas nos trabalhos seleccionados foi feito uma sumarização das atividades feitas para processamento de imagem relacionadas a técnicas de *data augmentation* (Figura 9). Esse detalhamento foi feito dado que *data augmentation* foi a técnica mais citada nos trabalhos seleccionados.

Figura 8 apresenta uma sumarização das técnicas mencionadas nos trabalhos relacionadas à *data augmentation* com o número de ocorrências das mesmas. Como pode ser observado a técnica *Flip* horizontal e Rotação aparecem indicadas em cinco trabalhos, seguido de Translação com quatro trabalhos.

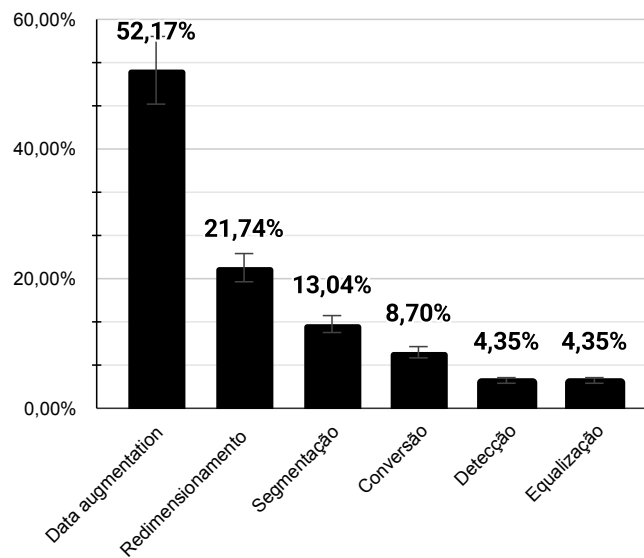


Figura 8 – Técnicas de processamento

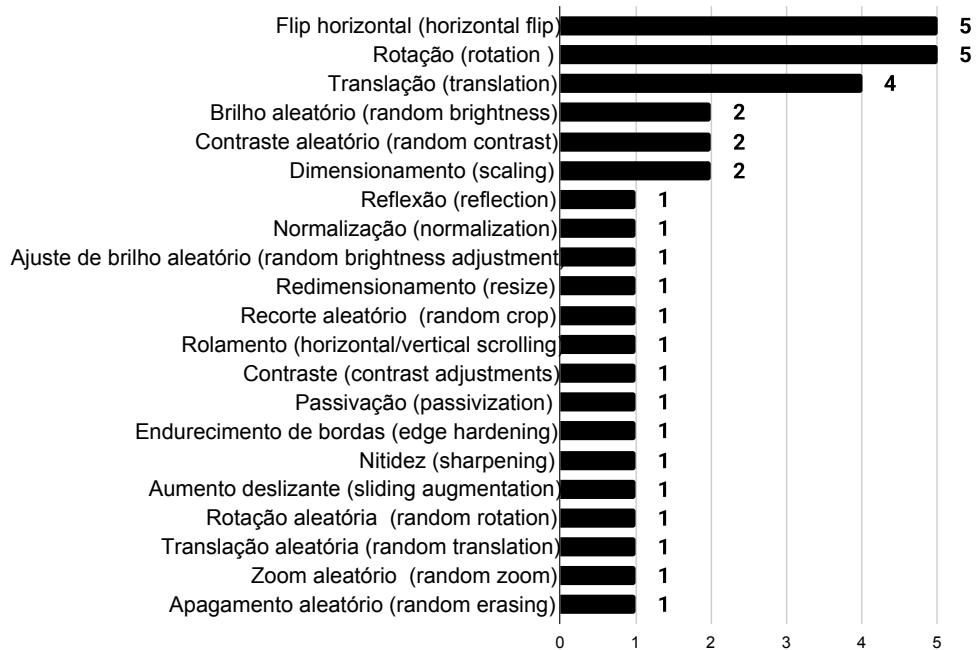
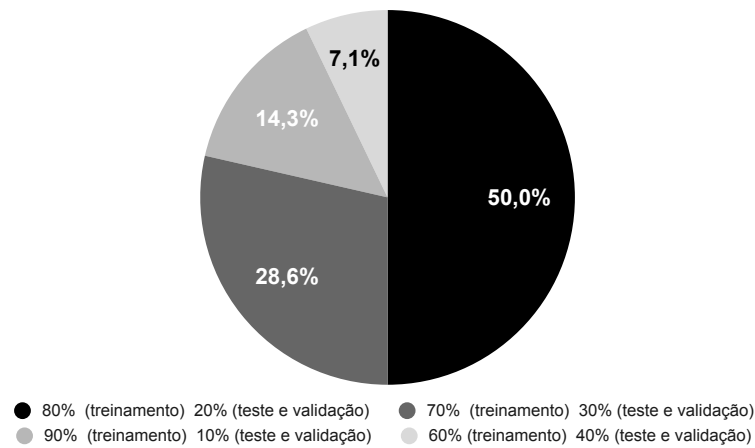
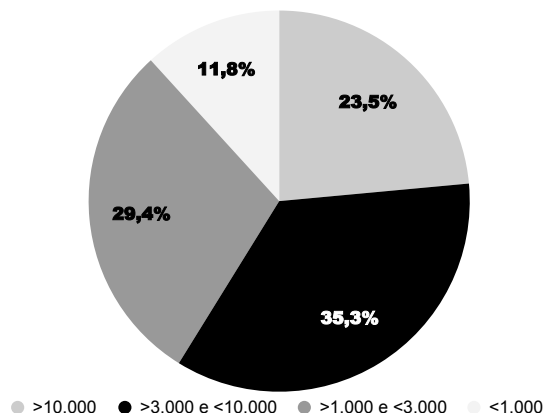


Figura 9 – Tipos de data augmentation

Além disso, foi observado a estratégia adotada nos trabalhos para escolher a porcentagem de imagens para teste, treinamento e validação. A Figura 10(a) apresenta sumarização dos dados dos 16 estudos (um trabalho não mencionou qual foi a divisão de dados utilizada). Como pode ser observado na Figura 10(a) cerca de 50% dos trabalhos utilizam a divisão da base de dados em 80% das imagens para treinamento e 20% restante são para teste e validação. Ainda, na Figura 10(b) definiu-se um agrupamento da quantidade de imagens utilizadas nos estudos. Este agrupamento foi composto de quatro grupos: estudos que apresentaram uma quantidade de imagens maior que 10.000 imagens (cerca de 23% dos estudos), estudos em que o número de imagens foram entre 3.000 e 10.000 (cerca de 35% dos estudos), estudos que apresentaram entre 1.000 e 3.000 mil imagens (cerca de 29% dos estudos). Finalmente, grupos que apresentaram menos que 1.000 imagens (cerca de 11% dos estudos). A Figura 10(b) apresenta o quantitativo de estudo em cada grupo. Como pode ser observado o grupo entre 3.000 e 10.000 imagens foi o grupo mais representativo, com cerca de 35% dos estudos.



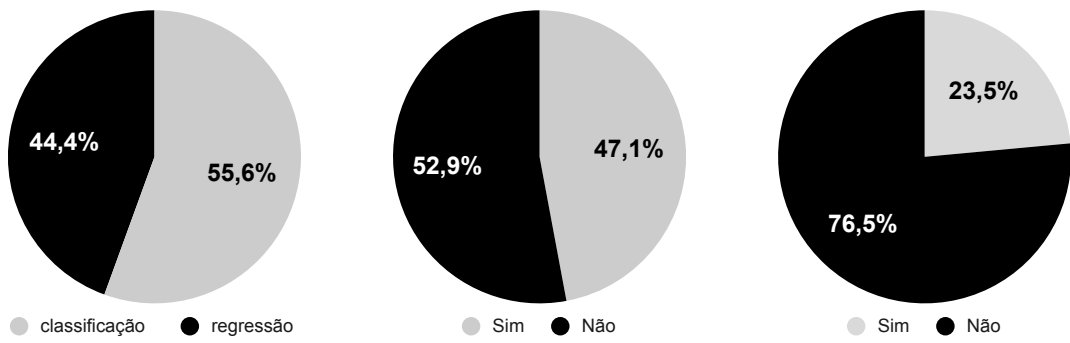
(a) Divisão de dados



(b) Imagens

Figura 10 – Divisão dos dados e quantidade de imagens

Ainda, a Figura 11 apresenta uma sumarização dos trabalhos em relação, o modelo utilizado (classificação ou regressão), (Figura 11(a)) se o trabalho utilizou *Transfer learning* (Figura 11(b)) e se o trabalho disponibiliza a base de dados utilizada no estudo (Figura 11(c)). Finalmente, os resultados desta seção são importantes pois além de caracterizar e sumarização os trabalhos selecionados, também podem apoiar a construção protocolos para futuros estudos.



(a) Modelo

(b) Transfer Learning

(c) Banco de dados

Figura 11 – Regressão/classificação, Transfer learning e Base de dados

3.3.2 Quais são as Principais RNP Utilizadas?

A Figura 12 apresenta as 25 redes neurais identificadas no presente estudo e a quantidade de ocorrências dessas redes. Além disso, foram identificadas as redes utilizadas para tarefas de regressão e classificação. Como pode ser observado a rede *VGG-116* foi a rede mais utilizada para regressão com três ocorrências e a rede *SOS-Net* para classificação, com quatro ocorrências. Este resultado pode apoiar pesquisadores que desejarem escolher redes artificiais para compor seus estudos.

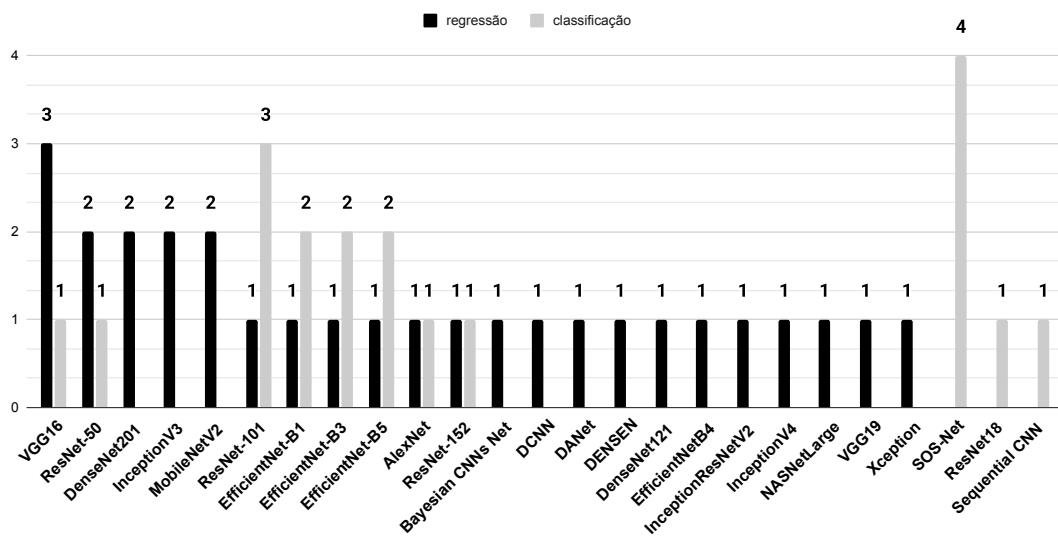


Figura 12 – Redes identificadas

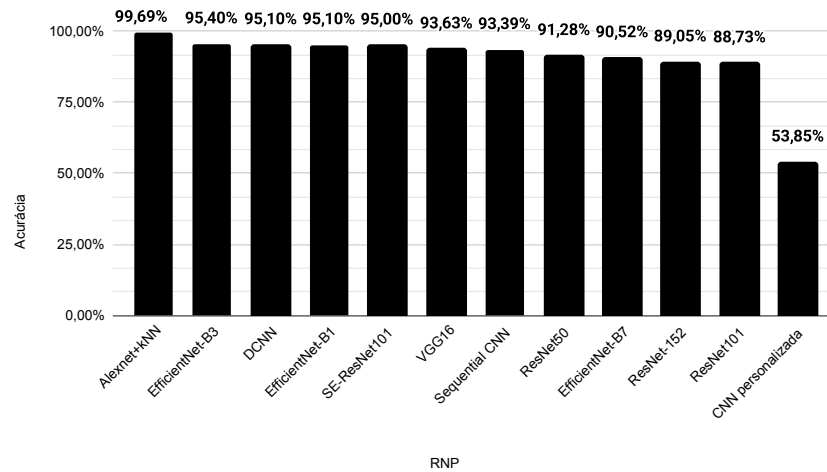
3.3.3 Quais são as Principais Características das RNP?

Nos trabalhos analisados, observou-se que as principais métricas reportadas incluem acurácia (todos os trabalhos), *recall* (75% dos trabalhos), precisão (50% dos trabalhos), MAE (*Mean Absolute Error*) (todos os trabalhos), MSE (*Mean Squared Error*) (55%) e RMSE (*Root Mean Squared Error*) (55%). Essas métricas permitem a avaliação do desempenho dos modelos, sendo essencial a escolha adequada de cada uma, especialmente considerando a diferença entre tarefas de classificação e regressão. A seleção da métrica impacta diretamente a interpretação dos resultados e a tomada de decisões sobre quais modelos apresentam melhor desempenho em diferentes contextos.

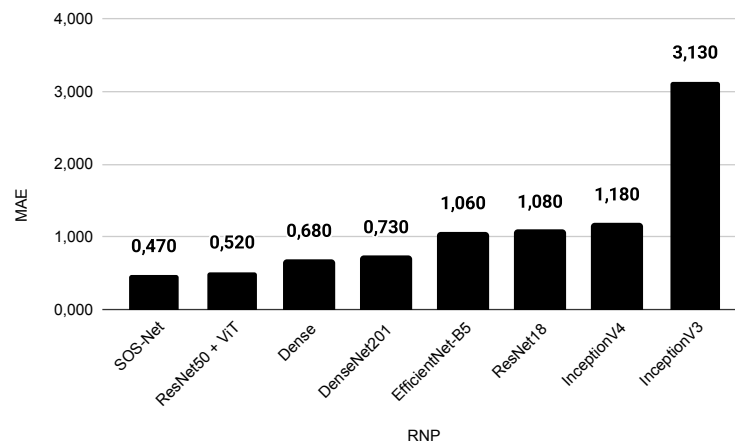
Como as métricas de acurácia e *Mean Absolute Error* (MAE) foram as mais utilizadas nos trabalhos, a Figura 13 apresenta os valores encontrados para estas métricas. A Figura 13(a) apresenta a sumarização dos melhores resultados de acurácia para RNP aplicadas à classificação, segundo a menção dos trabalhos selecionados. Uma vez que todos os estudos analisados reportaram a acurácia, optou-se por sumarizar os seus resultados. A acurácia é uma métrica comum e amplamente utilizada por pesquisadores da área para avaliar o desempenho de seus modelos e compará-los com resultados disponíveis na literatura. A rede neural com melhor ganho entre as demais redes analisadas foi *Alexnet+kNN* (com cerca de 99% de acurácia), conforme apresentado pela Figura 13(a). Além disso, a Figura 13(b) apresenta os melhores resultados da métrica MAE para os estudos de regressão. O MAE mede a média das diferenças absolutas entre os valores reais e os valores preditos pelo modelo.

A escolha dessa métrica se justifica porque, ao contrário do *Mean Squared Error* (MSE), que penaliza erros maiores de forma quadrática, o MAE expressa o erro médio em termos absolutos e na mesma unidade da variável predita. Essa característica facilita a análise comparativa entre diferentes abordagens e evita que *outliers* influenciem desproporcionalmente os resultados. Além disso, é uma métrica intuitiva. Conforme apresentado na Figura 13(b) a rede neural com melhor resultado foi a *SOS-NET* com o valor para a métrica MAE com cerca de 0,470 anos.

A Figura 14 apresenta o comportamento das redes neurais encontradas nos estudos selecionados, levando em conta a faixa etária. A Figura 14(a) apresenta, para as redes de classificação, os valores de acurácia. Já a Figura 14(b) apresenta o valores da métrica MAE para as redes de regressão. Como pode ser observado na Figura 14(a) as medianas dos três grupos de faixa etária estão muito próximas, ainda é possível observar que para o de 3 a 14 anos as



(a) Classificação



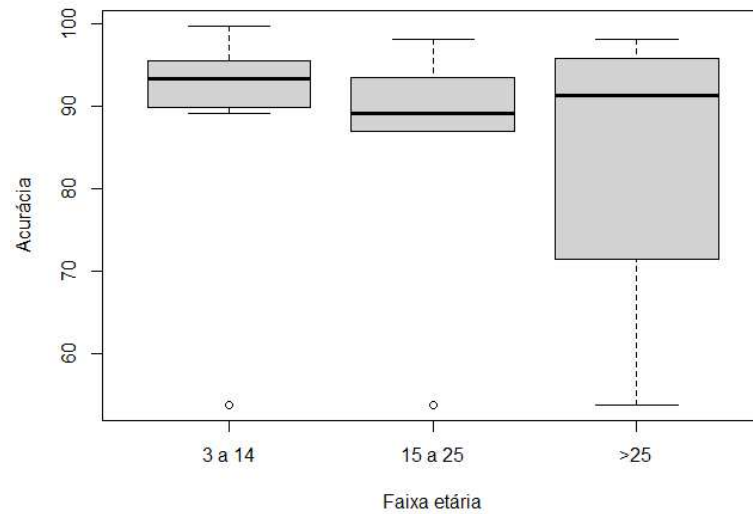
(b) Regressão

Figura 13 – Comportamento das redes neurais

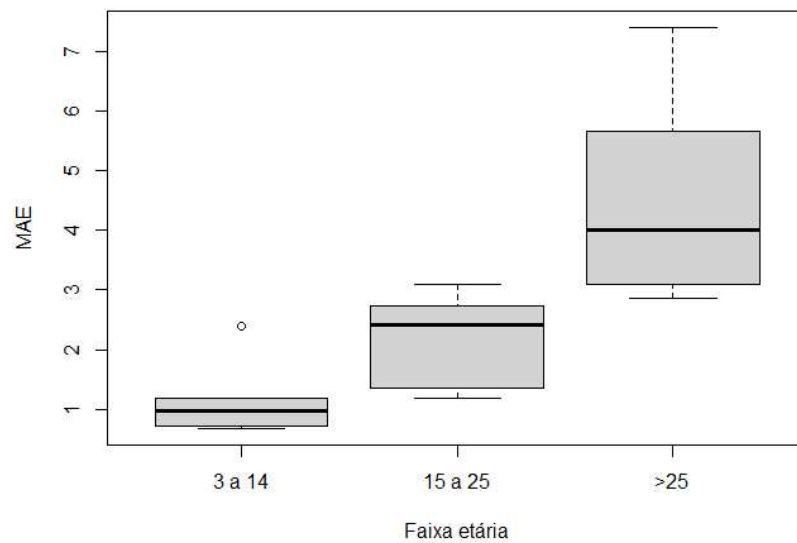
medidas de acurácia apresentam um valor simétrico. Para o grupo acima de 25 anos, houve maior dispersão dos valores de acurácia. Dessa forma, foi observado que, para as crianças (de 3 a 14 anos), as redes puderam apresentar melhores resultados em comparação as demais faixas etárias. Entretanto, foi percebido que para as idades acima de 25 anos os modelos apresentam uma queda nos valores das métricas de avaliação.

Como pode ser observado na Figura 14(b) existe uma variação considerável da mediana para os grupos. Os quartis para o grupo de 3 a 14 anos são menores que os quartis do grupo de 15 a 25 que também são menores que o grupo maior de 25 anos. Ainda, há um aumento significativo do erro em termos da mediana, por exemplo: de 3 a 14 anos (o erro foi de cerca de um ano), já para o grupo de 15 a 25 anos, o erro foi de cerca de 2,5 anos e para o grupo de maior que 25 anos o erro foi de cerca de 4 anos. Desta forma, os menores erros ocorrem na faixa etária de 3 a 14 anos. O MAE apresenta baixa variabilidade, indicando um bom desempenho

dos modelos nessa faixa e previsões mais precisas. No grupo de 15 a 25 anos, o erro médio é maior e há maior dispersão dos valores. Isso sugere que os modelos encontram mais dificuldades para prever a idade corretamente nesse intervalo. Para indivíduos acima de 25 anos, o MAE é ainda mais alto, com mediana elevada e maior variabilidade. As previsões tornam-se menos precisas e os erros são mais dispersos. Esse resultado pode estar relacionado à maior regularidade do desenvolvimento dentário em crianças, facilitando a estimativa da idade. A dificuldade de diferenciação entre 15 e 25 anos decorre da estabilidade da dentição, reduzindo sinais claros de variação etária. Acima dos 25 anos, fatores como desgaste dentário individual, implantes e uso de próteses introduzem maior variabilidade, tornando a estimativa mais complexa.



(a) Acurácia



(b) MAE

Figura 14 – Distribuição dos valores das métricas Acurácia e MAE

3.4 Ameaças a Validade

Foi realizado uma RSL resumindo os principais achados sobre as trabalhos que utilizam RNP para estimativa de idade, conforme apresentado nas seções anteriores. No entanto, apesar do esforço dos autores para condução deste trabalho existem, algumas ameaças podem ter afetado a validade dos dados encontrado. Nesta seção são discutidos algumas ameaças à validade que podem afetar os achados dessa pesquisa.

Replicabilidade do estudo: O estudo foi projetado para ser aplicado nas diferentes

bases de dados sem realizar alterações nos padrões de buscas, sendo necessário se adequar apenas às normas de utilização da biblioteca, a fim de evitar a descaracterização do estudo. Para mitigar os efeitos das discrepâncias entre as bases de dados, foi feita a definição das palavras chave para compor a *string* de busca adaptada para cada base de dados.

