



**UNIVERSIDADE FEDERAL DO CEARÁ
CURSO DE GRADUAÇÃO EM ENGENHARIA DE COMPUTAÇÃO**

IRIS CORDEIRO COSTA

**ANÁLISE COMPARATIVA DE ARQUITETURAS DE REDES NEURAIS
CONVOLUCIONAIS NA CLASSIFICAÇÃO ETÁRIA FORENSE A PARTIR DE
RADIOGRAFIAS PANORÂMICAS ODONTOLÓGICAS**

**SOBRAL
2025**

IRIS CORDEIRO COSTA

ANÁLISE COMPARATIVA DE ARQUITETURAS DE REDES NEURAIS
CONVOLUCIONAIS NA CLASSIFICAÇÃO ETÁRIA FORENSE A PARTIR DE
RADIOGRAFIAS PANORÂMICAS ODONTOLÓGICAS

Trabalho de Conclusão de Curso apresentado
ao Curso de Graduação em Engenharia de
Computação da Universidade Federal do
Ceará, como requisito parcial à obtenção do
grau de bacharel em Engenharia de Computação.

Orientador: Prof. Dr. Iális Cavalcante de
Paula Júnior

SOBRAL

2025

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

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

C872a Costa, Iris.

ANÁLISE COMPARATIVA DE ARQUITETURAS DE REDES NEURAIS CONVOLUCIONAIS NA
CLASSIFICAÇÃO ETÁRIA FORENSE A PARTIR DE RADIOGRAFIAS PANORÂMICAS
ODONTOLÓGICAS / Iris Costa. – 2025.

63 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Campus de Sobral,
Curso de Engenharia da Computação, Sobral, 2025.

Orientação: Prof. Dr. Iális Cavalcante de Paula Júnior.

1. Radiografias Panorâmicas. 2. Redes Neurais Convolucionais. 3. Odontologia Forense. 4. Classificação
etária. 5. Pré-processamento de Imagens. I. Título.

CDD 621.39

IRIS CORDEIRO COSTA

ANÁLISE COMPARATIVA DE ARQUITETURAS DE REDES NEURAIS
CONVOLUCIONAIS NA CLASSIFICAÇÃO ETÁRIA FORENSE A PARTIR DE
RADIOGRAFIAS PANORÂMICAS ODONTOLÓGICAS

Trabalho de Conclusão de Curso apresentado
ao Curso de Graduação em Engenharia de
Computação da Universidade Federal do Ceará,
como requisito parcial à obtenção do grau de
bacharel em Engenharia de Computação.

Aprovada em: 06/08/2025

BANCA EXAMINADORA

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

Prof. Dr. Marcelo Marques Simões de Souza
Universidade Federal do Ceará (UFC)

Prof. Dr. Evilásio Costa Júnior
Universidade Federal do Ceará (UFC)

AGRADECIMENTOS

Agradeço, primeiramente, a Deus por me guiar, iluminar e me permitir concluir este ciclo da minha vida acadêmica. Por sempre me dar forças para perseverar e continuar em busca dos meus objetivos, tanto pessoais quanto profissionais.

À minha família, em especial aos meus pais, Norma Suzana Melo Cordeiro Costa e Charly Pinto Costa, pelo amor incondicional, apoio constante e pelos ensinamentos preciosos que me orientarão por toda a vida. À minha irmã, Ingrid Cordeiro Costa, pela cumplicidade, apoio e companheirismo que sempre me deu. Sua presença, mesmo que de longe, ao meu lado tem sido uma fonte constante de força e carinho. À minha tia Maria das Graças Cordeiro, que, mesmo não estando mais entre nós, permanece viva em minhas lembranças e em meu coração. Sua crença em minha capacidade e a inspiração que sempre me deu para seguir em frente com força e sabedoria continuam a me guiar todos os dias.

Aos meus amigos de turma, João Paulo Militão e Jonas Fortes, que caminharam ao meu lado por quase todas as disciplinas, tornando os momentos de estudo e desafios mais leves e repletos de companheirismo.

Aos meus amigos de faculdade, Vitória Freitas, Joana Sthefanny, Jessica Marques, Nájala Kelly, Vinicius Tabosa, Wesley Fernandes, João Pedro Barbosa e Thainara Penha, que compartilharam comigo os desafios e as vitórias dessa jornada, tornando-a ainda mais especial.

Aos meus amigos e egressos do curso, Carlos Augusto, Andressa Gomes e Matheus Soares, que, em meio aos momentos mais desafiadores, sempre me ajudaram com sua generosidade e presença.

Aos meus amigos, Swanny Macêdo, Cândido Athayde, Clara Schmidt, Gabriela Mello e Lorena Veloso, que, apesar da distância física, nunca deixaram de me apoiar com suas palavras de incentivo e carinho.

Aos meus colegas de trabalho e amigos, Gisely Vasconcelos e Daniel D'alessandro, que acompanharam de perto toda a jornada do TCC. O apoio constante e a confiança no meu potencial foram fundamentais para que eu pudesse seguir em frente. A presença de vocês fez toda a diferença e sou muito grata por terem acreditado em mim.