Resultados alcançados: Foi realizada a sumarização das redes neurais encontradas, disponibilizadas na Seção 3.2. No entanto, esse procedimento pode ser afetado pela compreensão para agrupar as redes encontradas. Para minimizar esse problema, somente foram documentadas as informações que estavam explícitas nos artigos utilizados. As descobertas deste trabalho, também podem ser afetadas pelas informações disponíveis nas bases de dados, podendo não indexar todos os artigos relevantes para o foco do estudo. Para minimizar esse problema, foram selecionados seis bases de dados diferentes e discutido na Seção 3.1.

Generalização do estudo: Alguns fatores podem impedir a generalização dos dados encontrados na pesquisa. Por exemplo, as 25 redes neurais sumarizadas neste estudo, podem não representar adequadamente todas as redes neurais relevantes para a estimativa de idades por meio de imagens de radiografia panorâmica. As buscas realizadas neste estudo foram restritas ao período entre 2020 e 2024, podendo essa limitação afetar os resultados alcançados, uma vez que redes neurais desenvolvidas em outro período. No entanto, este estudo foi realizado com o objetivo de coletar informações recentes sobre novas redes neurais para estimativa de idade por meio de radiografia panorâmica. A *string* de busca utilizada para identificação primária dos artigos nas bases de dados utilizadas pode ter afetado nos trabalhos selecionados para compor o estudo, devido, uma vez que, um dado artigo não encontrado apresente uso de redes neurais para o propósito definido neste trabalho porém não tenha sido encontrado por meio da *string* de busca utilizada. Esse problema, pode ter ocorrido devido a alguma omissão neste artigo na informações de seus metadados (título, palavras-chave e resumo). Para mitigar esse problema foi certificado que o conjunto de artigo previamente selecionado na busca *ad hoc* se poderia ser encontrados por meio da *string* de busca performada. A data de pesquisa realizada na biblioteca, tendo em vista que um artigo pode ter sido indexado após a data de coleta das informações utilizadas neste estudo.

3.5 Considerações Finais

Como resultado desta RSL, foram identificadas 25 redes neurais. Dentre estas, as que tiveram melhor desempenho foram as redes *AlexNet* em trabalhos de classificação e *SOS-Net*

em trabalhos de regressão. Também foi observado que para o grupo das crianças (de 3 a 14 anos), as redes atingiram melhores resultados em comparação às demais faixas etárias. Ainda, as métricas acurácia e MAE foram as mais utilizadas para avaliar as redes neurais. Portanto, conclui-se que embora redes neurais sejam aplicadas na classificação de imagens radiográficas, e ainda com resultados promissores, faz-se necessário que os estudos avancem na identificação dos grupos etários acima de 25 anos de idade.

O próximo capítulo apresenta um banco de imagens criado para apoiar o presente estudo e demais estudos futuros que possam utilizar imagens de radiografias odontológicas panorâmicas. Desta forma, toda a comunidade de pesquisadores de redes neurais profundas, forense e de saúde bucal podem ser beneficiar com o banco de imagens proposto.

4 BANCO DE IMAGENS

Este capítulo apresenta as etapas seguidas para a construção do banco de imagens utilizado neste estudo. O banco de imagens disponibilizado neste trabalho pode ser utilizado para apoiar pesquisas na área de redes neurais profundas, bem como outros propósitos, tais como: métodos tradicionais de estimativa de idade por meio de radiografias odontológicas panorâmicas, como o método de *Demirjian* (SIVRI *et al.*, 2024; WALLRAFF *et al.*, 2021), detecção de lesões ou cistos (KWON *et al.*, 2020; SON *et al.*, 2021), classificação do sexo, masculino ou feminino, com base em características craniofaciais visíveis na radiografia (ARTHANARI *et al.*, 2024) e segmentação de dentes (RUBIU *et al.*, 2023).

A organização estrutural e metodológica do banco de imagens foi discutida com especialistas da área odontológica. A consultoria para criação deste banco de imagem se deu por meio de uma parceria técnico-científica estabelecida entre o Programa de Pós-Graduação em Engenharia Elétrica e Computação¹ e o curso de Odontologia², ambos pertencentes à Universidade Federal do Ceará, Campus de Sobral.

A principal motivação para a criação deste banco de imagens foi a falta de um banco de imagens de radiografia odontológica panorâmica público que pudesse ser utilizado nesta dissertação. Posteriormente à revisão sistemática da literatura, conduzida no Capítulo 3, foi encontrado um banco de imagens que pode ser acessado mediante liberação dos autores. Assim, este capítulo tem como principal contribuição a disponibilização do banco de imagens descrito. Os beneficiários deste capítulo são pesquisadores e profissionais da saúde que podem utilizar o banco de imagens, bem como o protocolo adotado para sua expansão. Este protocolo, discutido neste capítulo, possibilita a incorporação de novas imagens de radiografias odontológicas panorâmicas. O banco de imagem pode ser acessível no repositório aberto para toda a comunidade³.

As seções deste capítulo estão organizadas da seguinte forma: A Seção 4.1 apresenta o processo de aquisição de imagens. A Seção 4.2 detalha como as imagens foram preparadas para compor o banco de imagens. Ainda, a Seção 4.3 apresenta uma qualificação das imagens. As Seções 4.4 e 4.5 apresentam o banco de imagens e um exemplo de uso, respectivamente. Finalmente, as Seções 4.6 e 4.7 discutem as ameaças à validade, com suas tratativas para mitigação e as considerações finais do capítulo, respectivamente.

¹ <<https://ppgeec.ufc.br/pt/>>

² <<https://www.ufc.br/ensino/guia-de-profissoes/174-odontologia>>

³ <<https://rsl-redes-neurais-profundas.github.io/rsl/>>

4.1 Aquisição das Imagens

A Figura 15 apresenta as etapas realizadas para a aquisição das imagens de radiografias panorâmicas que compõem o banco de imagens deste estudo. Como pode ser observado na Figura 15, foram selecionadas imagens de três clínicas de dois estados distintos. A seleção das imagens foi feita diretamente do sistema disponibilizado pelas clínicas, foram selecionados no total 1.545 imagens. O processo de seleção das imagens ocorreu no período de Janeiro de 2024 a novembro de 2024, sendo realizado em três etapas, da seguinte forma:

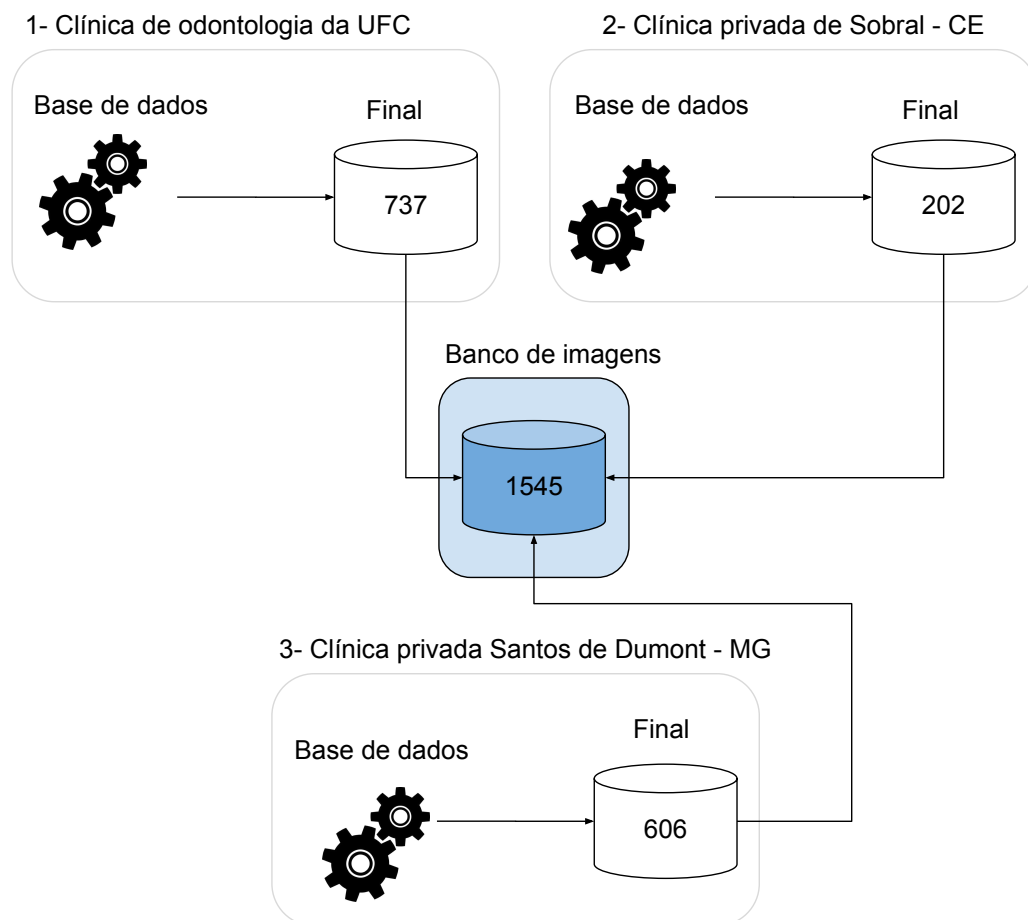


Figura 15 – Seleção de imagens para compor o banco de imagens de radiografias panorâmicas

1 - Clínica de odontologia da UFC: A construção do banco de imagens foi iniciada por meio de uma parceria institucional entre o curso de Engenharia da Computação e o curso de Odontologia da Universidade Federal do Ceará, Campus de Sobral. Essa colaboração viabilizou o acesso a um acervo de 868 radiografias panorâmicas odontológicas, extraídas do sistema da clínica universitária, todas no formato “.jpg”. O processo de aquisição envolveu a exportação individual das imagens a partir dos sistemas de armazenamento digital da clínica, com acompanhamento da equipe técnica local. Após a análise inicial, foram descartadas as imagens em que os pacientes apresentavam edentulismo total ou número inferior a seis dentes, independentemente de quais dentes estavam presentes. Essa etapa resultou na exclusão de 131 imagens, totalizando 737 imagens válidas provenientes da clínica da universidade.

2 - Clínica privada de Sobral - CE: Além das imagens coletadas na clínica odontologia da UFC-Sobral, também foram coletadas imagens odontológica em uma clínica privada na cidade de Sobral-CE. A coleta dessas imagens foi conduzida diretamente pela pesquisadora, sendo realizada em cerca de 25 dias, por meio do acesso à base de dados da clínica com a devida autorização da mesma. Cada imagem foi selecionada individualmente e manualmente na base de dados, omitindo quaisquer dados que pudesse identificar os pacientes. Juntamente com as imagens, foram coletados os dados de idade e sexo dos pacientes. Foram descartadas da seleção, as imagens que apresentaram as seguintes características: que não eram de radiografias panorâmicas, que apresentavam menos que seis dentes ou edentulismo total, que não apresentava a região bucal por completo e aquelas em que os dados de sexo e idade não estavam disponíveis. Finalmente, foram selecionadas 202 imagens para compor o banco de imagens radiográficas.

3 - Clínica privada de Santos Dumont - MG: A extração dos arquivos foi realizada manualmente e individualmente, acessando os sistemas clínicos. As imagens foram selecionadas e baixadas utilizando os mesmos critérios de seleção feito para a Clínica privada de Sobral. Foram selecionadas as imagens juntamente com os dados de idade e sexo do paciente. Nesta clínica, foram obtidas 606 imagens no total.

4.2 Preparação das Imagens

O primeiro passo da preparação das imagens consistiu na anonimização dos dados sensíveis. Muitas das imagens coletadas apresentavam o nome do paciente embutido diretamente na imagem. Para garantir a privacidade dos indivíduos, foi realizada a remoção dessas informações por meio de cortes na região correspondente.

As imagens também foram renomeadas, substituindo identificadores pessoais por uma numeração sequencial única. A estrutura de nomeação dos arquivos passou a conter apenas o número da imagem e de forma separada, uma planilha de metadados foi criada, contendo as colunas: Número da imagem, Idade e Sexo.

Posteriormente, foi aplicada uma etapa de padronização visual. Todas as imagens passaram por um recorte utilizando uma máscara com definição de Região de Interesse (ROI), com coordenadas proporcionais ao tamanho da imagem. Essa máscara oculta as áreas irrelevantes da imagem, de forma a destacar apenas a região anatômica de interesse para análise e treinamento dos modelos de aprendizado profundo.

A Figura 16 apresenta um exemplo de imagem de radiografia panorâmica que exibia o nome do paciente (omitido por questões de anonimização), e a imagem modificada, na qual foi aplicada ao redor de toda imagem. O processo de aplicação da máscara na imagem foi feito por meio de um *script Python*.

Para garantir a replicação e ampliação deste banco de imagens, o respectivo *script* encontra-se disponível no repositório do banco de imagens⁴. Desta forma, todas as 1.545 imagens que constituem o banco de imagens passaram por este processo de aplicação de uma máscara.

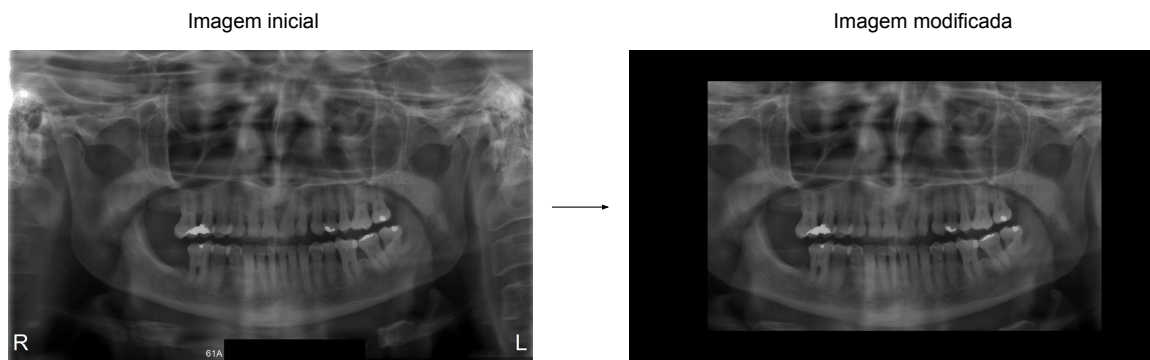


Figura 16 – Aplicação de máscara para seleção de interesse da imagem

⁴ <<https://rsl-redes-neurais-profundas.github.io/rsl/>>

4.3 Qualificação das Imagens

A etapa de qualificação consistiu em verificar a integridade visual das imagens (nitidez, ausência de artefatos, enquadramento da arcada dentária) e assegurar a presença de características anatômicas relevantes para os objetivos do projeto, como a presença de número suficiente de dentes e arcadas bem posicionadas.

Como métrica de qualificação, foi adotado um critério visual mínimo de qualidade para garantir que o modelo de rede neural receba dados com características consistentes. Imagens com baixa resolução, cortes anatômicos importantes ou distorções visuais foram desconsideradas. Todas as imagens do banco estão do formato *JPG*, totalizando o tamanho 3 gigabytes.

4.4 Banco de Imagens

A Figura 17 apresenta a distribuição das 1.545 imagens que compõem o banco de dados de radiografias panorâmicas odontológicas, organizadas conforme a idade dos pacientes no momento da realização do exame e sexo identificado no cadastro das clínicas. Como pode ser observado na Figura 17, a idade que apresenta mais imagens no banco é de paciente de 21 anos, com 45 imagens de pacientes do sexo feminino e 20 imagens de pacientes do sexo masculino, totalizando 65 imagens. Para os pacientes de idades: 1, 86, 89 e 91 o banco de imagens possuem apenas uma imagem de paciente do sexo feminino. Para a idade de 81 anos, o banco possui apenas uma imagem de paciente do sexo masculino.

Durante reuniões com pesquisadores da área de Odontologia Forense (por meio, da parceria técnico-científica estabelecida entre o Programa de Pós-Graduação em Engenharia Elétrica), ficou definido que a organização das imagens em subconjuntos etários seria a estratégia mais adequada. Considerando que as técnicas tradicionais de estimativa de idade têm melhores resultados considerado faixas etárias específicas. Essa divisão visa também refletir marcos relevantes tanto para o desenvolvimento dentário quanto para as implicações legais. Por exemplo, para o uso banco de imagens no treinamento de redes, presume-se que, com o agrupamento, as redes tenham maior capacidade de extrair características que distingue-se um grupo do outro. A classificação foi definida da seguinte forma:

- Crianças e adolescentes: 1 a 14 anos
- Jovens adultos: 15 a 25 anos
- Adultos: 26 anos ou mais

A divisão dos grupos etários em 1 a 14, 15 a 25 e 26 a 91 anos foi fundamentada com base nos marcos fisiológicos do desenvolvimento dentário. Entre 1 e 14 anos, os indivíduos apresentam intensa atividade de formação, erupção e mineralização dos dentes, o que permite a aplicação precisa de métodos como *Nolla* e *Demirjian*, mais eficazes para crianças e adolescentes (LUZ *et al.*, 2019).

Já a faixa dos 15 aos 25 anos abrange uma fase de transição entre o final do desenvolvimento dentário e o início das modificações regressivas. É nessa faixa que se concentram diversas demandas legais, como a determinação da maioridade penal e civil, além de ser um intervalo no qual o terceiro molar ainda pode estar em formação (JUNIOR; MOURA, 2014). Por isso, técnicas como o índice de maturação do terceiro molar (I3M), proposto por *Cameriere*, têm ampla aplicação (CORREIA *et al.*, 2020).

Por fim, a partir dos 26 anos, os métodos baseiam-se sobretudo em alterações degenerativas, como a deposição de dentina secundária, aposição de cimento, atrição, translucidez radicular e medidas pulpares. Características contempladas em métodos como *Gustafson*, *Kvaal* e *Olze* (SOUSA *et al.*, 2023; DEZEM *et al.*, 2021). Nessa fase, torna-se mais difícil estimar a idade de um indivíduo com base em técnicas de crescimento e desenvolvimento, pois as estruturas anatômicas já atingiram seu estágio final de maturação. Dessa forma, a estimativa de idade em adultos requer o uso de métodos específicos.

Apesar da concentração entre adolescentes e adultos jovens, o banco apresenta diversidade etária significativa, abrangendo desde crianças com 1 ano até idosos com mais de 90 anos. Observa-se uma predominância de amostras nas faixas etárias entre 15 e 25 anos. A distribuição, no entanto, não é uniforme, apresentando um decréscimo gradual na quantidade de amostras a partir dos 30 anos.

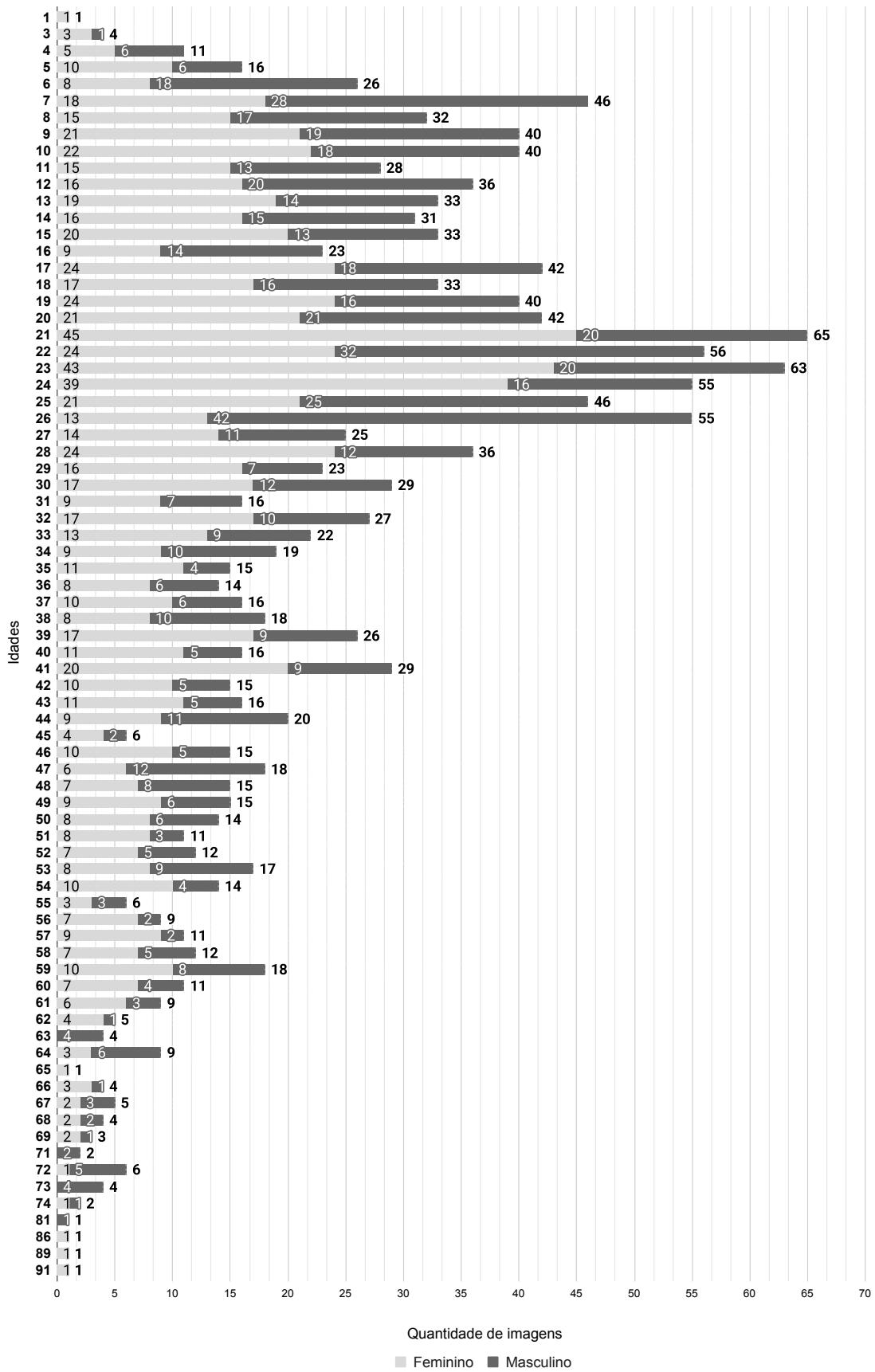


Figura 17 – Distribuição de imagens por idade e sexo

A Figura 18 apresenta a divisão do banco de imagens em três grupos etários: 1–14 anos, 15–25 anos e 26–91 anos. Nota-se que a faixa de 26 a 91 anos representa o maior volume de imagens disponíveis com 704 imagens, porém é o agrupamento de maior intervalo de idade com 66 idades diferentes. Já o segundo grupo, de 15 a 25 anos, tem proporcionalmente mais imagens com o total de 498 imagens com 10 idades. O menor grupo, em termos numéricos, é o de 1 a 14 anos, com 343 imagens, distribuídas em 14 idades. A maior concentração de imagens no grupo de 15 a 25 está em conformidade com a literatura relacionada. Estudos mencionam que nessa fase existe uma prevalência de tratamentos ortodônticos e avaliações relacionadas à formação dos terceiros molares, o que leva à maior solicitação de radiografias panorâmicas em jovens no final da adolescência, no início da vida adulta e também em adultos (FISCHBORN *et al.*, 2021). Além disso, em crianças, a indicação da radiografia panorâmica deve ser feita com parcimônia, uma vez que o tecido biológico infantil é mais sensível à radiação. Nessas situações, o exame é recomendado apenas diante de sinais clínicos relevantes (OLIVEIRA *et al.*, 2006).