À colega de faculdade e mestra Michele Faria de Oliveira, pela ajuda e experiência compartilhada na elaboração deste trabalho. Sua orientação foi de extrema importância para a conclusão desta pesquisa.

A todos aqueles que, de alguma forma, cruzaram meu caminho e contribuíram para

o meu crescimento, seja com ensinamentos, exemplos ou simplesmente com sua presença. Cada interação, cada experiência, ajudou a moldar a pessoa que sou hoje e a trajetória que percorri até aqui. Sou grata a todos que, direta ou indiretamente, influenciaram essa jornada.

Ao meu orientador, Prof. Dr. Iális Cavalcante, pela orientação, apoio e dedicação ao longo de todo este processo. Sua experiência e contribuição foram fundamentais para o aprimoramento e desenvolvimento deste trabalho.

Ao Programa de Educação Tutorial (PET), por me proporcionar experiências enriquecedoras que contribuíram diretamente para o meu crescimento acadêmico e pessoal.

Aos professores da Universidade Federal do Ceará (UFC), que, com seus ensinamentos, me proporcionaram uma formação sólida e de qualidade, me preparando para os desafios que virão.

RESUMO

A estimativa de idade com base em radiografias panorâmicas constitui uma prática consolidada na odontologia forense, sendo relevante em contextos como adoção, imigração, responsabilização penal e identificação de vítimas. Este trabalho propõe a utilização de redes neurais convolucionais (CNNs) aplicadas à tarefa de classificação etária forense a partir de ortopantomografias. Foram investigadas três arquiteturas de redes convolucionais utilizadas na literatura — VGG19, InceptionV3 e EfficientNetV2B0 — todas com pesos pré-treinados no ImageNet. O *pipeline* de preparação incluiu redimensionamento para 128×128 pixels, manutenção das imagens em RGB, normalização e aplicação de técnicas de *data augmentation*. A fim de explorar o impacto do pré-processamento das imagens no desempenho do modelo, oito variações de filtros foram consideradas, como filtros de borda (Sobel, Prewitt, Roberts, Laplacian), transformações morfológicas (*Top-hat* e *Black-hat*) e correção gama. O treinamento foi realizado por validação cruzada estratificada (*4-fold*), e as métricas de desempenho — *Accuracy*, *Precision*, *Recall* e *F1-score* — calculadas com base nas médias dos *folds*. Os resultados evidenciam variações de desempenho entre os filtros, indicando que determinadas técnicas de realce visual podem beneficiar ou comprometer a capacidade preditiva das CNNs. A comparação entre arquiteturas mostrou diferenças de sensibilidade ao pré-processamento, reforçando a importância dessa etapa em cenários clínicos e periciais.

Palavras-chave: Classificação Etária; Radiografias Panorâmicas; Redes Neurais Convolucionais; Odontologia Forense; Pré-processamento de Imagens.

ABSTRACT

Age estimation based on panoramic radiographs is a well-established practice in forensic dentistry, playing a key role in contexts such as adoption, immigration, criminal responsibility, and victim identification. This study proposed the use of convolutional neural networks (CNNs) applied to forensic age classification from orthopantomograms, exploring the impact of different image preprocessing strategies on model performance. Three architectures commonly employed in the literature were investigated — VGG19, InceptionV3, and EfficientNetV2B0 — all initialized with pre-trained ImageNet weights. The preprocessing pipeline included resizing to 128×128 pixels, maintaining images in RGB format, normalization, application of data augmentation techniques, and the use of eight processing variations, such as edge filters (Sobel, Prewitt, Roberts, Laplacian), morphological transformations (Top-hat and Black-hat), and gamma correction. Training was conducted using 4-fold stratified cross-validation, and performance metrics — accuracy, precision, recall, and F1-score — were computed based on the average across folds. The results revealed performance variations between filters, indicating that certain visual enhancement techniques can either benefit or impair the predictive capacity of CNNs. The comparative analysis among architectures demonstrated differences in sensitivity to preprocessing, underscoring the importance of this stage in clinical and forensic applications.

Keywords: Age Classification; Panoramic Radiographs; Convolutional Neural Networks; Forensic Dentistry; Image Preprocessing.

LISTA DE FIGURAS

Figura 1 – Estágios A–H do desenvolvimento dentário segundo o método de Demirjian.	21
Figura 2 – Radiografia panorâmica de uma criança de 8,9 anos conforme o método de Demirjian.	21
Figura 3 – Medidas realizadas no método de Cameriere.	22
Figura 4 – Radiografia panorâmica da arcada dentária.	24
Figura 5 – Sequência ilustrativa das etapas iniciais de pré-processamento.	26
Figura 6 – Visualização de uma radiografia panorâmica original e suas versões após aplicação dos filtros de pré-processamento considerados no estudo.	28
Figura 7 – Imagem original (ao centro) e variações geradas por técnicas de <i>data augmentation</i> , incluindo transformações geométricas e ajustes de intensidade.	30
Figura 8 – Neurônio biológico (A) e neurônio artificial (B) como unidade de processamento.	31
Figura 9 – Arquitetura de rede neural do tipo <i>feedforward</i> .	32
Figura 10 – Arquitetura típica de uma Rede Neural Convolucional, composta por camadas de convolução, ReLU, <i>pooling</i> , <i>flatten</i> e classificadores totalmente conectados.	35
Figura 11 – Curvas de aprendizado do modelo InceptionV3 com o filtro Gamma Correction.	47
Figura 12 – Curvas de aprendizado do modelo InceptionV3 com o filtro Top-hat (modo add).	48
Figura 13 – Curvas de aprendizado do modelo EfficientNetV2B0 com o filtro Gamma Correction.	50
Figura 14 – Curvas de aprendizado do modelo EfficientNetV2B0 com o filtro Subtract Prewitt.	50
Figura 15 – Curvas de aprendizado do modelo VGG19 com imagem original (sem filtros de pré-processamento).	52
Figura 16 – acurácia média dos modelos em função das técnicas de pré-processamento. As barras de erro representam o desvio-padrão obtido na validação cruzada (<i>4-fold</i>).	54

LISTA DE TABELAS

Tabela 1 – Desempenho médio do modelo InceptionV3 com diferentes técnicas de pré-processamento.	47
Tabela 2 – Desempenho médio do modelo EfficientNetV2B0 com diferentes técnicas de pré-processamento.	49
Tabela 3 – Desempenho médio do modelo VGG19 com diferentes técnicas de pré-processamento.	51
Tabela 4 – Filtros de pré-processamento que obtiveram o melhor desempenho em cada métrica por modelo. Os destaque em negrito representam os melhores resultados globais.	53

LISTA DE ABREVIATURAS E SIGLAS

BAA	<i>Bone Age Assessment</i>
CM	Método de Cameriere
CNNs	<i>Convolutional Neural Networks</i>
DL	<i>Deep Learning</i>
DM	Método de Demirjian
DSS	<i>Dental Scoring Systems</i>
GANs	<i>Generative Adversarial Networks</i>
IA	Inteligencia Artificial
ML	<i>Machine Learning</i>
MLPs	<i>Multilayer Perceptrons</i>
MP	McCulloch e Pitts
OPGs	Radiografias Panorâmicas Odontologicas
ReLU	<i>Rectified Linear Unit</i>
RF	<i>Random Forest</i>
RM	Ressonância Magnética
RNAs	Redes Neurais Articiais
RX	Raio-X
SVM	<i>Support Vector Machine</i>
TC	Tomografia Computadorizada
TCI	<i>Tooth Coronal Index</i>
US	Ultrassonografia

LISTA DE SÍMBOLOS

x_i	i -ésima entrada da rede neural ou razão entre abertura apical e comprimento do dente i
x_j	Saída do neurônio j após ativação
w_{ji}	Peso da conexão entre a entrada x_i e o neurônio j
a_j	Soma ponderada das entradas do neurônio j
b_j	Viés adicionado ao somatório no neurônio j
$\varphi(\cdot)$	Função de ativação aplicada ao somatório ponderado
X	Vetor de entrada com as variáveis independentes (camada de entrada)
Y	Valores desejados de saída (rótulos reais)
Y'	Valores de saída previstos pela rede
E	Erro entre os valores desejados e previstos: $E = Y - Y'$
Σ	Operador de somatório aplicado às entradas de um neurônio
$\sigma(x)$	Função sigmoide logística
$\tanh(x)$	Função tangente hiperbólica
$\text{ReLU}(x)$	Função de ativação <i>Rectified Linear Unit</i>
x_5	Razão entre abertura apical e comprimento do quinto dente (Cameriere)
N_0	Número de dentes com ápices completamente formados (Cameriere)
g	Sexo na equação de Cameriere (1 para meninos, 0 para meninas)
s	Soma das sete proporções normalizadas (Cameriere)
TP_c	Verdadeiros positivos da classe c
FP_c	Falsos positivos da classe c
FN_c	Falsos negativos da classe c
C	Número total de classes
n	Número total de amostras
$\mathbf{1}(\cdot)$	Função indicadora (retorna 1 se a condição for verdadeira, 0 caso contrário)