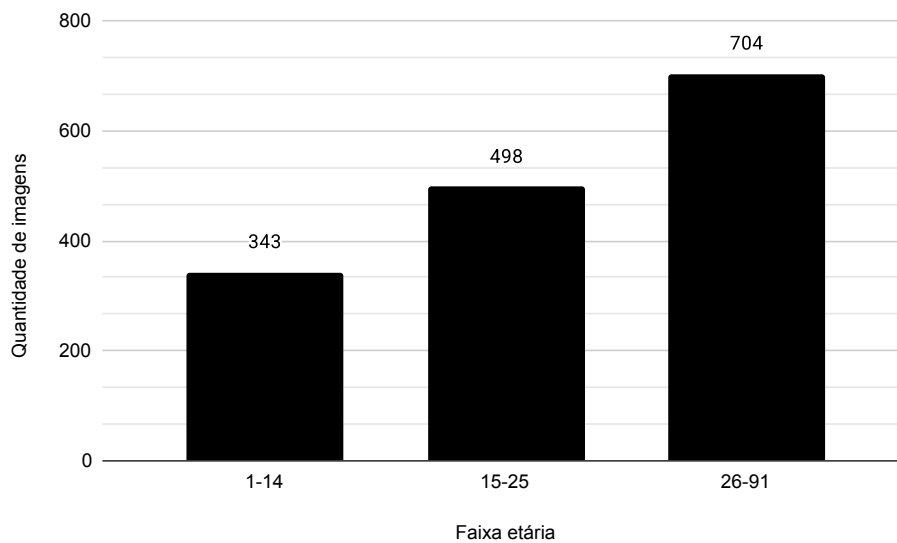


Figura 18 – Quantidade de imagens por faixa etária

A Figura 19 apresenta o quantitativo de imagens distribuídos entre os três grupos etários: 1-14, 15-25 e 26-91 anos, organizados nos sexos: feminino e masculino. Para ambos grupos etários, o sexo feminino apresenta mais imagens em comparação ao sexo masculino. Por exemplo, no grupo de 26 a 91 anos são 422 para o sexo feminino e 282 para o sexo masculino. Observa-se uma leve predominância de imagens do sexo feminino que representam cerca de 56% do total ($n = 876$) em relação ao masculino, em comparação com cerca de 43% de imagens do sexo masculino ($n = 669$). Essa diferença, embora não seja extrema, foi mantida na

composição do banco sem aplicação de técnicas de balanceamento neste estágio, uma vez que o objetivo principal nesta fase, foi consolidar o volume total de dados disponíveis a partir das coletas realizadas.

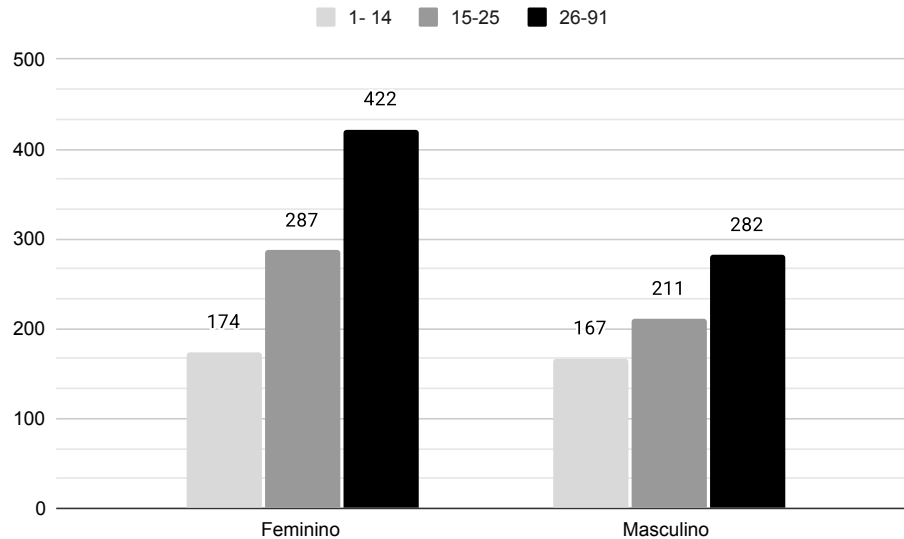


Figura 19 – Agrupamento das imagens por faixa etária e sexo

A Figura 20 apresenta o *boxplot* da variável idade, com base nas 1.545 imagens que compõem o banco de dados. Esse gráfico fornece uma visão da distribuição etária e permite identificar tendências centrais, dispersão e possíveis valores extremos no conjunto de dados. Visualmente, observa-se uma concentração das idades no intervalo entre aproximadamente 7 e 40 anos, com a mediana situada em torno dos 21 anos. A caixa interquartil (Q1 a Q3) está relativamente estreita, indicando que 50% das amostras estão concentradas em um intervalo reduzido de idade, o que reflete a densidade do banco em torno das faixas mais jovens. A presença de valores acima de 60 anos posicionados como *outliers* confirma a menor frequência de amostras em faixas etárias mais avançadas.

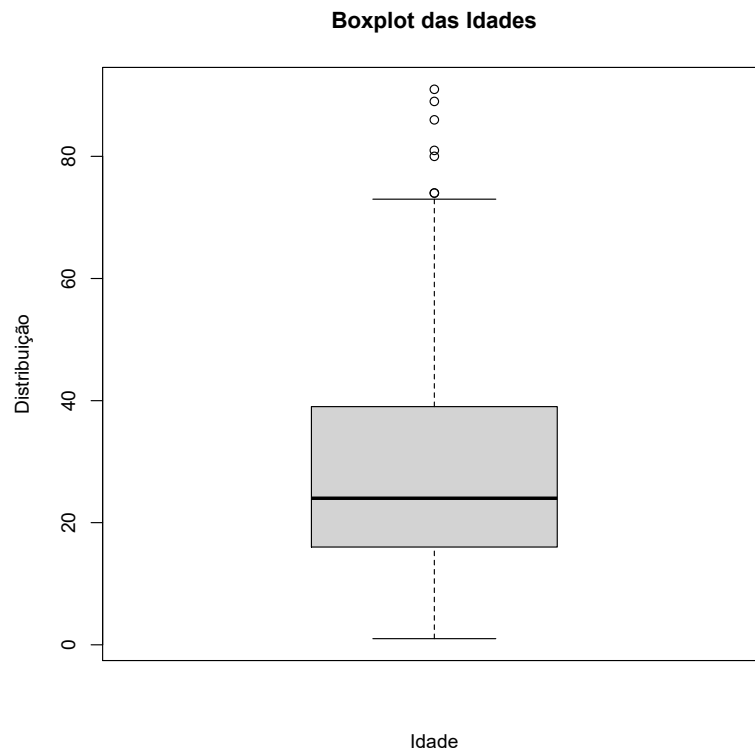


Figura 20 – Distribuição de imagens

A extensão da linha de dispersão inferior até aproximadamente 5 anos reforça a presença de dados desde a infância, embora em menor número. Já a linha de dispersão superior, que se estende até cerca de 50 anos, revela uma dispersão gradual das idades até a vida adulta, com distribuição assimétrica à direita, ou seja, com cauda longa nas maiores idades.

4.4.1 Identificação dos Dentes na Arcada Dentária

Por meio de ajuda de alunos do curso de Odontologia da Universidade Federal do Ceará, Campus de Sobral, foram analisadas as imagens de radiografia odontológica panorâmica deste banco de imagens. Realizou-se a contagem dos dentes permanentes identificados nas radiografias, bem como a identificação de cada dente presente nas arcadas dentárias. Assim, está

disponibilizada no banco de imagens uma planilha que constam os campos apresentados nas Tabelas 6 e 7. Por questões de espaço no texto, as tabelas estão apresentadas separadamente, porém, os dados referentes a estas tabelas correspondem a uma única radiografia.

Como exemplo as Tabelas 6 e 7 apresentam uma instância do banco de imagens. Como pode ser observado na Tabela 6 a numeração da imagem escolhida como exemplo é 27_60A_M, paciente de 60 anos de idade, do sexo Masculino e com a radiografia odontológica apresentando 25 dentes permanentes. A Tabela 7 apresenta a identificação dos dentes segundo a Nomenclatura Dental pela Federação Dentária Internacional, conforme apresentada na Seção 2.2. Desta forma, a primeira linha da tabela, identifica as regiões da arcada dentária (parte superior direita, parte superior esquerda, parte inferior direita e parte inferior esquerda). A segunda linha da tabela apresenta a numeração dos dentes permanentes, que são de 18 a 11 (parte superior direita), 21 a 28 (parte superior esquerda), 48 a 41 (parte inferior direita) e 31 a 38 (parte inferior esquerda). A terceira linha da Tabela 7 indica a presença do dente (1), não presença (0), dente incluso na gengiva (dente que não conseguiu nascer na posição correta, ficando retido ou impactado) (2), resto radicular (apenas a raiz do dente) (3), implante (4) e traumatismo dental (dente quebrado) (5). Como por exemplo para a radiografia 27_60A_M apresentada na Tabela 7, os dentes 15, 14, 48, 45, 41, 35 e 38 são ausentes, o dentes 23 está incluso.

Tabela 6 – Identificação da imagem

Numeração da imagem	Idade	Sexo	Quantidade de dentes
27_60A_M	60	M	25

Tabela 7 – Identificação dos dentes na imagem

Parte superior direita								Parte superior esquerda								Parte inferior direita								Parte inferior esquerda							
18	17	16	15	14	13	12	11	21	22	23	24	25	26	27	28	48	47	46	45	44	43	42	41	31	32	33	34	35	36	37	38
1	1	1	0	0	1	1	1	1	1	2	1	1	1	1	1	0	1	1	0	1	1	1	0	1	1	1	1	0	1	1	0

As Figuras 21 e 22 apresentam o quantitativo percentual da condição dos dentes nas arcadas dentárias em relação ao dentes ser qualificado como está presente, ausente, incluso, raiz, implante ou traumatismo dental. Desse modo, a Figura 21 apresenta a distribuição percentual das condições dos dentes na parte superior da arcada dentária, conforme observada nas imagens que compõem o banco de dados utilizado neste estudo. A análise foi feita individualmente para cada dente, numerados de 18 a 28, abrangendo as regiões superior direita (dentes 18 a 11) e superior esquerda (dentes 21 a 28) (ver a Seção 2.2).

Ainda, a cor verde indica a presença do dente, sendo o padrão mais prevalente em quase todos os casos, com destaque para os dentes centrais (11 e 21), que apresentam taxas de presença superiores a 95%. Já os terceiros molares (dentes 18 e 28) apresentam as menores taxas de presença, com 55% e 54% respectivamente, e as maiores taxas de ausência, com 38% e 39%.

As ausências (em azul claro) são mais significativas nos dentes 18, 17, 28 e 27. Já os casos de inclusão (em lilás), embora pouco frequentes, aparecem nos dentes 28 e 18. As categorias menos prevalentes são: traumatismo dental (vermelho), implante (cinza) e raiz remanescente (amarelo), que aparecem com porcentagens muito pequenas, inferiores a 3% em todos os dentes.

A Figura 22 apresenta a distribuição percentual das condições dos dentes na parte inferior da arcada dentária, de acordo com as imagens analisadas no banco de dados deste estudo. Os dentes estão numerados de 48 a 38, abrangendo as regiões inferior direita (dentes 48 a 41) e inferior esquerda (dentes 31 a 28) (ver a Seção 2.2). Assim como na arcada superior, observa-se que a presença do dente (em verde) é a condição predominante em praticamente todos os dentes, com destaque para os dentes 44 a 34, cuja presença é maior ou igual a 94%. O dente 22, por exemplo, apresenta 100% de presença.

Por outro lado, os terceiros molares inferiores (dentes 48 e 38) novamente se destacam por apresentarem as maiores taxas de ausência, com 40% e 41%, respectivamente. Além disso, esses dentes possuem as menores taxas de presença da arcada inferior (53% para o 48 e 52% para o 38), evidenciando um padrão semelhante ao observado na arcada superior. A porcentagem de dentes inclusos (em lilás) permanece baixa, com registros pontuais nos dentes 48 e 38. Já os casos de raiz remanescente (em amarelo), implante (cinza) e traumatismo dental (vermelho) são raros e aparecem em proporções muito pequenas (geralmente inferiores a 1%).

Dessa forma, as análises das Figuras 21 e 22 indicam que a maior parte das imagens do banco representa arcadas dentárias com estrutura dentária preservada ou com ausências simples, sem muitos casos de intervenções ou de traumas. Além disto, elas permitem compreender a distribuição anatômica e as condições dentárias mais comuns em cada região da boca dessa população. Essa análise é relevante tanto para fins de estudos clínicos quanto para o treinamento e validação de modelos de inteligência artificial voltados à análise de imagens odontológicas.

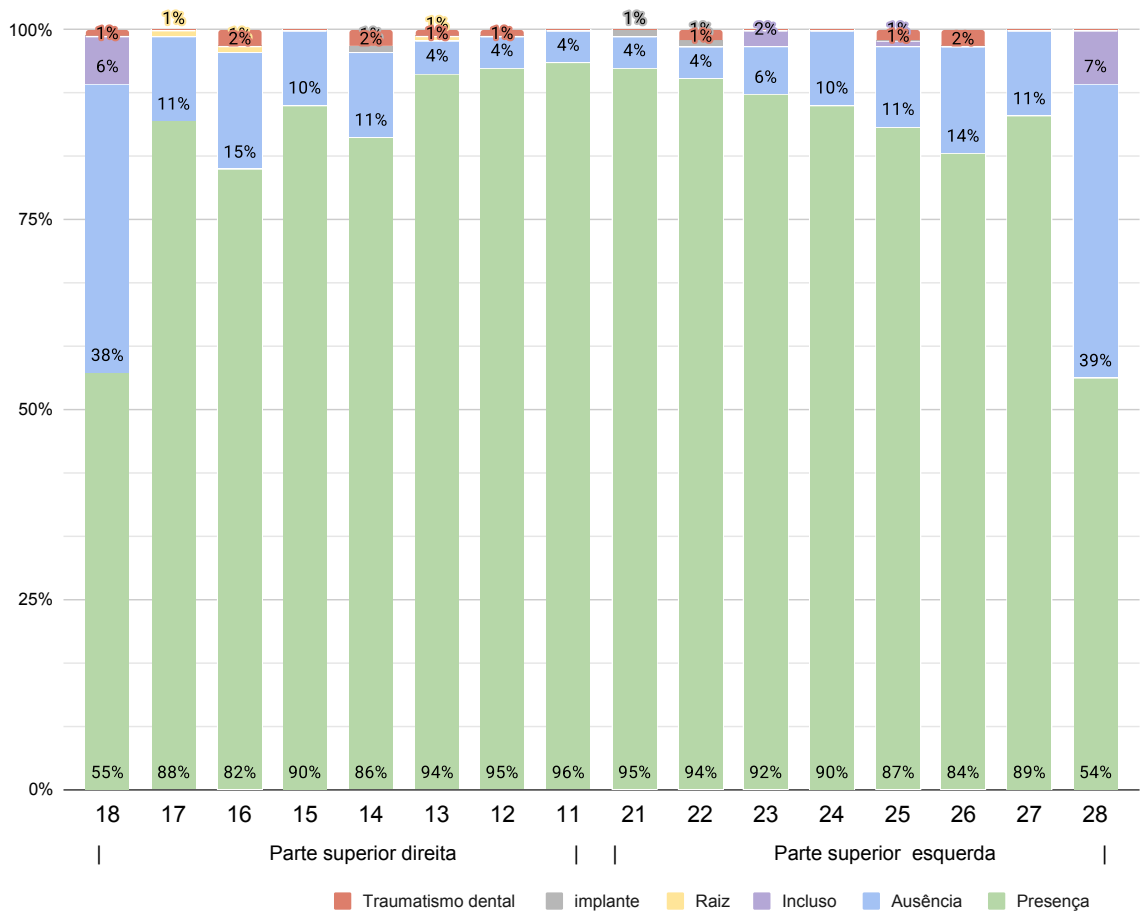


Figura 21 – Porcentagem do dentes parte superior da arcada dentária

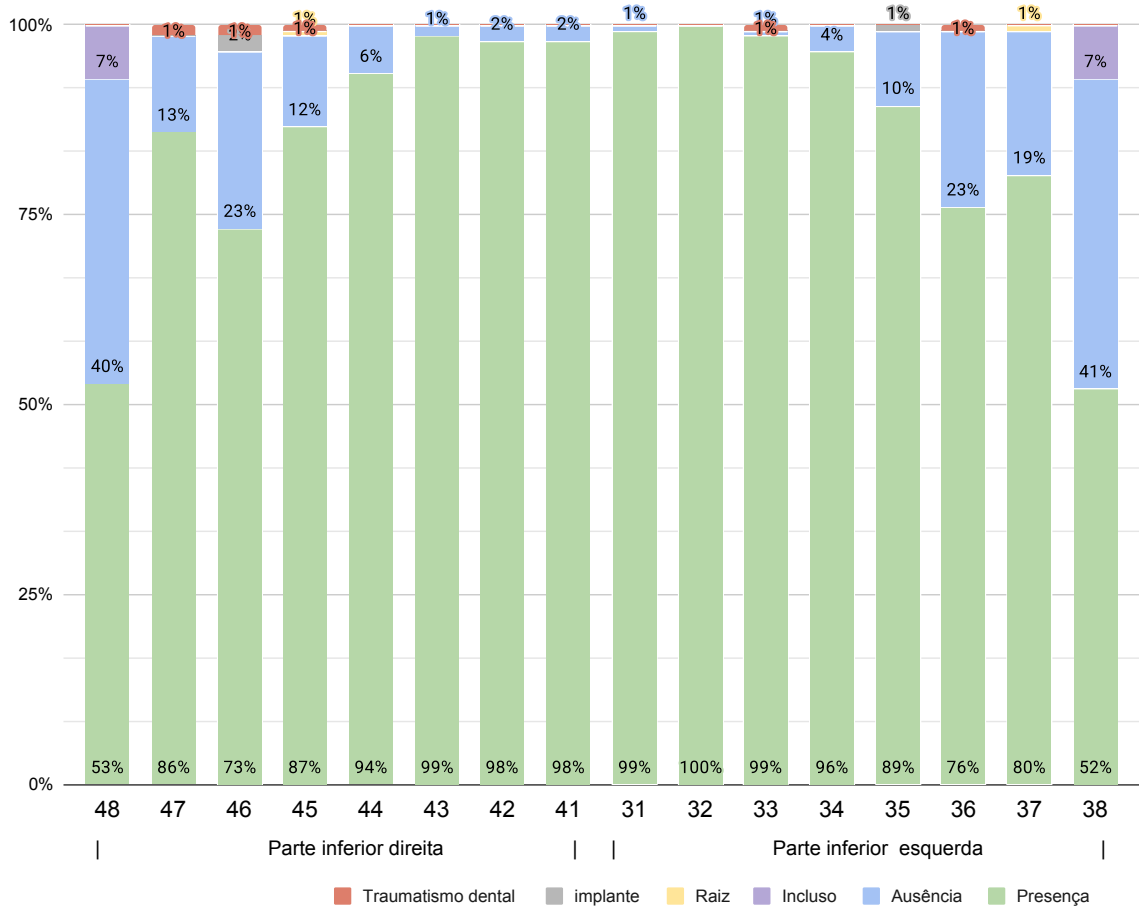


Figura 22 – Porcentagem do dentes parte inferior da arcada dentária

A Figura 23 apresenta a distribuição de dentes presentes nas arcadas dentárias das amostras do banco de imagens. Como pode ser observado na figura, a mediana é de 30 dentes. Ainda, pode ser observado alguns *outliers*, que representam pacientes que possuem um número reduzido de dentes. Como critério de exclusão, foram desconsideradas arcadas com menos de seis dentes; o boxplot mostra os *outliers* começando com seis e até aproximadamente 20 dentes.

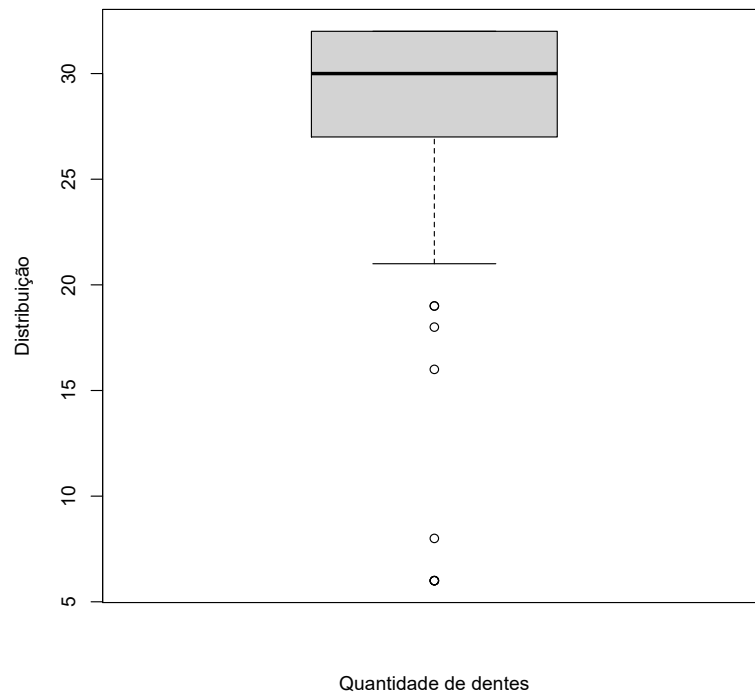


Figura 23 – Distribuição de imagens por quantidade de dentes nas arcadas dentárias

4.5 Exemplo de Uso do Banco de Imagens

Para exemplificar uma das aplicações práticas do banco de imagens construído neste estudo, esta seção apresenta um caso de uso voltado à classificação do sexo do paciente com base em radiografias panorâmicas odontológicas. Essa tarefa se insere no contexto da biometria médica, na qual modelos de aprendizado profundo são treinados para identificar padrões morfológicos associados ao sexo biológico, como o desenvolvimento mandibular, espessura óssea e proporções craniofaciais visíveis na imagem radiográfica.

4.5.1 Organização dos Dados e Pré-Processamento

Para o treinamento do modelo, as 1545 imagens foram rotuladas de acordo com o sexo do paciente, feminino ou masculino, com base nos dados fornecidos pelas clínicas. Os dados foram divididos da seguinte forma:

- Treinamento: 70% das imagens
- Validação: 15% das imagens
- Teste: 15% das imagens

A divisão foi realizada de forma estratificada para manter a proporção entre as classes nos diferentes subconjuntos. Antes do treinamento, todas as imagens foram convertidas para escala de cinza, redimensionadas para 224x224 pixels (tamanho compatível com redes pré-treinadas).

4.5.2 Arquitetura do Modelo de Rede Neural Utilizada

A arquitetura escolhida para este experimento foi a *MobileNetV2*, conhecida por sua eficiência computacional e bom desempenho em tarefas de classificação com conjuntos de dados limitados. A rede foi adaptada para um problema de classificação binária, com uma camada de saída contendo um único neurônio ativado por função sigmoide, que retorna a probabilidade da imagem pertencer à classe “Feminino”.

Foi utilizada *transfer learning*, com pesos iniciais da *MobileNetV2* treinada no *ImageNet*. As últimas camadas da rede foram ajustadas com os dados do banco de imagens por meio do processo de *fine-tuning*.

4.5.3 Pseudo Código Utilizado no Experimento

O Código-fonte 1 apresenta um pseudocódigo que descreve, de forma estruturada, o processo desenvolvido para a classificação automática do sexo dos pacientes a partir de radiografias panorâmicas odontológicas. Esse algoritmo foi implementado no ambiente *Google Colab* utilizando a arquitetura *MobileNetV2*, com o suporte da biblioteca *TensorFlow* e técnicas de *transfer learning*.

Nas Linhas 1 e 3 do código-fonte 1 realiza-se a montagem do *Google Drive* e o carregamento de todas as imagens da pasta *Dataset_FINAL*. Em seguida, cada arquivo é analisado para extrair as informações de idade e sexo, que são utilizadas como rótulos (Linha 3). Essas

informações são organizadas em um *DataFrame* que serve como base para o particionamento dos dados.

Para garantir uma avaliação justa do desempenho do modelo, foi adotada uma estratégia de balanceamento do conjunto de teste, com igual número de amostras masculinas e femininas. Dessa forma, na Linha 4, é realizada uma etapa de balanceamento do conjunto de teste. São selecionadas amostras com proporção igual entre as classes “Masculino” e “Feminino”, com o objetivo de evitar viés na avaliação do modelo e garantir uma análise justa da performance. Esse conjunto é separado antes da divisão em treino e validação.

As imagens restantes foram divididas em conjuntos de treino e validação de forma estratificada, respeitando a proporção original entre as classes. O conjunto de treinamento passou por uma etapa de *data augmentation*, com o objetivo de aumentar a robustez do modelo frente a variações naturais das imagens. Desta forma, as Linhas 5 e 6 descrevem o particionamento estratificado do restante dos dados em conjuntos de treino e validação, e a aplicação de *data augmentation* para o treinamento. Técnicas como rotação, deslocamento e ajuste de brilho foram empregadas com o objetivo de aumentar a generalização do modelo e reduzir o risco de *overfitting*.

A rede neural utilizada foi baseada na *MobileNetV2*, pré-treinada com pesos do *ImageNet*. A base da rede foi inicialmente congelada, sendo adicionadas camadas densas específicas para a tarefa de classificação binária (masculino/feminino). O treinamento foi realizado em duas fases: na primeira, com a base congelada, e na segunda, com as últimas camadas da rede base descongeladas para *fine-tuning*, utilizando uma taxa de aprendizado reduzida. Desta forma, na Linha 7, é carregada a arquitetura *MobileNetV2* com pesos pré-treinados no *ImageNet*. A estrutura original da rede é adaptada com a inclusão de camadas densas específicas para a tarefa binária de classificação de sexo (Linha 7), e as camadas da base são inicialmente congeladas (Linha 8) para preservar os conhecimentos previamente adquiridos durante o treinamento original.

A compilação do modelo com a função de perda *binary_crossentropy* e o otimizador Adam é descrita na linha 9, utilizando pesos de classe (Linha 10) para compensar possíveis desbalanceamentos no conjunto de treino, mesmo após o balanceamento do teste.

O treinamento é dividido em duas fases: na Linha 11, realiza-se o primeiro treinamento com a base congelada. Em seguida, na Linha 12, as últimas camadas da *MobileNetV2* são descongeladas para o processo de *fine-tuning*, refinando a rede para o domínio odontológico com

uma menor taxa de aprendizado (Linha 13). O segundo treinamento ocorre na Linha 14, agora com ajustes mais finos dos pesos da rede.

Por fim, nas Linhas 15 a 17, o modelo é avaliado no conjunto de teste balanceado, com a geração de métricas como precisão, *recall* e *F1-score*, além da matriz de confusão, que permite a análise visual dos acertos e erros do modelo em cada classe. Esse fluxo de trabalho busca garantir rigor metodológico, permitindo a reprodutibilidade do experimento e a transparência na avaliação do desempenho do modelo.

Código-fonte 1 – Pseudocódigo do experimento de classificação de sexo

```

1  Entrada: Imagens rotuladas com idade e sexo, organizadas em uma pasta do
    Google Drive
2  1. Montar o Google Drive
3  2. Carregar todas as imagens da pasta Dataset_FINAL
4  3. Para cada nome de arquivo:
5      - Extrair número da imagem, idade e sexo (M ou F)
6      - Armazenar caminho da imagem e rótulo em um DataFrame
7  4. Balancear o conjunto de teste:
8      - Selecionar a mesma quantidade de imagens para cada sexo (Feminino
    e Masculino)
9  5. Dividir o restante em treino e validação, mantendo a proporção de
    classes (estratificado)
10 6. Criar geradores de imagens:
11    - Aplicar data augmentation no treino
12    - Normalizar todas as imagens (rescale = 1/255)
13 7. Carregar o modelo base MobileNetV2 pré-treinado
14    - Remover camada de classificação original
15    - Adicionar camada densa de 128 neurônios (ReLU)
16    - Adicionar camada de saída com ativação sigmoide (1 neurônio)
17 8. Congelar as camadas da MobileNetV2
18 9. Compilar o modelo com otimizador Adam e loss binário
19 10. Calcular pesos de classe para compensar o desbalanceamento no treino
20 11. Treinar o modelo por 10 épocas com a base congelada
21 12. Descongelar as últimas 30 camadas da MobileNetV2
22 13. Recompilar o modelo com taxa de aprendizado menor (1e-5)
23 14. Continuar o treinamento por mais 10 épocas (fine-tuning)
24 15. Avaliar o modelo no conjunto de teste:
25    - Gerar métricas: acurácia, precisão, recall, F1-score
26    - Construir matriz de confusão
27 Saída: Modelo treinado e resultados de avaliação no conjunto de teste
    balanceado

```

4.5.4 Resultados e Análise

O modelo foi treinado por 30 épocas utilizando a função de perda *binary cross-entropy* e o otimizador Adam com taxa de aprendizado de 10^{-4} . A Tabela 8 apresenta os resultados

Tabela 8 – Métricas de desempenho para classificação do sexo a partir de radiografias panorâmicas

Classe	Precisão	Recall	F1-score	Suporte
Masculino	0,83	0,90	0,86	231
Feminino	0,89	0,82	0,85	231
Acurácia			0,86	462

de desempenho do modelo. A precisão indica que 89% das amostras classificadas como “Feminino” estavam corretas, enquanto 83% das classificadas como “Masculino” também estavam corretas. Em relação ao recall, que representa a proporção de amostras de cada classe corretamente identificadas pelo modelo, 90% dos indivíduos do sexo masculino foram corretamente classificados, assim como 82% dos indivíduos do sexo feminino. O F1-score, que corresponde à média harmônica entre precisão e recall, foi de 86% para a classe masculina e 85% para a classe feminina, indicando um bom equilíbrio entre as métricas. O suporte, que corresponde à quantidade de amostras reais de cada classe no conjunto de teste, foi balanceado com 231 imagens para cada sexo. A acurácia geral do modelo, considerando o total de acertos, foi de 86%.

Esses valores indicam que o modelo foi capaz de aprender características morfológicas associadas ao sexo com alta taxa de acerto. A Figura 24 apresenta a matriz de confusão obtida no conjunto de teste.

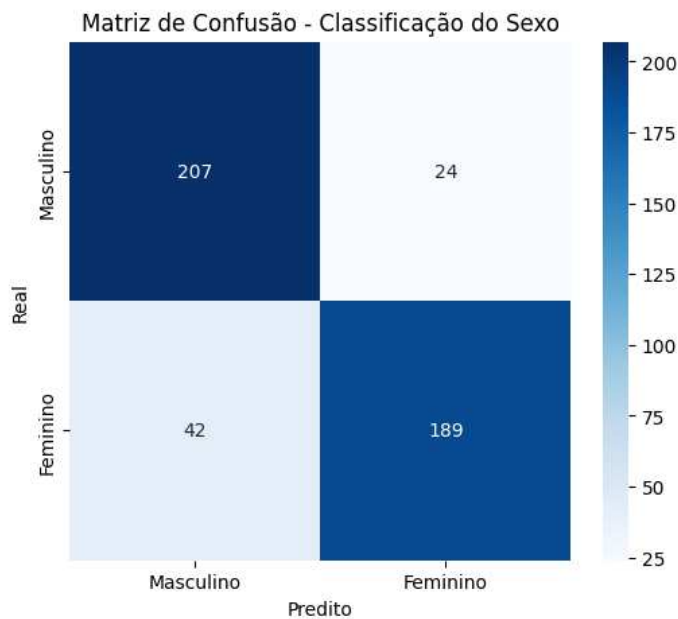


Figura 24 – Matriz de confusão para predição do sexo no conjunto de teste

Observa-se que a maior parte dos erros ocorreu na classificação de imagens do

sexo feminino, o que pode estar relacionado a maior variabilidade morfológica nas amostras femininas, sobreposição de características ou uso do *class weightin*. O modelo pode ter sido incentivado a dar mais atenção à classe minoritária, melhorando seu desempenho às custas da classe majoritária. Em estudos futuros, com ampliação do banco de imagens, o uso de técnicas de balanceamento poderá melhorar esse aspecto.

4.5.5 *Código e Reprodutibilidade*

Todo o código referente a esse experimento está disponível no repositório do banco de imagens⁵, incluindo:

- Scripts de carregamento e organização dos dados por sexo;
- Pipeline de pré-processamento das imagens;
- Implementação do modelo MobileNetV2 com ajuste para classificação binária;
- Códigos para treinamento, avaliação e geração de métricas de desempenho.

Este exemplo embora seja simples, demonstra o potencial do banco de imagens para aplicação em tarefas de classificação biomédica, sendo útil não apenas na estimativa de idade, mas também na predição de atributos demográficos como o sexo, e poderá contribuir para avanços na odontologia forense e na pesquisa em inteligência artificial aplicada à saúde.

4.6 Ameaças a Validade do Conjunto de Imagens

Apesar do planejamento criterioso e da adoção de um protocolo específico para a construção do banco de imagens, é importante reconhecer que sua validade pode ser impactada por diversos fatores. A seguir, são discutidos esses possíveis comprometimentos e as principais decisões adotadas com o objetivo de mitigar tais riscos, assegurando maior confiabilidade às pesquisas que venham a utilizar esse banco como base.

Uma das principais ameaças à validade está relacionada à representatividade da amostra. Como a base foi composta a partir de parcerias com clínicas específicas, existe o risco de viés regional, socioeconômico ou étnico, uma vez que os pacientes atendidos por essas instituições podem não refletir com precisão a diversidade populacional brasileira. Isso pode limitar a generalização dos resultados para outras populações. Para mitigar o problema de representatividade, as imagens foram coletadas em três clínicas diferentes em dois estados distintos.

⁵ <<https://rsl-redes-neurais-profundas.github.io/rsl/>>

Além disso, há riscos associados à qualidade técnica das radiografias. Embora tenham sido excluídas imagens com baixa resolução ou que arcada dentária estivesse cortada, pequenas variações na posição do paciente, no contraste ou definição da imagem podem interferir na acurácia dos modelos de aprendizado profundo. Ainda, foram tomados alguns cuidados na seleção das imagens, por exemplo, radiografias que apresentavam menos que seis dentes não foram incluídas, a fim de garantir uma mínima representatividade anatômica da arcada dentária. Esses fatores, embora muitas vezes sutis, podem afetar o desempenho das redes neurais treinadas com esse banco. Desta forma, como a seleção das imagens foi manual, todas as imagens que apresentavam algum desses problemas foram excluídas previamente.

Outro ponto relevante diz respeito à exatidão dos rótulos utilizados, em especial da idade dos pacientes. Essa informação foi extraída a partir do nome dos arquivos ou fornecida diretamente pelas clínicas, podendo estar sujeita a erros de digitação ou desatualizações, sobretudo nos casos em que as imagens não são recentes. A ausência de uma verificação cruzada com prontuários médicos representa uma potencial ameaça à validade dos rótulos de idade. Para mitigar esse risco, todos os dados disponíveis dos pacientes, embora não tenham sido coletados de forma sistemática, foram cuidadosamente analisados com o objetivo de identificar possíveis inconsistências relacionadas às informações de interesse.

Por fim, há ainda a possibilidade de duplicidade de imagens caso um mesmo paciente tenha realizado mais de uma radiografia em diferentes datas, o que poderia influenciar os dados de treino e teste, comprometendo a independência das amostras. Para mitigar esse problema todos os dados relativos às imagens foram inseridos em um arquivo *csv* para evitar possíveis duplicação de dados.

Essas limitações não invalidam o banco construído, mas devem ser consideradas tanto na interpretação dos resultados quanto em futuras expansões ou refinamentos da base de dados.

4.7 Consideração Finais

Este capítulo apresentou detalhadamente o processo de construção do banco de imagens utilizado neste estudo, descrevendo as etapas de aquisição, preparação, qualificação, organização e disponibilização das radiografias panorâmicas odontológicas. Ao todo, foram reunidas 1545 imagens provenientes de três clínicas localizadas em dois estados distintos do Brasil, representando uma diversidade etária ampla e informações demográficas relevantes, como

idade e sexo.

A preparação criteriosa incluiu a anonimização dos dados, padronização visual com aplicação de máscara de região de interesse e organização estruturada em faixas etárias e categorias de sexo, com o objetivo de garantir qualidade, consistência e facilidade de uso. Todas as imagens foram submetidas a uma etapa de qualificação, na qual foram avaliadas quanto à nitidez, presença anatômica mínima e conformidade técnica para uso em modelos computacionais.

Como uma das principais contribuições deste trabalho, destaca-se a disponibilização pública de um banco de imagens inédito, voltado à estimativa de idade com redes neurais profundas, mas aplicável a diversos outros propósitos em pesquisa odontológica e biomédica. Este repositório representa um avanço para a comunidade científica, especialmente considerando a escassez de bases públicas de radiografias panorâmicas no contexto nacional.

O próximo capítulo apresenta um estudo exploratório com sete redes neurais profundas. O estudo foi projetado e executado utilizando o banco de imagens apresentado neste capítulo.

5 ESTUDO EXPLORATÓRIO

Este capítulo apresenta um estudo exploratório para estimativa de idade considerando três grupos de faixas etárias (1-14, 15-25 e 26-96 anos) utilizando um subgrupo de nove redes neurais profundas identificadas na revisão sistemática da literatura, discutida no Capítulo 3. Na revisão sistemática da literatura foram encontradas 25 redes, mas por questões de limitações de recursos computacionais (unidades de computação do Google Colab) foi conduzido o estudo com nove redes.

Este capítulo está organizado da seguinte forma: a Seção 5.1 apresenta as configurações adotadas no estudo. A Seção 5.2 apresenta os resultados alcançados por meio deste estudo exploratório e discussões relacionadas os resultados alcançados. Ainda, a Seção 5.3 apresenta algumas implicações deste estudo. Finalmente, as Seções 5.4 e 5.5 discutem as ameaças a validade deste estudo e as considerações finais, respectivamente.

5.1 Configuração do Estudo

Este estudo foi estruturado em duas etapas principais. A primeira envolveu a classificação de imagens em três faixas etárias (1–14, 15–25 e 26–91 anos), utilizou-se para treinamento, validação e teste o banco de dados organizado e apresentado no Capítulo 4. A segunda etapa corresponde à validação externa, conduzida com um conjunto de dados independente. A Figura 25 ilustra o fluxo de trabalho proposto, organizado em quatro etapas: ARTEFATOS DO ESTUDO, CLASSIFICAÇÃO, VALIDAÇÃO EXTERNA e MÉTRICAS DE AVALIAÇÃO.

A etapa ARTEFATOS DO ESTUDO apresenta os principais recursos utilizados: o banco de imagens interno (Capítulo 4) e um segundo banco (OLIVEIRA *et al.*, 2024), identificado posteriormente à revisão sistemática descrita no Capítulo 3. Nesta etapa também está detalhado o ambiente computacional empregado.

A etapa de CLASSIFICAÇÃO representa a primeira fase do experimento. Nessa fase, o banco de dados apresentado no Capítulo 4 foi utilizado para treinar, validar e testar nove arquiteturas de redes neurais convolucionais. O fluxo metodológico dessa etapa seguiu as seguintes fases: (i) uso do banco de dados interno, (ii) divisão dos dados via validação cruzada do tipo *Holdout*, (iii) aplicação de técnicas de pré-processamento, (iii) aumento de dados, (iv) utilização da técnica de *Transfer Learning*, (v) treinamento, validação e teste das arquiteturas convolucionais e (vi) avaliação dos modelos por meio de métricas de avaliação:

acurácia, precisão, *recall*, *F1-score* e tempo de treinamento. Essas métricas foram utilizadas para comparar o desempenho entre as diferentes arquiteturas testadas.

Na etapa de VALIDAÇÃO EXTERNA, as três arquiteturas com melhor desempenho na primeira fase foram treinadas e validadas como o banco de dados de Oliveira *et al.* (2024). O conjunto de treino passou pelos mesmos procedimentos de pré-processamento e aumento de dados utilizados anteriormente. Em seguida, os modelos treinados foram testados com o banco de imagens interno (Capítulo 4), a fim de verificar sua robustez frente a dados distintos dos utilizados no treinamento inicial. Por fim, os resultados da validação externa foram analisados utilizando as mesmas MÉTRICAS DE AVALIAÇÃO. As seções seguintes apresentam detalhes sobre cada uma das etapas apresentadas na Figura 25.