SUMÁRIO

1	INTRODUÇÃO	14
1.1	Justificativa	15
1.2	Objetivo Geral	15
1.3	Objetivos Específicos	15
2	TRABALHOS RELACIONADOS	17
3	FUNDAMENTAÇÃO TEÓRICA	19
3.1	Estimativa de Idade na Odontologia Forense	19
3.1.1	<i>Métodos Tradicionais</i>	19
3.1.2	<i>Radiografias Panorâmicas</i>	24
3.2	Processamento Digital de Imagens Médicas	25
3.2.1	<i>Pré-processamento das Imagens</i>	26
3.2.2	<i>imageNet</i>	26
3.2.3	<i>Técnicas de Realce de Imagem</i>	27
3.2.4	<i>Aumento de Dados (Data Augmentation)</i>	28
3.2.5	<i>Avaliação do Impacto dos Processamentos</i>	29
3.3	Redes Neurais Artificiais e Deep Learning	29
3.4	Redes Neurais Convolucionais	35
3.4.1	<i>Arquitetura Fundamental</i>	36
4	METODOLOGIA	38
4.1	Conjunto de dados	38
4.1.1	<i>Critérios de Inclusão e Exclusão</i>	38
4.1.2	<i>Limitações do Dataset</i>	39
4.2	Pré-processamento	39
4.3	Ambiente Experimental	40
4.4	Arquiteturas de Redes Neurais Convolucionais	41
4.5	Modelagem e Treinamento	42
4.6	Métricas de Avaliação	43
4.7	Extração e Visualização dos Resultados	44
5	RESULTADOS E DISCUSSÕES	46
5.1	Apresentação dos Resultados	46

5.1.1	<i>InceptionV3</i>	46
5.1.2	<i>EfficientNET</i>	48
5.1.3	<i>VGG19</i>	50
5.2	Análise Comparativa	52
6	CONCLUSÕES E TRABALHOS FUTUROS	56
6.1	Trabalhos Futuros	56
	REFERÊNCIAS	58

1 INTRODUÇÃO

A odontologia legal exerce um papel fundamental na identificação humana, especialmente em cenários adversos, como desastres em massa, investigações criminais e na identificação de indivíduos a partir de vestígios humanos, nos quais a identificação visual se torna inviável ou insuficiente. Nesse contexto, a estimativa da idade cronológica é essencial, na identificação de vítimas, a determinação da maioridade penal e a resolução de questões de ordem civil e jurídica (SHEN *et al.*, 2024). Tal estimativa constitui um dos pilares na definição da identidade biológica de um indivíduo (YEOM *et al.*, 2023).

Historicamente, a estimativa de idade em ambientes forenses tem sido realizada por meio da análise de características morfológicas dentárias e ósseas, com ênfase nas Radiografias Panorâmicas Odontológicas (OPGs), dada sua durabilidade frente a condições ambientais adversas e sua capacidade de registrar padrões dentários únicos (PRETTY; SWEET, 2001). Métodos clássicos, a exemplo dos propostos por Demirjian, Willems e Cameriere, além de atlas — compilações visuais do desenvolvimento dentário ao longo das idades — baseiam-se na avaliação dos estágios de mineralização e desenvolvimento dentário.

No entanto, essas abordagens manuais apresentam limitações relevantes, como a subjetividade da interpretação, o tempo necessário para análise especializada e a variabilidade nos resultados, fatores que podem comprometer a confiabilidade do resultado (YAVUZ *et al.*, 2025). Além disso, a precisão dessas técnicas pode ser afetada por aspectos populacionais, como etnia, exigindo adaptações metodológicas específicas para diferentes grupos (SHEN *et al.*, 2021).

Tais limitações expõem a necessidade de soluções mais automatizadas, objetivas e escaláveis. Nesse cenário, a Inteligência Artificial (IA) tem se destacado, em especial as técnicas baseadas em *Deep Learning* (DL), com ênfase nas Redes Neurais Convolucionais, do inglês *Convolutional Neural Networks* (CNNs). Essas arquiteturas vêm demonstrando grande eficácia na automatização da análise de imagens, contribuindo para a elevação da acurácia na estimativa de idade por meio de radiografias odontológicas. As técnicas de Deep Learning superam as técnicas clássicas de reconhecimento de padrões tanto em desempenho quanto em eficiência (HIYARI *et al.*, 2024). Sua capacidade de extrair, características relevantes – sejam elas locais ou globais – diretamente das imagens, sem a necessidade de medições manuais ou da extração ad hoc, confere às CNNs uma aplicabilidade notável nesse tipo de tarefa (ATAŞ *et al.*, 2023).

1.1 Justificativa

Os métodos computacionais de inteligência artificial aplicados em ciência forense tem ganhado destaque nos últimos anos (SALEHI *et al.*, 2023). As CNNs destacam-se particularmente no campo da odontologia, com ênfase na análise de radiografias panorâmicas (ONG *et al.*, 2024). Essas redes vêm sendo aplicadas com sucesso em tarefas como a estimativa automatizada de idade, trazendo avanços importantes para a prática forense.(KOCH *et al.*, 2025)

Pesquisas como as de Vila-Blanco *et al.* (2020) e Milošević *et al.* (2022) demonstram que, apesar da evolução nessa área, ainda persistem limitações, como o alto erro médio absoluto, a exclusão de indivíduos com lesões ou faixas etárias incompletas e a baixa representatividade dos dados. Logo, é relevante investigar alternativas para o aprimoramento dos modelos existentes, tanto em termos de desempenho quanto de generalização.