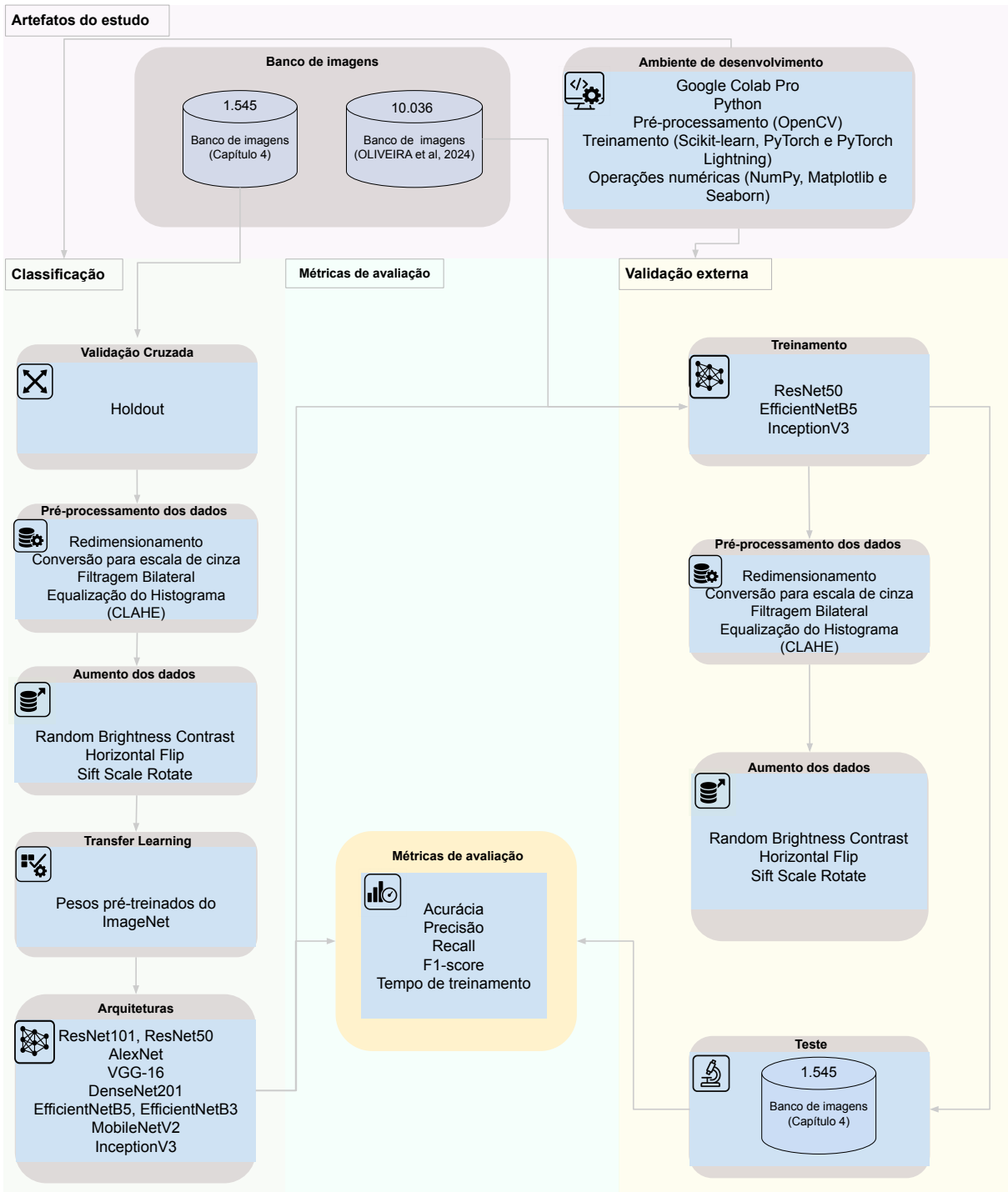


Figura 25 – Configuração do estudo

5.1.1 Base de Dados

Para compor este estudo foi usado dois bancos de imagens. Na etapa de classificação, foi utilizado o conjunto de imagens organizado especificamente para este estudo, o qual foi detalhadamente apresentado no Capítulo 4. Esse conjunto de dados foi empregado no treinamento das arquiteturas de redes neurais para a tarefa de classificação de imagens. Esta base é composta

por 1.545 imagens de radiografias panorâmicas no formato *Joint Photographic Experts Group* (JPEG) e abarca as faixas etárias de 1 a 91 anos. Para a validação externa dos modelos, foi empregada a combinação de dois conjuntos de dados distintos. Para isso, utilizou-se o conjunto de dados apresentado por Oliveira *et al.* (2024), em conjunto com a base de dados desenvolvida especificamente para este estudo.

O conjunto de dados de Oliveira *et al.* (2024) foi dividido em três faixas etárias: 1–14, 15–25 e 26–96 anos, sendo utilizado exclusivamente para o treinamento e validação dos modelos. Esse banco de imagens de radiografias panorâmicas está no formato JPEG e abarca as faixas etárias de 3 a 96 anos. Em seguida, os modelos treinados foram expostos ao banco de dados construído neste trabalho, que atuou como conjunto de teste na etapa de validação externa. A Tabela 9 apresenta um compilado com as principais informações do bancos de dados (OLIVEIRA *et al.*, 2024) incorporado neste estudo.

Tabela 9 – Distribuição das imagens por faixa etária

Grupos etários	Total de imagens
3–14	1.559
15–25	1.623
26–96	6.854
Total	10.036

5.1.2 Ambiente de Desenvolvimento

Todo o desenvolvimento e a execução dos experimentos foram realizados na plataforma *Google Colab Pro*, versão paga do ambiente colaborativo baseado em nuvem. A escolha por esse recurso foi por conta da limitação de acesso a máquinas locais com maior capacidade de processamento, sendo o *Colab Pro* uma alternativa viável e eficiente, especialmente pela disponibilidade de Unidade de Processamento Gráfico (GPU) de alto desempenho, essenciais para a execução de modelos de aprendizado profundo. Além disso, a integração direta com o *Google Drive* facilitou tanto o armazenamento e o carregamento do conjunto de dados de imagens radiográficas quanto o salvamento automático dos modelos treinados e dos resultados obtidos.

As aplicações foram desenvolvidas em *Python*¹, linguagem bastante difundida na

¹ Python Software Foundation. Disponível em: <<https://www.python.org>>

área de *deep learning* devido à sua sintaxe acessível e ao suporte de bibliotecas especializadas. O pré-processamento das imagens foi realizado com a biblioteca *OpenCV*², enquanto a construção e o treinamento dos modelos contaram com ferramentas como *Scikit-learn*³, *PyTorch*⁴ e *PyTorch Lightning*⁵. Já operações numéricas, visualizações e gráficos foram conduzidas com o auxílio do *NumPy*⁶, *Matplotlib*⁷ e *Seaborn*⁸, permitindo uma visualização gráfica dos resultados obtidos.

5.1.3 Classificação

A etapa de classificação teve como objetivo avaliar o desempenho de diferentes arquiteturas de redes neurais convolucionais na predição de atributos a partir das radiografias panorâmicas, especificamente voltadas à identificação da faixa etária dos pacientes. Nesse estudo, foram utilizados modelos CNN pré-treinados em conjuntos de dados, aplicando o conhecimento já adquirido.

Nesse contexto, foram definidos diversos hiperparâmetros para otimizar o desempenho dos modelos. A escolha dos hiperparâmetros, como taxa de aprendizado, função de ativação, otimizador e tamanho do batch, foi realizada com base em experimentações empíricas e em boas práticas recomendadas na literatura. Cada modelo foi treinado por 50 épocas, o que representa o número de iterações durante o treinamento. Foi definida a função de perda Entropia Cruzada, amplamente empregada em problemas multiclasse. Por fim, foi estabelecida uma paciência (*patience*) igual a 10, o que indica o número de épocas consecutivas em que, na ausência de melhora no desempenho do modelo, o treinamento pode ser interrompido. A Tabela 10 apresenta o conjunto de valores definidos para cada hiperparâmetro na construção dos modelos.

5.1.4 Validação Cruzada

Neste procedimento, foi adotada a estratégia de validação cruzada do tipo *Holdout*. Nessa abordagem, o conjunto de dados foi inicialmente dividido em 80% para o treinamento e 20% para teste e validação. Posteriormente, os 20% reservados foram subdivididos de forma equitativa, destinando-se 10% para a validação do modelo durante o treinamento e 10% para a

² OpenCV: Open Source Computer Vision Library. Disponível em: <<https://opencv.org>>

³ Scikit-learn: Machine Learning in Python. Disponível em: <<https://scikit-learn.org>>

⁴ PyTorch. Disponível em: <<https://pytorch.org>>

⁵ PyTorch Lightning. Disponível em: <<https://www.pytorchlightning.ai>>

⁶ NumPy: Fundamental package for scientific computing. Disponível em: <<https://numpy.org>>

⁷ Matplotlib: Visualization with Python. Disponível em: <<https://matplotlib.org>>

⁸ Seaborn: Statistical Data Visualization. Disponível em: <<https://seaborn.pydata.org>>

Tabela 10 – Parâmetros de configuração utilizados no treinamento das redes neurais

Rede	LR	Otimizador	Batch Size	WD	Épocas	Função de Perda	Patience
ResNet101	0,0005	Adam	8	0,01	50	crossentropy	10
ResNet50	0,0005	Nadam	8	0,01	50	crossentropy	10
AlexNet	0,0005	Nadam	8	0,01	50	crossentropy	10
VGG-16	0,0005	Nadam	16	0,01	50	crossentropy	10
DenseNet201	0,0005	Nadam	8	0,01	50	crossentropy	10
EfficientNet B5	0,0005	Nadam	8	0,01	50	crossentropy	10
EfficientNetB3	0,0005	Nadam	8	0,01	50	crossentropy	10
MobileNetV2	0,0005	Nadam	16	0,01	50	crossentropy	10
InceptionV3	0,001	Adam	32	0,01	50	crossentropy	5

avaliação final de desempenho. Essa divisão foi realizada de forma estratificada, a fim de garantir a representatividade das classes em todas as partições e assegurar maior robustez aos resultados obtidos. Além disso, essa estratégia busca garantir que os dados utilizados na etapa de teste permaneçam inéditos para a rede neural, evitando qualquer tipo de sobreajuste ou viés durante o treinamento.

5.1.5 Pré-Processamento dos Dados

As imagens utilizadas neste estudo passaram por etapas de pré-processamento com o objetivo de padronizar o conjunto de dados, melhorar a qualidade das informações e preparar os dados para o treinamento dos modelos de aprendizado profundo. A base de dados original foi composta por radiografias panorâmicas com diferentes resoluções, níveis de contraste e regiões anatômicas não uniformemente evidenciadas.

Inicialmente, todas as imagens foram redimensionadas para uma resolução fixa de 224×224 pixels, atendendo aos requisitos de entrada das redes convolucionais utilizadas. Em seguida, foi aplicada uma máscara para delimitação da Região de Interesse (ROI), com o intuito de isolar a área anatômica mais relevante para a tarefa preditiva, minimizando a interferência de regiões não informativas da imagem.

A conversão para escala de cinza foi realizada a fim de reduzir a complexidade computacional e eliminar redundâncias provenientes dos canais de cor. Para atenuar ruídos e preservar detalhes estruturais, foi utilizada a filtragem bilateral, que permite suavizar regiões homogêneas sem comprometer as bordas. Por fim, a técnica de CLAHE foi aplicada com o objetivo de realçar estruturas internas da imagem, otimizando o contraste local.

Essas etapas visaram tornar o conjunto de dados mais homogêneo e condizente com os padrões esperados pelos modelos, contribuindo para uma aprendizagem mais estável e com

menor risco de *overfitting*.

5.1.6 Aumento de Dados

Com o intuito de aumentar a capacidade de generalização dos modelos, ampliar a variabilidade do banco de dados e reduzir a ocorrência de *overfitting*, foram empregadas técnicas de *data augmentation*. Esse procedimento é relevante por permitir a simulação de variações comuns em ambientes clínicos reais, como pequenas alterações de posicionamento, escala e condições de iluminação.

Destaca-se que as transformações foram aplicadas exclusivamente ao conjunto de treinamento, de modo que as imagens destinadas à validação e ao teste foram mantidas inalteradas, preservando sua integridade para avaliação do desempenho dos modelos.

Foram adotadas transformações geométricas do tipo lineares afins, que mantêm a estrutura original da imagem enquanto permitem modificações como translação, rotação e redimensionamento. Especificamente, foram utilizadas alterações de brilho e contraste (*RandomBrightnessContrast*), inversão horizontal (*HorizontalFlip*), bem como mudanças de escala, rotação e deslocamento espacial (*ShiftScaleRotate*). A Tabela 11 apresenta os parâmetros configurados para cada uma dessas transformações.

Tabela 11 – Técnicas para aumento de dados

Técnica	Parâmetro
<i>RandomBrightnessContrast</i>	Limite de brilho = 0,3, Limite de contraste = 0,3, Probabilidade de aplicar a Transformada: p = 0,5
<i>HorizontalFlip</i>	Probabilidade da imagem ser invertida: p = 0,5
<i>ShiftScaleRodar</i>	Deslocamento = 0,05, Escala = 0,1, Rotação = 15, Probabilidade de aplicar a Transformada: p = 0,5

5.1.7 Transfer Learning

O treinamento de uma arquitetura CNN do zero é um processo extremamente oneroso em termos de tempo e recursos computacionais. Para mitigar essa limitação, é comum a aplicação da técnica de *Transfer Learning*. Dessa forma, utilizou-se redes com pesos pré-treinados no *ImageNet*.

5.1.8 *Arquitetura de Redes Neurais Profundas*

No Capítulo 3 de revisão sistemática da literatura foram encontradas 25 arquitetura de redes neurais profundas. Porém, devido as limitações de recursos computacionais e considerando que este trabalho tem como objetivo a realização de um estudo de classificação, não foi viável analisar todas as arquiteturas encontradas. Desta forma, foram escolhidas as nove redes que apresentaram melhores resultados segundo os achados da revisão sistemática. Assim, para realizar a tarefa classificação em faixas etárias, foram selecionadas as seguintes arquiteturas: *ResNet101*, *ResNet50*, *AlexNet*, *VGG-16*, *DenseNet201*, *EfficientNetB5*, *EfficientNetB3*, *MobileNetV2*, *InceptionV3*. As arquiteturas *EfficientNetB1* e *Sequential CNN* não foram incluídas devidos a problemas técnicos durante o experimento.

5.1.9 *Validação Externa*

Com o objetivo de verificar a capacidade de generalização dos modelos, foi realizada uma etapa de validação externa. Foi utilizado um conjunto de imagens distinto do empregado na primeira fase do estudo.

Para essa etapa, os três modelos (*ResNet50*, *EfficientNetB5* e *InceptionV3*) que apresentaram melhores resultados na primeira fase foram treinados exclusivamente com o conjunto de dados descrito por Oliveira *et al.* (2024), composto por 10.036 imagens. Esse conjunto foi previamente organizado em três faixas etárias (1–14, 15–26 e 26–99 anos), respeitando a proposta metodológica adotada neste estudo. As imagens passaram pelo mesmo pré-processamento descrito nas Seções 5.1.5 e 5.1.6. Após o treinamento, os modelos foram expostos ao banco de dados apresentado no Capítulo 4, sem qualquer tipo de pré-processamento.

Essa abordagem visa demonstrar a robustez dos modelos propostos e fornece evidências de sua aplicabilidade em cenários clínicos reais. Ou seja, nesta etapa, avaliou-se a capacidade do modelo de lidar com imagens totalmente novas e em estado bruto, simulando sua utilização em contextos reais nos quais as imagens podem variar em qualidade, contraste ou padronização.

5.1.10 *Métricas de Avaliação*

Para a análise do desempenho dos modelos de classificação, foram utilizadas métricas amplamente reconhecidas na literatura, com foco na interpretação da performance preditiva em

problemas multiclasse. Essas métricas permitiram uma avaliação mais completa dos modelos além da simples taxa de acerto, levando em consideração a distribuição entre classes e os tipos de erro cometidos.

A acurácia foi utilizada como métrica principal para indicar a proporção de previsões corretas sobre o total de amostras. No entanto, considerando os desequilíbrios entre classes, também foram computadas a precisão, o *recall* e o *F1-score*, que fornecem uma visão mais detalhada sobre o comportamento do modelo em relação a falsos positivos e falsos negativos. Além disso, também foi calculada a especificidade métrica que mede, para uma classe específica, a capacidade do modelo de identificar corretamente os casos que não pertencem a ela, sendo especialmente relevante em cenários que exigem maior rigor na separação entre categorias.

As métricas foram calculadas a partir da matriz de confusão, que organiza as previsões em categorias de verdadeiros positivos, falsos positivos, verdadeiros negativos e falsos negativos. Em todos os casos, as métricas foram obtidas com base no conjunto de teste. Foi assegurado que conjunto de teste não foi exposto ao modelo durante o treinamento. Dessa forma, pode-se garantir uma avaliação imparcial da capacidade de generalização dos modelos.

5.2 Resultados e Discussão

Esta seção apresenta e discute os resultados alcançados para as nove arquiteturas de redes utilizadas no estudos. Desta forma, são apresentados os valores das métricas e algumas observações importantes sobre as características das redes.

5.2.1 Utilização das Métricas Essenciais no Contexto de Estimativa de Idade

No contexto da classificação etária a partir de radiografias panorâmicas, a obtenção de boas taxas para as métricas de avaliação é fundamental para garantir a eficácia e a confiabilidade dos modelos desenvolvidos. As métricas de avaliação orientam a escolha da arquitetura mais adequada e também asseguram que o modelo seja generalista, preciso e aplicável em cenários reais.

Dessa forma, a acurácia indica a capacidade geral do modelo em realizar classificações corretas, considerando todas as faixas etárias avaliadas. A precisão mede a proporção de classificações corretas entre todas as previsões positivas de uma determinada classe etária. Já o *recall* avalia a capacidade do modelo em identificar corretamente todos os indivíduos que

realmente pertencem a uma determinada classe, estando diretamente relacionado à minimização de erros por omissão. O *F1-score*, por sua vez, é a média harmônica entre precisão e *recall*. No contexto deste estudo, é crucial que o modelo seja capaz não apenas de identificar corretamente os indivíduos de uma determinada classe (alto *recall*), mas também de evitar atribuições incorretas dessa classe, a indivíduos que não pertencem a ela (alta precisão). A especificidade complementa essa análise ao medir a capacidade do modelo de reconhecer corretamente os indivíduos que não pertencem a determinada faixa etária, contribuindo para evitar falsos positivos e reforçando a confiabilidade do sistema. Por fim, o tempo de treinamento de cada modelo, medido em segundos, é apresentado como critério adicional, permitindo avaliar o custo computacional associado ao uso de cada arquitetura.

5.2.2 Resultados Encontrados para as Métricas Essenciais

A Tabela 12 exibe os resultados para as métricas de avaliação das nove arquiteturas avaliadas. De maneira geral, os valores alcançados indicam que as arquiteturas avaliadas apresentam desempenho promissor na tarefa de distinção das faixas etárias, evidenciando o potencial dos modelos em classificar corretamente a classe etária à qual cada indivíduo pertence com base em radiografias panorâmicas. Dentre os modelos analisados, a *EfficientNetB5* obteve as melhores taxas gerais (em destaque na Tabela 12), com uma acurácia de teste de 86,36%, precisão de 87,47%, *recall* de 86,36%, *F1-Score* de 86,59% e especificidade de 93,03%. Ademais, as arquiteturas *InceptionV3* e *Resnet50* apresentaram desempenho semelhante em relação às métricas analisadas, na qual, a arquitetura *InceptionV3* apresentou uma acurácia de 83,77%, precisão de 84,62%, *recall* de 83,77%, *F1-Score* de 84,01% e especificidade de 91,66%. Enquanto, a rede *ResNet50* obteve uma acurácia e 83,77%, precisão de 84,97%, *recall* de 83,77%, *F1-Score* de 84,01% e especificidade de 91,73%.

Tabela 12 – Comparação de desempenho entre diferentes arquiteturas de redes neurais

Rede	Acurácia (%)	Precisão (%)	Recall (%)	F1-Score (%)	Especificidade (%)	Tempo (s)
ResNet101	80,52	81,35	80,52	80,78	90,06	1.219
ResNet50	83,77	84,97	83,77	84,01	91,73	2.911
AlexNet	77,92	81,19	77,92	78,41	88,53	3.536
VGG-16	77,92	81,05	77,92	78,51	88,53	3.575
DenseNet201	76,62	77,72	76,62	76,62	87,48	3.029
EfficientNetB5	86,36	87,47	86,36	86,59	93,05	2.122
EfficientNetB3	78,57	78,67	78,57	78,62	88,68	1.915
MobileNetV2	81,82	82,65	81,82	81,99	90,64	1.280
InceptionV3	83,77	84,62	83,77	84,01	91,66	1.506

Observando os dados apresentados na Tabela 12, verifica-se que a arquitetura *EfficientNetB5* obteve o melhor desempenho global entre os modelos avaliados. Os valores elevados das métricas de avaliação indicam que essa rede foi capaz de identificar corretamente a classe etária dos indivíduos, tanto em termos de cobertura quanto de exatidão, com baixo comprometimento entre falsos positivos e falsos negativos. Sua especificidade de 93,05% reforça a sua robustez, sugerindo uma excelente capacidade de rejeitar classificações incorretas, característica particularmente relevante no contexto deste estudo, nos qual decisões equivocadas podem ter graves implicações legais.

Além disso, a *EfficientNetB5* tem um arquitetura otimizada, que combina profundidade, largura e resolução de maneira equilibrada, por meio de uma técnica conhecida como *compound scaling*. Essa abordagem permite que a rede seja escalada de forma eficiente, preservando a performance sem incorrer em custos computacionais excessivos. Essa particularidade também confere a rede a capacidade de extração de detalhes finos, característica especialmente relevante na análise de radiografias panorâmicas, onde pequenas variações morfológicas podem ser determinantes para a classificação etária. Apesar de toda família *EfficientNet* compartilhe da mesma técnica de *compound scaling*, a *EfficientNetB3*, versão menor da B5, apresentou resultados inferiores, o que evidencia, sua capacidade de representação é mais limitada. Isso reforça a importância de selecionar o nível adequado da família *EfficientNet*, de acordo com a complexidade da tarefa e a disponibilidade de recursos computacionais.

A arquitetura *InceptionV3* se destaca por sua estratégia de processamento multiescalar, implementada por meio de módulos *Inception*, que permitem à rede capturar simultaneamente padrões em diferentes escalas espaciais. Essa capacidade é particularmente útil para imagens médicas, pois permite reconhecer estruturas anatômicas que variam em tamanho e forma ao longo das diferentes faixas etárias. Embora sua performance tenha sido ligeiramente inferior à da *EfficientNetB5*, o menor tempo de execução observado evidencia uma boa relação entre complexidade e eficiência.

A *ResNet50* também apresentou métricas competitivas, que podem ser justificadas por sua arquitetura residual, que introduz conexões de atalho entre camadas. Essas conexões mitigam o problema do desaparecimento do gradiente, comum em redes muito profundas, e permitem que a rede aprenda representações mais complexas sem degradação de desempenho. No entanto, o tempo de execução mais elevado sugere que o custo computacional associado à profundidade da rede é mais expressivo, o que pode ser um fator limitante.

Entretanto, arquiteturas como *AlexNet* e *VGG-16*, embora, tenham desempenhado um papel histórico no avanço da visão computacional, apresentam limitações na classificação etária deste estudo. A *AlexNet*, por exemplo, é uma rede relativamente rasa e com filtros de grandes dimensões, o que pode levar à perda de detalhes. A *VGG-16*, embora mais profunda, utiliza apenas convoluções 3x3 e não incorpora mecanismos de atenção ou atalhos residuais, o que pode prejudicar a capacidade de capturar padrões mais refinados sem gerar redundância. Isso justifica tanto os resultados das suas métricas mais modestas quanto o tempo de execução elevado.

A *DenseNet201*, embora tecnicamente avançada por suas conexões densas entre todas as camadas, demonstrou desempenho inferior. Uma possível justificativa é que, apesar de sua capacidade de reutilização de características e gradientes fluírem com facilidade, o grande número de conexões pode introduzir redundância e dificultar a generalização quando o conjunto de dados é limitado ou altamente específico, como no caso das radiografias panorâmicas.

Ainda, ressalta-se também o desempenho da *MobileNetV2* que apesar de intermediário, é um modelo projetado com foco em eficiência e portabilidade, sendo altamente otimizada para dispositivos móveis. Sua estrutura baseada em *depthwise separable convolutions* reduz drasticamente o número de parâmetros e o custo computacional, o que explica seu tempo de execução reduzido. Contudo, essa economia estrutural implica em uma capacidade menor de modelar padrões complexos, o que pode justificar a performance intermediária observada. Ainda assim, seus resultados são importantes, pois em contexto de limitação *hardware* ela demonstra grande potencial.

5.2.3 Resultados Encontrados para Especificidade

Para complementar a análise das métricas globais, a Figura 26 apresenta a comparação da especificidade por classe etária entre os diferentes arquiteturas avaliadas. Observa-se que a maioria das redes atingiu níveis elevados de especificidade para a classe 1–14. Esse resultado sugere que as arquiteturas foram particularmente eficazes em evitar falsos positivos na classificação de indivíduos muito jovens, o que pode estar relacionado à presença de padrões morfológicos mais distintivos nessa faixa etária.

A classe 15–25, por sua vez, apresentou os menores valores de especificidade de forma recorrente, como pode ser observado na Figura 26. Essas variações consideráveis entre as arquiteturas, indicam maior dificuldade das redes em distinguir essa faixa etária das adjacentes,

possivelmente devido à sobreposição de características morfológicas que ocorrem no período de transição entre adolescência e idade adulta. A *EfficientNetB5*, se destacou com especificidade acima de 90% para essa classe, mantendo-se como uma das arquiteturas mais equilibradas entre as métricas avaliadas.

Já na classe 26–99, observou-se uma tendência geral de desempenho consistente entre os modelos, com destaque a grande maioria do modelos apresenta especificidade acima de 90% para essa classe. Isso pode ser justificado pelo fato de que indivíduos mais velhos apresentam estruturas anatômicas mais desenvolvidas e características dentárias mais consolidadas, facilitando a correta rejeição de classificações indevidas.

Embora algumas redes tenham apresentado especificidades muito altas em uma determinada classe, seu desempenho global não necessariamente acompanha esse padrão, o que reforça a importância de se considerar a performance por categoria etária.

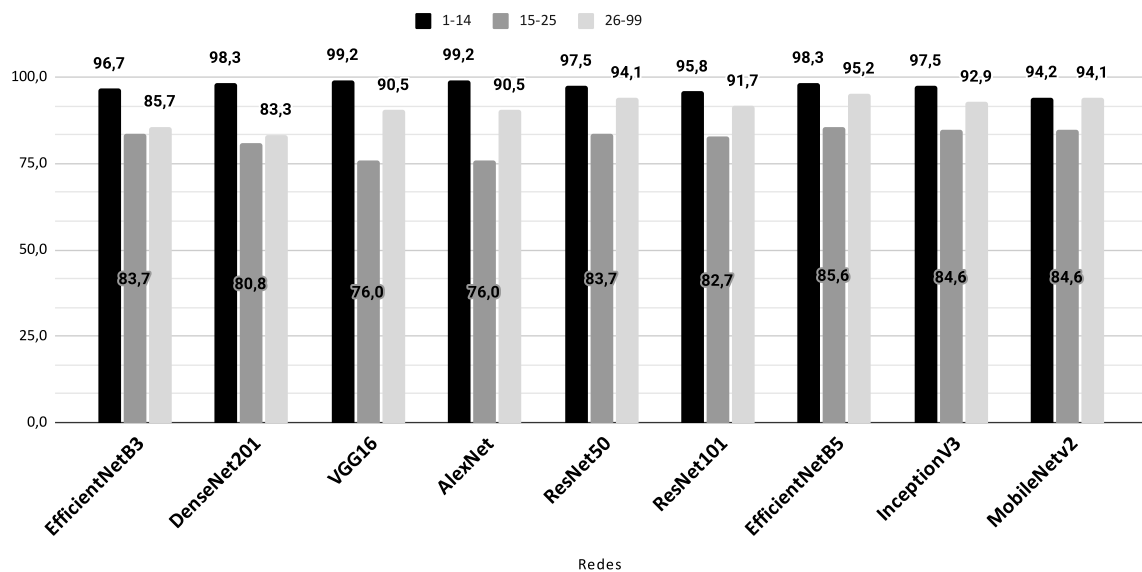


Figura 26 – Especificidade por faixa etária

5.2.4 Resultados Encontrados para Matriz de Confusão

A Figura 27 apresenta as matrizes de confusão para a rede *Resnet101* (Figura 27(a)) e a *Resnet50* (Figura 27(b)). Como pode ser observado nas matrizes de confusão são demonstrados três categorias: 1–14 anos, 15–25 anos e 26–99 anos.

A rede *Resnet101*, demonstrou desempenho satisfatório nas classes mais extremas, com acurácia de 85,3% para a faixa 1–14 anos (29 acertos de 34 amostras) e 81,4% para a faixa 26–99 anos (57 de 70). Entretanto, a classe intermediária (15–25 anos) apresentou maior taxa

de confusão, com 76% de acerto (38 de 50), e erros distribuídos tanto para a classe inferior (5) quanto para a superior (7), refletindo a dificuldade comum em classificar zonas etárias de transição. A Figura 27(b), demonstra desempenho superior da rede *ResNet50* em duas das três classes. Para a classe 1–14 anos, atingiu 91,2% de acurácia (31 acertos de 34), sem confusão com a classe 26–99. A classe 15–25 apresentou 84% de acerto (42 de 50), com menor dispersão de erros em relação à *ResNet101*. Já a classe 26–99 manteve desempenho próximo (80%, com 56 acertos).

Comparando com a rede anterior (*ResNet101*), observa-se uma melhoria significativa na classe 1–14 e na classe 15-25. Embora a *ResNet101* seja uma arquitetura mais profunda, com um número maior de camadas e, conseqüentemente, maior capacidade teórica de extração de características, essa complexidade adicional nem sempre se traduz em melhor desempenho prático. A *ResNet50*, por ser mais leve e convergir mais rapidamente, tendeu a ajustar os pesos de forma mais estável, reduzindo erros.

Portanto, a análise da matriz de confusão sugere que a *ResNet50* é eficaz para a tarefa, apresentando um desempenho geral satisfatório. No entanto, ainda há espaço para aprimoramentos na redução das ambigüidades entre as faixas intermediárias, possivelmente com ajustes de arquitetura, técnicas de balanceamento amostral ou abordagens de aprendizado focalizado, dado que existe sobreposição morfológica entre faixas etárias próximas.

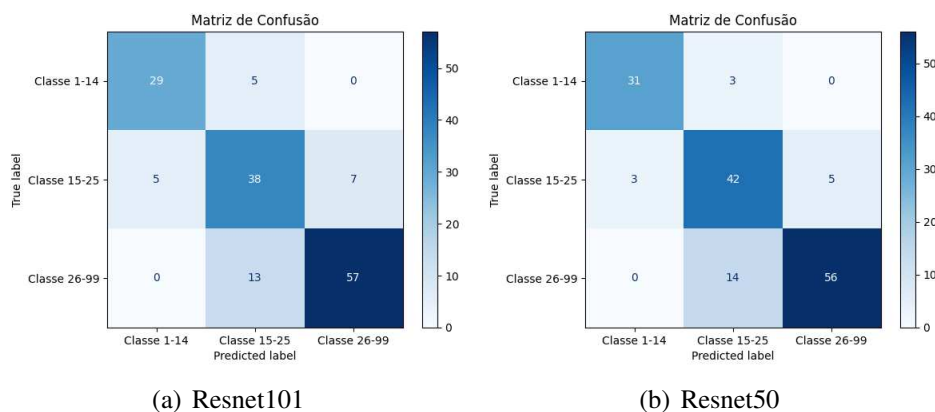


Figura 27 – Matriz de confusão *Resnet*

A Figura 28 apresenta as matrizes de confusão para a rede *EfficientNet B3* (Figura 28(a)) e a *EfficientNet B5* (Figura 28(b)). Ambas apresentaram elevada acurácia na faixa 1–14 anos, com 30 acertos para a *B3* e 31 para a *B5*, sendo os erros concentrados na classe 15–25. Essa concentração indica que a rede aprendeu as transições naturais entre faixas adjacentes. Dessa forma, a classe intermediária foi novamente a mais desafiadora, com 68% de acurácia na

B3 (34 de 50) e 88% na *B5* (44 de 50), evidenciando a maior robustez da *B5* na identificação de padrões associados ao final da adolescência e início da vida adulta. Na faixa 26–99, a *B3* obteve 57 acertos (81,4%) e a *B5* 58 acertos (82,9%), confirmando a vantagem geral da *B5*. De forma geral, a *EfficientNet-B5* superou a *B3* em todas as faixas etárias e métricas avaliadas. Sua arquitetura mais robusta permite a extração mais eficaz de características morfológicas sutis, refletindo em menor dispersão dos erros e maior estabilidade preditiva. Portanto, embora a *EfficientNet-B3* consiga realizar a tarefa de classificação, ainda há margem significativa para novos testes e aprimoramentos.

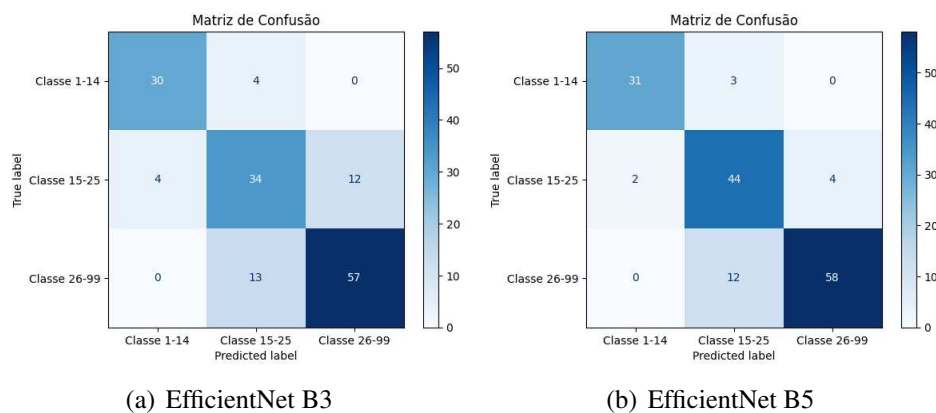


Figura 28 – Matriz de confusão *EfficientNet*

A Figura 29 apresenta as matrizes de confusão para a rede *InceptionV3* (Figura 29(a)) e a *MobileNet* (Figura 29(b)). A *MobileNetV2* alcançou excelente desempenho na classe 1–14 anos (91,2%), com erros restritos à classe 15–25 e nenhuma confusão com a classe 26–99. Para a faixa intermediária, o modelo apresentou 76% de acerto (38 de 50), com 7 erros para a classe inferior e 5 para a superior. Na classe 26–99, obteve 81,4% de acurácia (57 de 70), com erros apenas para a faixa adjacente. Já a *InceptionV3* demonstrou 85,3% de acurácia na classe 1–14, 82% na 15–25 e 84,3% na 26–99. Sua maior na faixa 15–25 em relação a *MobileNetV2*, indica uma melhor separação dos padrões morfológicos intermediários.

A Figura 30 apresenta as matrizes de confusão para a rede *VGG16* (Figura 30(a)), a *Dense201* (Figura 30(b)) e *AlexNet* (Figura 30(c)). De maneira geral, essas arquiteturas apresentaram desempenho inferior em relação às redes supracitadas, especialmente na classe 1–14 anos. A *AlexNet* obteve apenas 67,7% de acurácia nessa faixa, a *VGG-16* alcançou 70,6% e a *DenseNet201* 73,5%. Os erros estiveram majoritariamente associados à confusão com a faixa 15–25, indicando limitações na capacidade dessas redes em capturar padrões morfológicos associados a infância. Todavia, a *AlexNet* e a *VGG-16* apresentaram desempenho razoável na

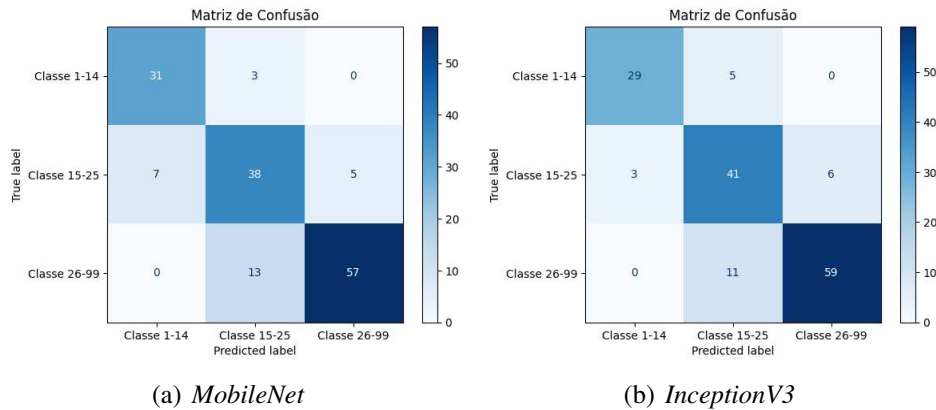


Figura 29 – Matriz de confusão *MobileNet* e *InceptionV3*

classe 15–25 (84% e 82%, respectivamente), e a *DenseNet201* destacou-se na faixa 26–99 com acurácia de 82,9%. Os resultados mostram que todas as redes tiveram melhor desempenho nas classes 1–14 e 26–99, com menor taxa de confusão entre os extremos etários.

A faixa 15–25, por outro lado, concentrou a maior parte dos erros, o que é justificável, dada sua posição como zona de transição morfológica. Esse padrão foi recorrente em todas as arquiteturas, com acurácias variando de 68% (*EfficientNet-B3*) a 88% (*EfficientNet-B5*), e sugere que ajustes na forma de agrupar as faixas ou a utilização de abordagens como aprendizado focalizado podem ser estratégias promissoras. Ressalta-se, que nenhum dos modelos cometeu erros críticos entre as classes 1–14 e 26–99, reforçando a robustez geral das redes para classificações mais extremas. Tal característica é particularmente relevante em aplicações clínicas ou forenses, nas quais equívocos graves entre crianças e adultos podem comprometer a validade das análises. Dessa, embora existam margens para aperfeiçoamentos, os modelos avaliados demonstram bom potencial na tarefa proposta, com destaque para a *EfficientNet-B5* e a *ResNet50*.

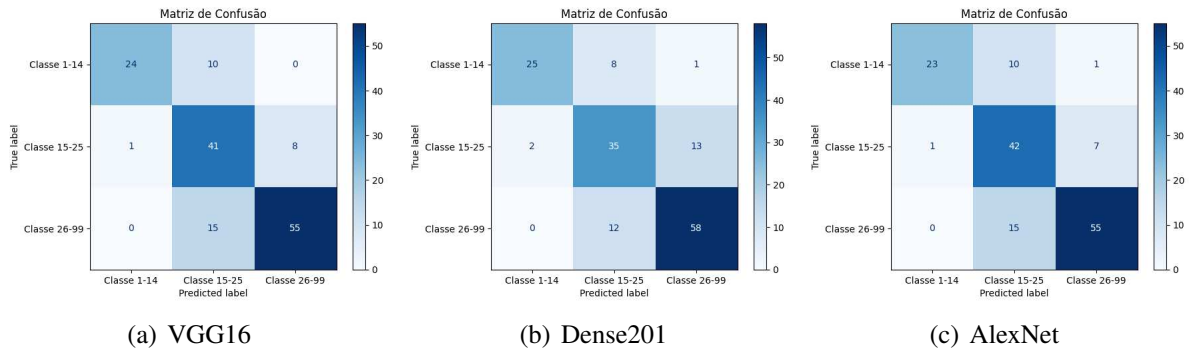


Figura 30 – Matriz de confusão *VGG16*, *Dense201* e *AlexNet*

5.2.5 Resultados para Validação Externa

A Tabela 13 exibe os resultados para as métricas de avaliação das três arquiteturas escolhidas para a validação externa: *EfficientNetB5*, *ResNet50* e *InceptionV3*. A arquitetura *EfficientNetB5* obteve o melhor desempenho, alcançando acurácia de 94,50%, precisão de 91,55%, *recall* de 94,14% e F1-score de 92,83% e especificidade 96,92%, apesar de ter obtido o maior tempo de treinamento 384 minutos.

A *ResNet50* também demonstrou desempenho satisfatório, com acurácia de 92,88%, precisão de 92,91%, *recall* de 92,88%, F1-score de 92,89% e especificidade de 95,14%. O tempo de treinamento foi de 350 minutos.

Ressalta-se, que o treinamento dos modelos com um conjunto de dados numericamente maior, usado na etapa de validação externa, contribuiu para aumentar taxas dos resultados das métricas de avaliação. Possivelmente, ampliação da variabilidade dos dados de treino permitiu que as redes neurais generalizassem melhor as características relevantes para a classificação etária.

Dessa forma, os resultados obtidos na validação externa confirmam a robustez dos modelos treinados, evidenciando sua capacidade de adaptação e generalização frente a dados brutos não utilizados no treinamento inicial. A análise conjunta das métricas demonstra que a *EfficientNetB5* apresentou superioridade em desempenho, enquanto a *InceptionV3* destacou-se pelo menor tempo de treinamento.

A Figura 31 mostra as matrizes de confusão de cada.

Tabela 13 – Resultados de desempenho dos modelos avaliados na avaliação externa

Rede	Acurácia (%)	Precisão (%)	Recall (%)	F1-Score (%)	Especificidade (%)	Tempo (min)
EfficientNetB5	94,50	91,55	94,14	92,83	96,92	384
ResNet50	92,88	92,91	92,88	92,89	95,14	350
InceptionV3	91,91	91,86	91,91	91,87	94,16	300

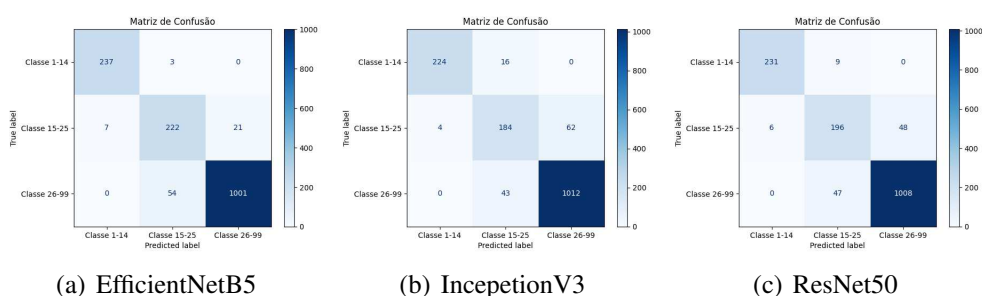


Figura 31 – Matriz de confusão *EfficientNetB5*, *InceptionV3* e *ResNet50*

5.3 Implicações do Estudo

Os resultados obtidos neste estudo apresentam implicações práticas e científicas relevantes para pesquisadores, desenvolvedores de ferramentas, e profissionais que atuam na área de aprendizado profundo aplicado à área da saúde e à visão computacional.