Este trabalho faz uma análise comparativa entre diferentes arquiteturas de redes neurais convolucionais aplicadas à classificação etária forense com base em radiografias panorâmicas odontológicas. A partir da avaliação do desempenho dos modelos em um conjunto de dados específico, busca-se compreender quais arquiteturas apresentam maior precisão e robustez frente aos desafios apontados na literatura. Além disso, a proposta de investigação inclui a análise da influência de diferentes técnicas de pré-processamento de imagens, como filtros de realce e normalização, no desempenho dos modelos. Assim, espera-se contribuir para o avanço da área, oferecendo subsídios que possam orientar futuras pesquisas e aplicações práticas na estimativa etária forense.

1.2 Objetivo Geral

O objetivo geral deste trabalho é realizar experimentos comparativos com diferentes arquiteturas de Redes Neurais Convolucionais e técnicas de pré-processamento de imagem, a fim de identificar a metodologia mais adequada para a classificação etária forense de indivíduos com base em imagens de radiografias panorâmicas odontológicas, considerando as características específicas do conjunto de dados selecionado.

1.3 Objetivos Específicos

Com base no objetivo geral, foram definidos os seguintes objetivos específicos, que orientam as etapas do estudo:

- Realizar uma análise das principais arquiteturas de CNNs utilizadas na literatura para tarefas de classificação de imagens, destacando suas características, vantagens e limitações.
- Treinar diferentes arquiteturas de CNNs utilizando o conjunto de dados de radiografias panorâmicas odontológicas selecionado para este estudo.
- Validar o desempenho dos modelos com as métricas: *acurácia, precision, recall* e *F1-score*.
- Comparar os desempenhos das diferentes arquiteturas, identificando aquela que melhor performar na classificação etária forense, de acordo com as características do dataset utilizado.

2 TRABALHOS RELACIONADOS

A estimativa da idade cronológica, amplamente empregada em contextos pediátricos e forenses, tem passado por uma transição significativa: de métodos manuais e laboriosos, como os atlas de Greulich-Pyle e Tanner-Whitehouse, para abordagens mais avançadas baseadas em inteligência artificial. Essa evolução visa mitigar as limitações associadas às técnicas tradicionais, como a subjetividade na interpretação e a variabilidade interobservador.

Nesse cenário, os primeiros avanços com aprendizado de máquina, do inglês *Machine Learning* (ML), mostraram-se promissores na otimização de metodologias consagradas na odontologia forense. Como demonstrado por Shen *et al.* (2021), algoritmos como *Random Forest* (RF) e *Support Vector Machine* (SVM), podem ser usados para aprimorar o método de Cameriere, resultando em uma redução expressiva no erro estimativo da idade dentária. De forma complementar, Shen *et al.* (2024) aplicou ML para prever idades legais críticas em adolescentes, contribuindo para a minimização da subjetividade nas avaliações convencionais. O estudo de Bizjak e Robić (2024), apresenta um modelo de deep learning para estimativa etária a partir de radiografias panorâmicas dentárias, além de destacar fatores que contribuíram para erros de previsão, como restaurações protéticas, reabsorção óssea e perda dentária.

No campo da estimativa da idade óssea, do inglês *Bone Age Assessment* (BAA), por meio de radiografias de mão e punho, os modelos DL promoveram uma mudança expressiva em termos de precisão e eficiência. O estudo de Iglovikov *et al.* (2018) propôs um *pipeline* robusto, utilizando redes U-Net para segmentação e arquiteturas do tipo VGG para a estimativa, obtendo resultados satisfatórios. Nessa mesma linha, a revisão sistemática conduzida por Pérez *et al.* (2025) destacou a superioridade das CNNs, especialmente quando integradas a estratégias de *Transfer Learning*, como no caso dos modelos VGG16 e EfficientNetV2B0. Apesar do desempenho elevado, os autores enfatizam a importância da calibração populacional e da padronização das imagens para garantir maior generalização dos modelos.

Na área odontológica, a aplicação do DL tem revolucionado a estimativa de idade com base em radiografias panorâmicas. Modelos personalizados, como o proposto por Koch *et al.* (2025), e redes específicas como a DENSEN desenvolvida por Wang *et al.* (2022), alcançaram elevados níveis de acurácia de amplas faixas etárias. Estudos como o de Alam *et al.* (2025) reforçam a eficácia de arquiteturas como *DenseNet*, enquanto Oliveira *et al.* (2024) adaptou a arquitetura InceptionV4 para a população brasileira. Além disso, o modelo *DentAge*, desenvolvido por Bizjak e Robić (2024), consolidou o potencial das técnicas de DL aplicadas à

odontologia forense.

Apesar dos avanços observados na literatura recente, algumas limitações ainda persistem. A capacidade de generalização dos modelos para diferentes populações, segue como um desafio relevante, especialmente em faixas etárias extremas e em contextos com dados escassos. Esse problema se agrava pela suscetibilidade dos algoritmos a vieses presentes no conjunto de dados, o que pode comprometer a acurácia em grupos menos representados.