5.3.1 Implicações para Pesquisadores

Este estudo reforça a importância da revisão sistemática da literatura como base para a escolha fundamentada de arquiteturas de redes neurais profundas, assegurando respaldo metodológico e maior credibilidade à seleção realizada. A transparência do protocolo experimental adotado também representa um ponto relevante que beneficia a comunidade científica. Desde a definição dos critérios de escolha dos modelos até a execução dos testes, o trabalho oferece um ponto de partida validado, tanto teoricamente quanto empiricamente, facilitando a replicação, adaptação e expansão por outros pesquisadores.

Além disso, o cálculo padronizado das métricas por faixa etária oferece uma visão mais abrangente do desempenho dos modelos. Com base nos achados deste trabalho, torna-se possível investigar quais arquiteturas apresentam melhor desempenho em faixas etárias específicas, ampliando a aplicabilidade dos modelos de acordo com o público-alvo. Isso é especialmente relevante considerando que, nesse tipo de pesquisa, é muito comum selecionar apenas uma faixa etária específica para análise.

A disponibilização pública dos bancos de dados e do código-fonte promove boas práticas de ciência aberta e facilita a replicação de estudos. Isso permite que pesquisadores de diferentes instituições testem a robustez dos modelos em cenários diversos e com populações distintas, contribuindo para o avanço da visão computacional aplicada à saúde. Além disso, a escolha de métricas validação padronizadas permite que estes resultados sejam comparados diretamente com outros trabalhos da literatura.

Há espaço para investigações futuras que explorem arquiteturas alternativas, incluindo redes recentes como os *Vision Transformers*. Além das escolhas arquiteturais, destacam-se como promissoras as abordagens que envolvem técnicas de balanceamento de classes, segmentação automática de regiões de interesse e o uso de aprendizado auto-supervisionado. Este último representa uma estratégia eficaz para reduzir a dependência de conjuntos de dados rotulados manualmente, permitindo que os modelos aprendam representações úteis a partir de dados não

anotados. Outra vertente possível é a adaptação das redes para tarefas de regressão, especialmente quando o objetivo for estimar a idade de forma contínua. Além disso, a utilização de mapas de calor pode auxiliar na identificação das regiões anatômicas mais relevantes para a discriminação entre faixas etárias. Com base nessas informações, é viável realizar um treinamento direcionado, focando nas regiões de interesse identificadas, o que pode reforçar o aprendizado da rede e potencialmente melhorar seu desempenho na estimativa de idade.

É importante destacar que, embora plataformas como o *Google Colab Pro* represente uma alternativa viável economicamente e amplamente utilizada na comunidade científica para experimentação com redes neurais profundas, seu uso apresentou limitações. Uma das principais restrições enfrentadas foi o limite mensal de unidades computacionais disponíveis, que se mostrou insuficiente para a execução completa dos experimentos propostos. Dessa forma, no desenvolvimento deste trabalho, foi necessário recorrer à versão *Google Colab Pro+* em diversos momentos, a fim de viabilizar o treinamento dos modelos, evitar interrupções e garantir maior estabilidade no acesso a recursos computacionais. Esse aspecto ressalta a importância de considerar, já na fase de planejamento experimental, as demandas computacionais associadas ao uso de redes profundas, bem como a necessidade de infraestrutura adequada ou investimento complementar para execução eficiente dos experimentos.

5.3.2 Implicações para Desenvolvedores de Ferramentas

O *pipeline* experimental apresentado, desde o pré-processamento até a validação dos resultados, pode servir de base para o desenvolvimento de ferramentas automatizadas voltadas à classificação de imagens médicas. Desenvolvedores podem aproveitar a padronização dos procedimentos adotados, como o uso de ambientes baseados em nuvem (*Google Colab*) e bibliotecas específicas (*PyTorch*, *OpenCV*), para criar soluções reproduzíveis e portáteis. Além disso, a disponibilização de todos os artefatos do estudo (scripts, modelos, logs e métricas) oferece um ponto de partida para construir plataformas que auxiliem médicos legistas, odontologistas ou pesquisadores a validar redes em novos conjuntos de imagens com facilidade.

Os desenvolvedores podem utilizar este estudo como referência para selecionar arquiteturas que melhor atendam às suas necessidades, seja em termos de tempo de treinamento, acurácia, precisão, tempo de inferência ou outros critérios relevantes. Dessa forma, essas arquiteturas podem ser incorporadas a aplicações práticas voltadas ao apoio de profissionais, como médicos legistas e odontologistas.

As métricas também oferecem insumos para a criação de módulos de validação automática em ferramentas baseadas em IA. Por exemplo, um software pode ser desenvolvido para, antes de aplicar o modelo, verificar se os padrões do conjunto de dados de entrada se assemelham às faixas etárias e condições dos dados utilizados neste estudo. Isso contribui para evitar uso indevido ou interpretações erradas dos resultados por parte de usuários finais.

5.3.3 Implicações para o Ensino e Formação Profissional

As informações e os resultados deste estudo possuem também um importante papel formativo. A organização clara dos experimentos, a coleta e análise das métricas, e a discussão dos resultados compõem um exemplo completo de aplicação de redes neurais profundas em um problema do mundo real. Esse material pode ser explorado em disciplinas de aprendizado de máquina, inteligência artificial aplicada à saúde, ciência de dados biomédicos e cursos de especialização em odontologia legal ou radiologia. Para os estudantes, ter acesso a métricas reais permite entender não apenas o funcionamento dos modelos, mas também as limitações, *trade-offs* e critérios técnicos usados para selecionar ou rejeitar uma arquitetura. Isso ajuda a formar profissionais mais críticos, capazes de interpretar relatórios de IA com base em evidências concretas e tomar decisões técnicas mais fundamentadas.

5.4 Ameaças a Validade Relacionadas ao Estudo Exploratório

O estudo exploratório apresentado neste capítulo foi planejado e conduzido cuidadosamente. No entanto, algumas ameaças à validade podem ter influenciado os resultados e as discussões. Essas ameaças foram consideradas durante o planejamento e a execução do estudo, sendo tratadas conforme diretrizes da literatura. A seguir, são discutidas as principais ameaças à validade, organizadas segundo as quatro categorias propostas por Wohlin *et al.* (2012): validade de construção, validade interna, validade externa e validade de conclusão.

Validade de Construção: Uma ameaça importante à validade de construção refere-se à seleção das arquiteturas de redes neurais profundas. Embora exista uma ampla variedade de arquiteturas descritas na literatura, este estudo não teve como objetivo explorar todas as possibilidades. Para mitigar essa limitação, foi realizada uma revisão sistemática da literatura, apresentada no Capítulo 3, que fundamentou a escolha das redes utilizadas nos experimentos, garantindo alinhamento com o estado da arte e relevância para o problema proposto. Outra

ameaça possível está relacionada ao desenho dos protocolos do estudo. Para reduzir esse risco, os procedimentos experimentais foram definidos com base em estudos anteriores e discutidos entre os pesquisadores, com apoio na literatura especializada.

Validade Interna: As ameaças à validade interna estão relacionadas à influência de fatores não controlados nos resultados do experimento. Uma preocupação relevante diz respeito à possibilidade de inconsistências de software ou hardware impactarem os resultados. Para mitigar esse risco, todos os experimentos foram realizados no ambiente Google Colab, que oferece uma infraestrutura padronizada, com acesso a unidades de processamento gráfico (GPUs) e controle de versões de bibliotecas. Isso contribuiu para a reprodutibilidade e estabilidade dos experimentos, reduzindo o impacto de variáveis externas.

Validade Externa: A ameaça à validade externa diz respeito à generalização dos resultados obtidos para outros contextos. Ainda que dois conjuntos de dados públicos e distintos tenham sido utilizados, compostos por radiografias panorâmicas odontológicas, não se pode garantir que os modelos apresentariam o mesmo desempenho em bases provenientes de outras populações ou dispositivos. Para minimizar esse risco, os conjuntos de dados foram escolhidos de modo a abranger diferentes faixas etárias e variabilidades visuais. Além disso, a disponibilização dos bancos de dados e de todos os artefatos do estudo permite que novos trabalhos realizem validações adicionais em diferentes cenários.

Validade de Conclusão: Em relação à validade de conclusão, uma ameaça relevante é a possibilidade de interpretações equivocadas ou generalizações indevidas dos resultados. Para reduzir esse risco, os dados foram apresentados por meio de diferentes gráficos e tabelas, favorecendo a análise comparativa entre os modelos. O pré-processamento das imagens incluiu etapas padronizadas, como normalização e recorte da região de interesse (ROI), o que contribuiu para a consistência dos dados e reduziu ruídos que pudessem impactar as conclusões. Além disso, foram utilizadas métricas amplamente reconhecidas na literatura: acurácia, precisão, *recall*, especificidade e *F1-score*, facilitando a comparação com estudos relacionados e fortalecendo a robustez das análises. Todos os códigos, modelos e resultados estão disponíveis para consulta, garantindo transparência e permitindo replicação.

5.5 Considerações Finais do Estudo Exploratório

Este capítulo documentou a condução de um estudo exploratório voltado à classificação de faixas etárias com base em radiografias panorâmicas odontológicas, utilizando

nove arquiteturas de redes neurais convolucionais previamente selecionadas por meio de uma revisão sistemática. A metodologia envolveu etapas bem definidas de configuração experimental, pré-processamento das imagens, aumento de dados, aplicação de técnicas de *transfer learning*, definição de hiperparâmetros e validação externa, sempre com foco na reprodutibilidade e rigor metodológico.

Os modelos foram avaliados segundo métricas amplamente reconhecidas na literatura, com destaque para acurácia, precisão, *recall*, *F1-score* e especificidade. Os resultados evidenciaram que, embora todas as arquiteturas tenham apresentado desempenho funcional, algumas se destacaram por sua robustez e equilíbrio entre as classes. A *EfficientNetB5*, obteve os melhores resultados gerais em termos de desempenho preditivo e menor dispersão de erros, enquanto outras arquiteturas, como *ResNet50* e *InceptionV3*, também demonstraram resultados consistentes. Por outro lado, observou-se que a faixa etária intermediária (15–25 anos) foi a mais desafiadora para todos os modelos, sendo recorrente a confusão com as faixas adjacentes.

A etapa de validação externa permitiu observar a capacidade de generalização dos modelos, destacando o comportamento das redes diante de dados inéditos e com maior heterogeneidade. A análise das matrizes de confusão confirmou tendências de acerto mais expressivas nas classes extremas e maior sobreposição na classe intermediária, indicando um limite da capacidade discriminativa das arquiteturas testadas para essa faixa de transição.

Em síntese, este estudo reforça a viabilidade da aplicação de redes neurais profundas na estimativa de idade a partir de imagens radiográficas, aponta caminhos para refinamentos metodológicos e oferece subsídios concretos para pesquisas futuras, desenvolvimento de ferramentas aplicadas e uso em contextos educacionais e clínicos.

O próximo capítulo faz um resumo das principais resultados desta dissertação além de apresentar os trabalhos futuros.

6 CONCLUSÃO E TRABALHOS FUTUROS

Este capítulo resume os resultados desta dissertação, em relação aos seus objetivos, contribuições e trabalho futuro. A Seção 6.1 resume os principais resultados alcançados por meio deste trabalho. A Seção 6.2 analisa as principais contribuições. A Seção 6.3 descreve possíveis ideias para trabalho futuro.

6.1 Resumo dos Principais Resultados

Este trabalho se destinou a alcançar os seguintes objetivos específicos (OE) conforme descrito na Seção 1.2.

- *OE1* Investigar as estratégias de redes neurais profundas apresentadas na literatura, por meio de uma revisão sistemática da literatura.
- *OE2* Selecionar um grupo das melhores redes neurais profundas utilizadas para estimativa de imagens.
- *OE3* Criar um repositório para disponibilização dos dados encontrados sobre as redes neurais identificadas.
- *OE4* Criar um banco de imagens com radiografias odontológicas panorâmicas.
- *OE5* Estabelecer um protocolo para padronização das expansões futuras do banco de imagens.
- *OE6* Criar um repositório para disponibilização do banco de imagens.
- *OE7* Realizar um estudo exploratório sobre as principais redes neurais profundas para estimativa de idades.
- *OE8* Criar um repositório para disponibilização do pacote de replicação do estudo, para apoiar pesquisas futuras.

Para os objetivos específicos (OE1, OE2 e OE3), foi conduzida uma revisão sistemática da literatura para sumarizar as redes neurais profundas e suas características na perspectiva de identificação de idade de pacientes por meio de imagens radiográficas. Com este propósito, foram coletados os dados disponíveis em bibliotecas digitais citadas na literatura e que indexam periódicos de eventos de impacto na área de estudo deste trabalho. A fim de obter informações mais atuais foi mantido o foco em publicações que foram realizadas entre os anos de 2020 e 2024. E, para coleta e seleção dos trabalhos primários foi definido um protocolo de revisão sistemática da literatura, sendo que seis bases de dados foram utilizadas. Foram selecionados, segundo os

critérios de inclusão e exclusão, 17 trabalhos para responder as questões de pesquisa performadas no estudo. Como resultado desta RSL, foram identificadas 25 redes neurais. Dentre estas, as que tiveram melhor desempenho foram as redes *ResNet* e *SOS-Net*. Também foi observado que para o grupo das crianças (de 3 a 14 anos), as redes atingiram melhores resultados em comparação às demais faixas etárias. Ainda, as métricas acurácia e MAE foram as mais utilizadas para avaliar as redes neurais. Portanto, conclui-se que embora redes neurais sejam aplicadas na classificação de imagens radiográficas, e ainda com resultados promissores, faz-se necessário que os estudos avancem na identificação dos grupos etários acima de 25 anos de idade.

Ainda, para os objetivos específicos (OE4, OE5 e OE6), foi conduzido a criação de um banco de imagens para o qual foram discutidas as etapas de aquisição, preparação, qualificação, organização e disponibilização das radiografias panorâmicas odontológicas. Ao todo, foram reunidas 1.545 imagens provenientes de três clínicas localizadas em dois estados distintos do Brasil. Foram levantadas para cada imagem do banco de imagens metadados constituídos das seguintes informações: sexo, idade, número de dentes permanentes, identificação dos dentes na arcada dentária segundo a nomenclatura dental pela Federação Dentária Internacional (DFI). Foi disponibilizado um repositório público para que novas pesquisas possam utilizar o conjunto de imagens construído neste trabalho.

Finalmente para os objetivos específicos (OE7 e OE8) foi projetado e executado um estudo exploratório com um subgrupo de nove redes neurais profundas encontradas no Capítulo 3. As redes analisadas foram qualificadas em sete métricas de comparação, tais como: acurácia e precisão. Em conformidade com os demais capítulos desta dissertação, que apresentam e disponibilizam todos os artefatos para apoiar a replicabilidade do estudo, o protocolo e etapas do estudo exploratório, apresentados no Capítulo 5, foram projetados e seguidos para a condução do estudo. Desta forma, futuras pesquisas relacionadas podem replicar o estudo exploratório conduzido, por meio do pacote de replicação do estudo, disponibilizado no repositório do trabalho.

6.2 Principais Contribuições

O presente trabalho foi dividido em três partes principais sendo elas: revisão sistemática da literatura, criação de um banco de imagens público para toda comunidade acadêmica e um estudo exploratório com arquiteturas de redes neurais para estimativa de idade. Algumas contribuições importantes foram alcançadas, bem como o suporte para trabalhos futuros que

podem ser feitos por meio dos desdobramentos e artefatos de pesquisa disponibilizados neste trabalho. Desta forma, esta dissertação resultou nas seguintes contribuições:

Parte 1: Revisão Sistemática da Literatura

- Uma breve visão geral sobre as características das redes neurais profundas encontradas;
- Repositório público com as redes neurais profundas encontradas neste trabalho, além, de disponibilização de estrutura para que novas redes desenvolvidas no futuro poder ser incluídas no repositório.

Parte 2: Banco da imagens público para toda comunidade

- Protocolo para ampliação do banco de imagens de radiografias odontológica panorâmicas;
- Repositório público com as imagens e seus metadados.

Parte 3: Estudo exploratório com nove redes neurais profundas

- Medição dos desempenho das nove redes neurais profundas utilizadas no estudo;
- Definição de protocolo para comparação das redes neurais profundas para estimativa de idade por meio de radiografia odontológicas panorâmicas.

O repositório com os dados desta pesquisa está disponível no seguinte endereço:

[<https://rsl-redes-neurais-profundas.github.io/rsl/>](https://rsl-redes-neurais-profundas.github.io/rsl/)

6.3 Trabalhos Futuros

Foram apresentados os achados na revisão sistemática conduzida, a estrutura e possibilidades do conjunto de imagens proposto e o estudo exploratório sobre redes neurais profundas para a estimativa de idade por meio de radiografias odontológicas panorâmicas. No entanto, novas possibilidades de estudos relacionadas para as três partes principais desta dissertação são discutidas nesta seção.

Inicialmente, a revisão sistemática da literatura encontrou 25 arquiteturas de redes utilizadas para estimativa de idade por meio de radiografia odontológica panorâmica. Foi feito um repositório que podem conter novas arquiteturas de rede desenvolvidas no futuro. Desta

forma, o repositório pode servir uma das principais bases de consulta para arquiteturas de redes relacionadas ao tema. Como trabalho futuro, é desejado expandir este repositório com outras arquiteturas de redes neurais profundas utilizadas para identificação de imagens em outros contextos médicos, tais como identificação de tumores, radiografia de outras partes do corpo e ressonância magnética.

Para o banco de imagens apresentados neste trabalho, deste a definição dos protocolos do banco foi tomando o cuidado de criar uma base de dados que possa apoiar não apenas o presente trabalho, mas também outros trabalhos relacionados a redes neurais profundas e de outras áreas, até mesmo no contexto de técnicas forense manuais de estimativa de idade. É desejado como trabalho futuro que novas imagens sejam incorporadas ao banco de imagens, seguindo a estrutura feita. Como por exemplo, incorporar as imagens do banco de dados de Oliveira *et al.* (2024) na estrutura proposta do presente banco de imagens.

Em relação ao estudo exploratório feito com as redes neurais profundas os resultados são preliminares e precisam ser validados em outros contextos, tais como com outras arquiteturas de redes neurais profundas. Portanto, é necessário expandir o escopo de arquiteturas antes da generalização adequada dos resultados encontrados, com outros experimentos adicionais comparando outras arquiteturas além das arquiteturas utilizadas neste trabalho. No trabalho futuro nessa direção, é planejado investigar diferentes arquiteturas de redes neurais profundas relacionadas a área de visão computacional.

Além disso, para uma análise mais precisa e individualizada das idades estimadas a partir de radiografias panorâmicas, torna-se necessário o desenvolvimento de estudos de regressão. Diferentemente das abordagens de classificação, que agrupam as idades em faixas etárias, os modelos de regressão possibilitam a predição de valores contínuos, o que pode representar um avanço significativo na estimativa de idade cronológica. Futuramente, sugere-se a implementação e avaliação de diferentes arquiteturas de redes neurais voltadas para tarefas de regressão, comparando seus desempenhos com métodos tradicionais da literatura odontológica forense. Essa abordagem pode contribuir para maior acurácia na estimativa individual, além de ampliar o potencial de aplicação clínica e pericial dos sistemas baseados em inteligência artificial.

REFERÊNCIAS

- AKTER, M. S.; SHAHRIAR, H.; SNEHA, S.; CUZZOCREA, A. Multi-class skin cancer classification architecture based on deep convolutional neural network. In: IEEE. **2022 IEEE International Conference on Big Data (Big Data)**. [S.l.], 2022. p. 5404–5413.
- ARTHANARI, A.; RAVINDRAN, V.; RAMALINGAM, K.; PRATHAP, L.; RAJ, S. Gender determination through mandibular features on orthopantomograms: A preliminary study. **Cureus**, Cureus, v. 16, n. 7, 2024.
- ATAŞ, İ. Performance evaluation of jaccard-dice coefficient on building segmentation from high resolution satellite images. **Balkan Journal of Electrical and Computer Engineering**, MUSA YILMAZ, v. 11, n. 1, p. 100–106, 2023.
- ATAŞ, İ.; ÖZDEMİR, C.; ATAŞ, M.; DOĞAN, Y. Forensic dental age estimation using modified deep learning neural network. **Balkan Journal of Electrical and Computer Engineering**, MUSA YILMAZ, v. 11, n. 4, p. 298–305, 2022.
- ATAŞ, M.; YEŞILNACAR, M. İ.; YETİŞ, A. D. Novel machine learning techniques based hybrid models (lr-knn-ann and svm) in prediction of dental fluorosis in groundwater. **Environmental Geochemistry and Health**, Springer, v. 44, n. 11, p. 3891–3905, 2022.
- ATHANASIADOU, E.; GERADTS, Z.; EIJK, E. V. Camera recognition with deep learning. **Forensic sciences research**, Oxford University Press, v. 3, n. 3, p. 210–218, 2018.
- BEAINI, T. L.; DIAS, P. E. M.; MELANI, R. F. H. Importância pericial dos sistemas de notação dental–revisão de literatura. **Revista Brasileira de Odontologia Legal**, v. 3, n. 1, 2016.
- BISHOP, C. M.; NASRABADI, N. M. **Pattern recognition and machine learning**. [S.l.]: Springer, 2006. v. 4.
- BULGARELI, J. V.; FARIA, E. T. d.; CORTELLAZZI, K. L.; GUERRA, L. M.; MENEGHIM, M. d. C.; AMBROSANO, G. M. B.; FRIAS, A. C.; PEREIRA, A. C. Fatores que influenciam o impacto da saúde bucal nas atividades diárias de adolescentes, adultos e idosos. **Revista de saude publica**, SciELO Brasil, v. 52, p. 44, 2018.
- COLLOBERT, R.; WESTON, J.; BOTTOU, L.; KARLEN, M.; KAVUKCUOGLU, K.; KUKSA, P. Natural language processing (almost) from scratch. 2011.
- CORREIA, A. de M.; BARBOSA, D. da S.; ALCANTARA, J. A. da S.; OLIVEIRA, P. M. da C.; SILVA, P. G. de B.; FRANCO, A.; BEZERRA, T. P. Performance and comparison of the london atlas technique and cameriere’s third molar maturity index (i3m) for allocating individuals below or above the threshold of 18 years. **Forensic Science International**, Elsevier, v. 317, p. 110512, 2020.
- DEMIRJIAN, A.; GOLDSTEIN, H.; TANNER, J. M. A new system of dental age assessment. **Human biology**, JSTOR, p. 211–227, 1973.
- DENG, J.; DONG, W.; SOCHER, R.; LI, L.-J.; LI, K.; FEI-FEI, L. Imagenet: A large-scale hierarchical image database. In: IEEE. **2009 IEEE conference on computer vision and pattern recognition**. [S.l.], 2009. p. 248–255.