Além disso, observa-se uma escassez de estudos que comparem, de forma sistemática, modelos de aprendizado profundo aplicados a diferentes modalidades de imagem, como radiografias dentárias e de mão/punho. Questões como a robustez clínica, a interpretabilidade dos algoritmos e a falta de bases de dados amplas e balanceadas continuam sendo obstáculos importantes. Para que esses sistemas avancem em direção a aplicações clínicas de fato viáveis, será necessário investir em processos de curadoria mais criteriosos, além de estratégias de validação mais amplas, que considerem a variabilidade populacional e as condições práticas de uso.

3 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta os fundamentos teóricos que sustentam a proposta deste estudo, com foco nos conceitos essenciais às técnicas empregadas. São exploradas, de forma articulada, as bases que orientam o uso de redes neurais convolucionais em tarefas de classificação etária forense, incluindo aspectos do processamento digital de imagens médicas e dos modelos de aprendizagem profunda. As seções que se seguem aprofundam os referenciais necessários para contextualizar a escolha metodológica e respaldar a análise dos resultados.

3.1 Estimativa de Idade na Odontologia Forense

A estimativa da idade cronológica constitui um dos pilares centrais em múltiplos contextos forenses e legais, abrangendo desde a identificação de indivíduos em cenários de desastres até a definição da imputabilidade penal de menores, além de sua aplicação em processos migratórios e procedimentos de adoção. Em situações nas quais não há documentação civil válida ou quando os registros de nascimento são inexistentes, a utilização de indicadores biológicos torna-se essencial para fundamentar decisões jurídicas e administrativas. Dentre os diversos marcadores de maturidade somática, o desenvolvimento dentário destaca-se por sua previsibilidade e robustez (SCHMELING *et al.*, 2016).

A importância dos dentes como indicadores biológicos de idade está associada à sua elevada resistência à degradação pós-morte, além de sua relativa independência frente a variáveis ambientais e genéticas. Em comparação ao processo de mineralização óssea, a formação dentária apresenta menor suscetibilidade a influências externas, o que contribui para sua confiabilidade como parâmetro cronológico. A sequência de desenvolvimento dos dentes, desde a formação inicial do germe até o fechamento completo do ápice radicular, ocorre de maneira ordenada e previsível, permitindo uma correlação consistente com a idade cronológica do indivíduo (VILA-BLANCO *et al.*, 2023).

3.1.1 Métodos Tradicionais

Historicamente, a estimativa da idade dental tem se baseado em métodos manuais, os quais consistem na análise dos estágios de desenvolvimento dentário ou na observação de características morfológicas visíveis em radiografias. Essas abordagens tradicionais vêm sendo amplamente validadas em diferentes populações ao redor do mundo. No entanto, reconhece-se

que a acurácia desses métodos pode variar entre distintos grupos étnicos, sendo suscetível à introdução de vieses, especialmente quando aplicados fora de seus contextos populacionais originais (VILA-BLANCO *et al.*, 2023).

No caso de crianças e jovens adultos, a estimativa da idade cronológica baseia-se, predominantemente, no desenvolvimento dentário, aproveitando a previsibilidade dos processos de erupção e mineralização dos dentes. Métodos que utilizam a erupção como referência podem ser úteis para avaliações rápidas, mas tendem a apresentar menor precisão, uma vez que tais mudanças ocorrem em janelas temporais relativamente curtas. Para contornar essa limitação, atlas dentários como o "London Atlas", proposto por AlQahtani *et al.* (2010), oferecem representações gráficas dos estágios de desenvolvimento e da erupção dentária, possibilitando uma comparação visual direta entre o estado atual da dentição e padrões de referência pré-estabelecidos.

Em um nível de melhor detalhamento, destacam-se os Sistemas de Pontuação Dental, do inglês *Dental Scoring Systems* (DSS), que classificam o desenvolvimento de cada dente em estágios específicos por pontuação. Entre os sistemas mais utilizados encontram-se os de Nolla, Gleiser e Hunt, Demirjian e Cameriere. Os dois últimos, por sua relevância e afinidade com o escopo deste estudo, merecem atenção especial.

O Método de Demirjian (DM) é uma abordagem amplamente empregada para estimativa de idade em indivíduos na infância e adolescência. Baseia-se na avaliação de oito estágios de desenvolvimento observáveis em sete dentes mandibulares, do lado esquerdo, a partir de radiografias panorâmicas. Os estágios (Figura 1) vão desde a formação da cripta dentária (estágio A) até o fechamento completo dos ápices radiculares (estágio H), cada um associado a uma pontuação específica (APAYDIN; YASAR, 2018).

A Figura 2 apresenta uma radiografia panorâmica real, na qual é possível observar os diferentes graus de desenvolvimento dentário em uma criança de 8,9 anos. A soma total dessas pontuações resulta na idade dental estimada.

Embora amplamente reconhecido por sua objetividade, o método apresenta acurácia sensível à população em que é aplicado. Em um estudo com crianças turcas, Apaydin e Yasar (2018) identificaram tendência à superestimação da idade, com desvio médio de 0,304 anos — resultado estatisticamente significativo, que reforça a necessidade de ajustes ao perfil populacional. Ainda que padronizado, o método exige atenção a fatores demográficos e biológicos, sobretudo em contextos forenses, nos quais pequenos desvios podem comprometer a interpretação pericial.

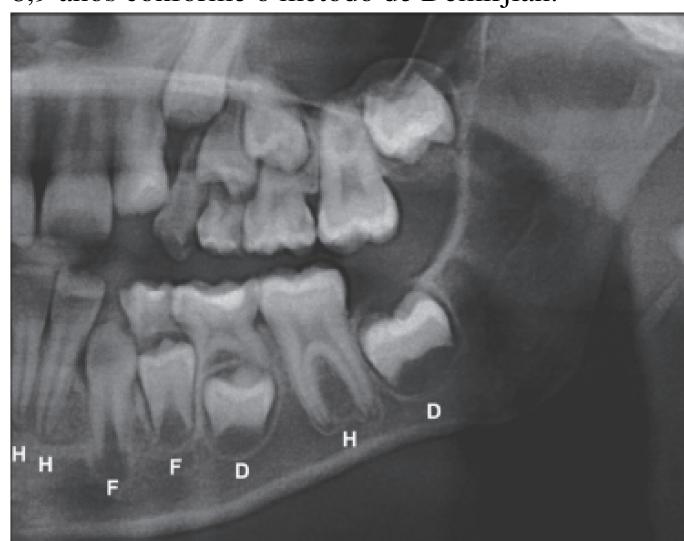
Em contraste, o Método de Cameriere (CM) adota uma abordagem métrica, funda-

Figura 1 – Estágios A–H do desenvolvimento dentário segundo o método de Demirjian.

	Molares	Pré-molares	Caninos	Incisivos
A				
B				
C				
D				
E				
F				
G				
H				

Fonte: Apaydin e Yasar (2018)

Figura 2 – Radiografia panorâmica de uma criança de 8,9 anos conforme o método de Demirjian.



Fonte: Apaydin e Yasar (2018)

mentada em medições diretas sobre radiografias panorâmicas. Ao contrário do DM, que utiliza critérios qualitativos, o CM quantifica a abertura apical e o comprimento dos dentes mandibulares permanentes do lado esquerdo, normalizando essas medidas pela altura do respectivo dente. Essas proporções normalizadas são representadas por $x_i = \frac{A_i}{L_i}$, onde A_i corresponde à distância entre as faces internas do ápice aberto e L_i ao comprimento total do dente. No caso de dentes com duas raízes, soma-se as distâncias dos dois ápices ($A_i = A_{i1} + A_{i2}$). A Figura 3 ilustra essas medições (ISHWARKUMAR *et al.*, 2022).

Os valores obtidos são então inseridos em uma equação de regressão linear que fornece a estimativa da idade dentária. Uma das fórmulas amplamente adotada, proposta por Cameriere e colaboradores, é expressa da seguinte forma:

$$\text{Idade} = 8,971 + 0,375g + 1,631x_5 + 0,674N_0 - 1,034s - 0,176sN_0 \quad (3.1)$$

em que g representa o sexo (1 para meninos, 0 para meninas), x_5 é a razão entre a abertura apical e o comprimento do quinto dente, N_0 é o número de dentes com ápices completamente formados, e s corresponde à soma das sete proporções normalizadas (ISHWARKUMAR *et al.*, 2022).

Embora apresente desempenho satisfatório em diversas populações, o método também é sensível ao perfil demográfico da amostra. No mesmo estudo, verificou-se tendência oposta à do DM, com subestimação da idade cronológica e desvio médio de $-0,580$ anos, novamente com significância estatística.

Figura 3 – Medições realizadas no método de Cameriere.



Fonte: Shen *et al.* (2021)

Para adultos, faixa etária em que os dentes já estão completamente formados, os métodos tradicionais concentram-se em alterações morfológicas associadas ao envelhecimento, visíveis em exames radiográficos. Entre as abordagens métricas, destaca-se a proposta de Kvaal *et al.* (1995), que estima a deposição de dentina secundária com base na redução da radiopacidade

da polpa dental, medida por meio de proporções lineares entre a estrutura dentária e o canal pulpar. Outra métrica relevante é o Índice Coronal do Dente, do inglês *Tooth Coronal Index* (TCI), proposto por Drusini (2008), que utiliza a razão entre a altura da coroa e da cavidade pulpar como estimativa da idade. Adicionalmente, critérios baseados na visibilidade do ligamento periodontal, na redução da câmara pulpar e em alterações como atrito oclusal, translucidez radicular e aposição de cimento vêm sendo empregados em sistemas de estadiamento, aplicáveis tanto em dentes extraídos como em radiografias (TIMME *et al.*, 2017; SCHMELING *et al.*, 2016).