DEZEM, T. U.; FRANCO, A.; PALHARES, C. E. M.; DEITOS, A. R.; SILVA, R. H. A. da; SANTIAGO, B. M.; RIBEIRO, I. L. A.; JUNIOR, E. D. Testing the olze and timme methods for dental age estimation in radiographs of brazilian subadults and adults. **Acta stomatologica Croatica: International journal of oral sciences and dental medicine**, Sveučilište u Zagrebu Stomatološki fakultet, v. 55, n. 4, p. 390–396, 2021.

DONG, W.; YOU, M.; HE, T.; DAI, J.; TANG, Y.; SHI, Y.; GUO, J. An automatic methodology for full dentition maturity staging from opg images using deep learning. **Applied Intelligence**, Springer, v. 53, n. 23, p. 29514–29536, 2023.

FABBRI, S.; SILVA, C.; HERNANDES, E.; OCTAVIANO, F.; THOMMAZO, A. D.; BELGAMO, A. Improvements in the start tool to better support the systematic review process. In: **Proceedings of the 20th international conference on evaluation and assessment in software engineering**. [S.l.: s.n.], 2016. p. 1–5.

FISCHBORN, A. R.; PEDROSO, C. M.; VAZ, C.; MIGLIORINI, E.; CLAUDINO, M.; FRANCO, G. C. Prevalência das posições e grau de dificuldade cirúrgica de terceiros molares inferiores: um estudo radiográfico retrospectivo. **Revista da Faculdade de Odontologia de Porto Alegre**, v. 62, n. 2, p. 64–73, 2021.

FORTES, A. B. C.; LIMA, L. N. C.; FURTADO, F. M. S. Análise da importância da documentação odontológica no processo de identificação humana no iml de são luís, maranhão. **Revista Brasileira de Odontologia Legal**, v. 7, n. 2, 2020.

FRID-ADAR, M.; DIAMANT, I.; KLANG, E.; AMITAI, M.; GOLDBERGER, J.; GREENSPAN, H. Gan-based synthetic medical image augmentation for increased cnn performance in liver lesion classification. **Neurocomputing**, Elsevier, v. 321, p. 321–331, 2018.

GALIBOURG, A.; CUSSAT-BLANC, S.; DUMONCEL, J.; TELMON, N.; MONSARRAT, P.; MARET, D. Comparison of different machine learning approaches to predict dental age using demirjian's staging approach. **International Journal of Legal Medicine**, Springer, v. 135, p. 665–675, 2021.

GOBBO, S. F. R.; ALONSO, M. B. C. C.; KAWAMOTO, K. K. M.; TEIXEIRA, D. d. B.; SILVA, R. H. A. d.; COMAR, L. P. Estimativa da idade dental pelo método de nicodemo em uma população da região sudeste do brasil. **Revista Criminalística e Medicina Legal**, v. 6, n. 1, p. 10–18, 2021.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A.; BENGIO, Y. **Deep learning**. [S.l.]: MIT press Cambridge, 2016. v. 1.

GUO, Y.-C.; HAN, M.; CHI, Y.; LONG, H.; ZHANG, D.; YANG, J.; YANG, Y.; CHEN, T.; DU, S. Accurate age classification using manual method and deep convolutional neural network based on orthopantomogram images. **International journal of legal medicine**, Springer, v. 135, p. 1589–1597, 2021.

HAYKIN, S. **Redes neurais: princípios e prática**. [S.l.]: Bookman Editora, 2001.

HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2016. p. 770–778.

HOWARD, A. G.; ZHU, M.; CHEN, B.; KALENICHENKO, D.; WANG, W.; WEYAND, T.; ANDREETTO, M.; ADAM, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. **arXiv preprint arXiv:1704.04861**, 2017.

HUANG, A.; JIANG, L.; ZHANG, J.; WANG, Q. Attention-vgg16-unet: a novel deep learning approach for automatic segmentation of the median nerve in ultrasound images. **Quantitative imaging in medicine and surgery**, v. 12, n. 6, p. 3138, 2022.

HUANG, G.; LIU, Z.; MAATEN, L. V. D.; WEINBERGER, K. Q. Densely connected convolutional networks. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2017. p. 4700–4708.

HUANG, G. B.; MATTAR, M.; BERG, T.; LEARNED-MILLER, E. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. In: **Workshop on faces in 'Real-Life' Images: detection, alignment, and recognition**. [S.l.: s.n.], 2008.

IRVIN, J.; RAJPURKAR, P.; KO, M.; YU, Y.; CIUREA-ILCUS, S.; CHUTE, C.; MARKLUND, H.; HAGHGOO, B.; BALL, R.; SHPANSKAYA, K. *et al.* Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. In: **Proceedings of the AAAI conference on artificial intelligence**. [S.l.: s.n.], 2019. v. 33, n. 01, p. 590–597.

ISLAM, S.; ELMEEKKI, H.; ELSEBAI, A.; BENTAHAR, J.; DRAWEL, N.; RJOUB, G.; PEDRYCZ, W. A comprehensive survey on applications of transformers for deep learning tasks. **Expert Systems with Applications**, Elsevier, v. 241, p. 122666, 2024.

JUNIOR, E. F.; MOURA, L. C. L. de. A importância dos arcos dentários na identificação humana. **Revista Brasileira de Odontologia**, v. 71, n. 1, p. 22, 2014.

KEBAILI, A.; LAPUYADE-LAHORGUE, J.; RUAN, S. Deep learning approaches for data augmentation in medical imaging: a review. **Journal of imaging**, MDPI, v. 9, n. 4, p. 81, 2023.

KEELE, S. *et al.* **Guidelines for performing systematic literature reviews in software engineering**. [S.l.], 2007.

KIM, S.; LEE, Y.; NOH, Y.; PARK, F.; AUH, Q. Age-group determination of living individuals using first molar images based on artificial intelligence. *sci rep* 11 (1): 1073. **doi. org/10**, v. 1038, 2021.

KIM, Y.-R.; CHOI, J.-H.; KO, J.; JUNG, Y.-J.; KIM, B.; NAM, S.-H.; CHANG, W.-D. Age group classification of dental radiography without precise age information using convolutional neural networks. In: MDPI. **Healthcare**. [S.l.], 2023. v. 11, n. 8, p. 1068.

KOHAVI, R. *et al.* A study of cross-validation and bootstrap for accuracy estimation and model selection. In: MONTREAL, CANADA. **Ijcai**. [S.l.], 1995. v. 14, n. 2, p. 1137–1145.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. **Advances in neural information processing systems**, v. 25, 2012.

KVAAL, S. I.; KOLLTVEIT, K. M.; THOMSEN, I. O.; SOLHEIM, T. Age estimation of adults from dental radiographs. **Forensic science international**, Elsevier, v. 74, n. 3, p. 175–185, 1995.

KWON, O.; YONG, T.-H.; KANG, S.-R.; KIM, J.-E.; HUH, K.-H.; HEO, M.-S.; LEE, S.-S.; CHOI, S.-C.; YI, W.-J. Automatic diagnosis for cysts and tumors of both jaws on panoramic radiographs using a deep convolution neural network. **Dentomaxillofacial Radiology**, Oxford University Press, v. 49, n. 8, p. 20200185, 2020.

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

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

LIN, T.-Y.; MAIRE, M.; BELONGIE, S.; HAYS, J.; PERONA, P.; RAMANAN, D.; DOLLÁR, P.; ZITNICK, C. L. Microsoft coco: Common objects in context. In: SPRINGER. **Computer vision–ECCV 2014: 13th European conference, zurich, Switzerland, September 6-12, 2014, proceedings, part v 13**. [S.l.], 2014. p. 740–755.

LIU, Z.; LUO, P.; QIU, S.; WANG, X.; TANG, X. Deepfashion: Powering robust clothes recognition and retrieval with rich annotations. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2016. p. 1096–1104.

LUZ, L. C. P. da; ANZULOVIC, D.; BENEDICTO, E. N.; GALIC, I.; BRKIC, H.; BIAZEVIC, M. G. H. Accuracy of four dental age estimation methodologies in brazilian and croatian children. **Science & Justice**, Elsevier, v. 59, n. 4, p. 442–447, 2019.

MIENYE, I. D.; SWART, T. G.; OBAIDO, G.; JORDAN, M.; ILONO, P. Deep convolutional neural networks in medical image analysis: A review. **Information**, Mdpi, v. 16, n. 3, p. 195, 2025.

MOHAMED, E. G.; REDONDO, R. P. D.; KOURA, A.; EL-MOFTY, M. S.; KAYED, M. Dental age estimation using deep learning: A comparative survey. **Computation**, MDPI, v. 11, n. 2, p. 18, 2023.

MOHAMMAD, N.; AHMAD, R.; GAUS, M. H. A.; KURNIAWAN, A.; YUSOF, M. Y. P. M. Accuracy of automated forensic dental age estimation lab (f-dentest lab) on large malaysian dataset. **Forensic Science International**, Elsevier, v. 361, p. 112150, 2024.

MOORREES, C. F.; FANNING, E. A.; JR, E. E. H. Age variation of formation stages for ten permanent teeth. **Journal of dental research**, SAGE Publications Sage CA: Los Angeles, CA, v. 42, n. 6, p. 1490–1502, 1963.

MORENO, M. B. P.; PONTES, T. J. P.; RABELLO, P. M. Utilização da tabela de cronologia de mineralização dental de nicodemo, Moraes e médico filho na estimativa da idade de paraibanos. **Saúde Ética & Justiça**, v. 19, n. 1, p. 35–44, 2014.

MU, C. C.; GANG, L. Age estimation using panoramic radiographs by transfer learning. **Chinese Journal of Dental Research**, v. 25, n. 2, 2022.

MUALLA, N.; HOUSSEIN, E. H.; HASSAN, M. Dental age estimation based on x-ray images. **Computers, Materials & Continua**, v. 62, n. 2, 2020.

MUJAHID, M.; RUSTAM, F.; ÁLVAREZ, R.; MAZÓN, J. L. V.; DÍEZ, I. d. I. T.; ASHRAF, I. Pneumonia classification from x-ray images with inception-v3 and convolutional neural network. **Diagnostics**, MDPI, v. 12, n. 5, p. 1280, 2022.

MUMUNI, A.; MUMUNI, F. Data augmentation: A comprehensive survey of modern approaches. **Array**, Elsevier, v. 16, p. 100258, 2022.

NICODEMO, R. A.; MORAES, L. C. d.; FILHO, E. M. Tabela cronológica da mineralização dos dentes permanentes entre brasileiros. **Rev Fac Odontol São José dos Campos**, v. 3, n. 1, p. 55–6, 1974.

NOLLA, C. M. *et al.* **The development of permanent teeth**. Tese (Doutorado) — University of Michigan Ann Arbor, 1952.

ODZHAKOV, F.; APOSTOLOV, A. Dental age estimation—literature review. **J. Transl. Sci**, v. 6, p. 1–5, 2019.

OLIVEIRA, M. d. M. N.; CORREIA, M. F.; BARATA, J. S. Aspectos relacionados ao emprego da radiografia panorâmica em pacientes infantis. **Revista da Faculdade de Odontologia de Porto Alegre**, v. 47, n. 1, 2006.

OLIVEIRA, W.; SANTOS, M. A.; BURGARDT, C. A. P.; PONTUAL, M. L. A.; ZANCHETTIN, C. Estimation of human age using machine learning on panoramic radiographs for brazilian patients. **Scientific Reports**, Nature Publishing Group UK London, v. 14, n. 1, p. 19689, 2024.

OLVERES, J.; GONZÁLEZ, G.; TORRES, F.; MORENO-TAGLE, J. C.; CARBAJAL-DEGANTE, E.; VALENCIA-RODRÍGUEZ, A.; MÉNDEZ-SÁNCHEZ, N.; ESCALANTE-RAMÍREZ, B. What is new in computer vision and artificial intelligence in medical image analysis applications. **Quantitative imaging in medicine and surgery**, v. 11, n. 8, p. 3830, 2021.

ONG, S.-H.; KIM, H.; SONG, J.-S.; SHIN, T. J.; HYUN, H.-K.; JANG, K.-T.; KIM, Y.-J. Fully automated deep learning approach to dental development assessment in panoramic radiographs. **BMC Oral Health**, Springer, v. 24, n. 1, p. 426, 2024.

OZDEMIR, C. Classification of brain tumors from mr images using a new cnn architecture. **Traitement du Signal**, International Information and Engineering Technology Association (IIETA), v. 40, n. 2, p. 611, 2023.

PAN, Y.; LIU, J.; CAI, Y.; YANG, X.; ZHANG, Z.; LONG, H.; ZHAO, K.; YU, X.; ZENG, C.; DUAN, J. *et al.* Fundus image classification using inception v3 and resnet-50 for the early diagnostics of fundus diseases. **Frontiers in Physiology**, Frontiers Media SA, v. 14, p. 1126780, 2023.

RUBIU, G.; BOLOGNA, M.; CELLINA, M.; CÈ, M.; SALA, D.; PAGANI, R.; MATTAVELLI, E.; FAZZINI, D.; IBBA, S.; PAPA, S. *et al.* Teeth segmentation in panoramic dental x-ray using mask regional convolutional neural network. **Applied Sciences**, MDPI, v. 13, n. 13, p. 7947, 2023.

SANDLER, M.; HOWARD, A.; ZHU, M.; ZHMOGINOV, A.; CHEN, L.-C. Mobilenetv2: Inverted residuals and linear bottlenecks. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2018. p. 4510–4520.

SARE, R. T. International numbering. **Journal of the American Dental Association (1939)**, v. 128, n. 9, p. 1198, 1997.

SARVAMANGALA, D.; KULKARNI, R. V. Convolutional neural networks in medical image understanding: a survey. **Evolutionary intelligence**, Springer, v. 15, n. 1, p. 1–22, 2022.

SATHYANARAYANAN, S.; TANTRI, B. R. Confusion matrix-based performance evaluation metrics. **African Journal of Biomedical Research**, p. 4023–4031, 2024.

SATHYAVATHI, S. Human age estimation using deep convolutional neural network based on dental images (orthopantomogram). **IETE Journal of Research**, Taylor & Francis, v. 70, n. 2, p. 1585–1592, 2024.

SCHMIDHUBER, J. Deep learning in neural networks: An overview. **Neural networks**, Elsevier, v. 61, p. 85–117, 2015.

SHI, Y.; YE, Z.; GUO, J.; TANG, Y.; DONG, W.; DAI, J.; MIAO, Y.; YOU, M. Deep learning methods for fully automated dental age estimation on orthopantomograms. **Clinical Oral Investigations**, Springer, v. 28, n. 3, p. 198, 2024.

SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. **arXiv preprint arXiv:1409.1556**, 2014.

SIVRI, M. B.; TAHERI, S.; ERCAN, R. G. K.; YAĞCI, Ü.; GOLRIZKHATAMI, Z. Dental age estimation: A comparative study of convolutional neural network and demirjian's method. **Journal of Forensic and Legal Medicine**, Elsevier, v. 103, p. 102679, 2024.

SON, D.-M.; YOON, Y.-A.; KWON, H.-J.; AN, C.-H.; LEE, S.-H. Automatic detection of mandibular fractures in panoramic radiographs using deep learning. **Diagnostics**, v. 11, n. 6, p. 933, 2021.

SOUSA, D. Pereira de; LIMA, E. D.; PAULINO, J. A. S.; PONTUAL, M. L. dos A.; BENTO, P. M.; MELO, D. P. Age determination on panoramic radiographs using the kvaal method with the aid of artificial intelligence. **Dentomaxillofacial Radiology**, Oxford University Press, v. 52, n. 4, p. 20220363, 2023.

SZE, V.; CHEN, Y.-H.; YANG, T.-J.; EMER, J. S. Efficient processing of deep neural networks: A tutorial and survey. **Proceedings of the IEEE**, Ieee, v. 105, n. 12, p. 2295–2329, 2017.

SZEGEDY, C.; LIU, W.; JIA, Y.; SERMANET, P.; REED, S.; ANGUELOV, D.; ERHAN, D.; VANHOUCHE, V.; RABINOVICH, A. Going deeper with convolutions. In: **Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)**. [S.l.: s.n.], 2015. p. 1–9.

SZEGEDY, C.; VANHOUCHE, V.; IOFFE, S.; SHLENS, J.; WOJNA, Z. Rethinking the inception architecture for computer vision. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2016. p. 2818–2826.

TAN, M.; LE, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In: PMLR. **International conference on machine learning**. [S.l.], 2019. p. 6105–6114.

TANG, W.; SUN, J.; WANG, S.; ZHANG, Y. Review of alexnet for medical image classification. **arXiv preprint arXiv:2311.08655**, 2023.

TASYUREK, M.; GUL, E. A new deep learning approach based on grayscale conversion and dwt for object detection on adversarial attacked images. **The Journal of Supercomputing**, Springer, v. 79, n. 18, p. 20383–20416, 2023.

TU, Z.; MILANFAR, P.; TALEBI, H. Muller: multilayer laplacian resizer for vision. In: **Proceedings of the IEEE/CVF International Conference on Computer Vision**. [S.l.: s.n.], 2023. p. 6877–6887.

UKWANDU, O.; HINDY, H.; UKWANDU, E. An evaluation of lightweight deep learning techniques in medical imaging for high precision covid-19 diagnostics. **Healthcare Analytics**, Elsevier, v. 2, p. 100096, 2022.

VARGAS, A. C. G.; PAES, A.; VASCONCELOS, C. N. Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. In: SN. **Proceedings of the xxix conference on graphics, patterns and images**. [S.l.], 2016. v. 1, n. 4.

VAZ, J. M.; BALAJI, S. Convolutional neural networks (cnns): concepts and applications in pharmacogenomics. **Molecular diversity**, Springer, v. 25, n. 3, p. 1569–1584, 2021.

WALLRAFF, S.; VESAL, S.; SYBEN, C.; LUTZ, R.; MAIER, A. Age estimation on panoramic dental x-ray images using deep learning. In: SPRINGER. **Bildverarbeitung für die Medizin 2021: Proceedings, German Workshop on Medical Image Computing, Regensburg, March 7-9, 2021**. [S.l.], 2021. p. 186–191.

WANG, J.; DOU, J.; HAN, J.; LI, G.; TAO, J. A population-based study to assess two convolutional neural networks for dental age estimation. **BMC Oral Health**, Springer, v. 23, n. 1, p. 109, 2023.

WANG, X.; LIU, Y.; MIAO, X.; CHEN, Y.; CAO, X.; ZHANG, Y.; LI, S.; ZHOU, Q. Densen: a convolutional neural network for estimating chronological ages from panoramic radiographs. **BMC bioinformatics**, Springer, v. 23, n. Suppl 3, p. 426, 2022.

WILIMITIS, D.; WALSH, C. G. Practical considerations and applied examples of cross-validation for model development and evaluation in health care: tutorial. **Jmir ai**, JMIR Publications Toronto, Canada, v. 2, p. e49023, 2023.

WILLEMS, G.; OLMEN, A. V.; SPIESSENS, B.; CARELS, C. Dental age estimation in belgian children: Demirjian's technique revisited. **Journal of forensic sciences**, ASTM International, v. 46, n. 4, p. 893–895, 2001.

WOHLIN, C.; RUNESON, P.; HOST, M.; OHLSSON, M. C.; REGNELL, B.; WESSLEN, A. **Experimentation in Software Engineering**. [S.l.]: Springer Berlin / Heidelberg, 2012.

WU, T.-J.; TSAI, C. L.; HUANG, Y.-H.; FAN, T.-Y.; CHEN, Y.-P. Efficacy of machine learning assisted dental age assessment in local population. **Legal Medicine**, Elsevier, v. 59, p. 102148, 2022.

XU, M.; YOON, S.; FUENTES, A.; PARK, D. S. A comprehensive survey of image augmentation techniques for deep learning. **Pattern Recognition**, Elsevier, v. 137, p. 109347, 2023.

YADAV, S. S.; JADHAV, S. M. Deep convolutional neural network based medical image classification for disease diagnosis. **Journal of Big data**, Springer, v. 6, n. 1, p. 1–18, 2019.

YEOM, H.-G.; LEE, B.-D.; LEE, W.; LEE, T.; YUN, J. P. Estimating chronological age through learning local and global features of panoramic radiographs in the korean population. **Scientific Reports**, Nature Publishing Group UK London, v. 13, n. 1, p. 21857, 2023.