Apesar de sua relevância histórica e ampla aplicação, os métodos tradicionais enfrentam limitações significativas, que motivam a busca por alternativas mais objetivas e escaláveis. Um dos principais entraves reside na subjetividade da análise: a interpretação visual dos estágios dentários depende da experiência do avaliador e da qualidade da imagem, o que pode comprometer a reproduzibilidade dos resultados e introduzir vieses. Soma-se a isso o alto custo de tempo e esforço associados à aplicação manual dessas técnicas, fator que dificulta sua utilização em larga escala, especialmente em situações como desastres em massa.(APAYDIN; YASAR, 2018)

Outro desafio recorrente é a variabilidade populacional. Métodos desenvolvidos a partir de amostras específicas nem sempre se traduzem precisos em outros contextos étnicos ou geográficos, sendo comum a ocorrência de superestimações ou subestimações sistemáticas. Tais limitações reforçam a importância de validações regionais e de metodologias mais flexíveis.(YAVUZ *et al.*, 2025)

As soluções baseadas em inteligência artificial têm ganhado destaque como alternativa promissora. Segundo Bizjak e Robić (2024), técnicas automatizadas demonstram potencial para superar barreiras operacionais, oferecendo ganhos em eficiência sem prejuízo da acurácia. Em particular, redes neurais convolucionais vêm se mostrando eficazes na estimativa de idade a partir de radiografias panorâmicas, conforme apontado por Koch *et al.* (2025). Tais redes possibilitam a extração automatizada de características morfológicas relevantes, reduzindo a dependência da interpretação humana e promovendo maior padronização. Nesse contexto, o uso de aprendizado profundo se apresenta como um caminho viável para superar os desafios impostos pelas abordagens convencionais — tema aprofundado nas próximas seções deste trabalho.

3.1.2 Radiografias Panorâmicas

As radiografias panorâmicas odontológicas, também denominadas ortopantomografias, são uma fonte de dados importante para a estimativa da idade dentária. Esse tipo de imagem bidimensional oferece uma visualização ampla da arcada dentária, permitindo a análise conjunta dos arcos superior e inferior, bem como de estruturas ósseas adjacentes, como mandíbula, fossa nasal e vértebras cervicais (JAYARAMAN *et al.*, 2018). A Figura 4 apresenta um exemplo típico desse tipo de imagem.

Figura 4 – Radiografia panorâmica da arcada dentária.



Fonte: Ong *et al.* (2024)

Sua utilização na estimativa etária se justifica por diversas razões. A primeira é que estas radiografias permitem examinar simultaneamente todos os dentes, em diferentes estágios de desenvolvimento e erupção, o que viabiliza a observação contínua de processos como mineralização, formação coronária e radicular, além da própria cronologia eruptiva — elementos diretamente relacionados à idade cronológica (BIRCHLER *et al.*, 2019).

Outro fator que contribui para seu uso é a acessibilidade. Trata-se de um exame rotineiro na prática odontológica, de baixo custo e fácil obtenção, o que favorece a formação de bases de dados extensas e diversificadas. Além disso, por se tratar de um procedimento não invasivo, que exige apenas uma única exposição para capturar toda a arcada, o exame é seguro e bem tolerado por pacientes de diferentes faixas etárias (KIM *et al.*, 2023).

A extensa disponibilidade de registros históricos provenientes de contextos clínicos favorece o desenvolvimento de modelos computacionais. O volume expressivo de imagens já digitalizadas permite o treinamento de algoritmos de inteligência artificial capazes de interpretar características morfológicas de forma automatizada, indo além das medições pontuais utilizadas nos métodos tradicionais (KOKOMOTO *et al.*, 2024).

Cabe destacar, a versatilidade diagnóstica dessas imagens, que podem ser aplicadas tanto na avaliação do desenvolvimento dentário de crianças e adolescentes quanto na identificação de alterações associadas ao envelhecimento em adultos — como a aposição de dentina ou modificações no osso alveolar (LEE *et al.*, 2021).

Dessa forma, as ortopantomografias consolidam-se como recurso estratégico em estimativa etária, sendo úteis tanto em abordagens manuais, conduzidas por especialistas, quanto em soluções baseadas em aprendizado profundo.

3.2 Processamento Digital de Imagens Médicas

O processamento digital de imagens médicas consolida-se como um dos pilares da medicina de precisão, ao empregar técnicas de *machine learning* e *deep learning* para extrair informações quantitativas e clinicamente relevantes a partir de dados visuais complexos, como os provenientes de exames de Ressonância Magnética (RM), Raio-X (RX), Tomografia Computadorizada (TC) e Ultrassonografia (US) (NARIN *et al.*, 2021).

Diferentemente da interpretação manual, caracterizada por subjetividade e maior demanda de tempo, as abordagens computacionais permitem automatizar processos críticos para a análise e interpretação diagnóstica. Entre essas abordagens, destacam-se a classificação de patologias, a segmentação de estruturas anatômicas, o registro multimodal e a geração de imagens sintéticas (SHEN *et al.*, 2021). Tais processos não apenas ampliam a acurácia dos modelos, como também contribuem para otimizar o fluxo de trabalho clínico, promovendo maior eficiência e padronização na prática médica.

Segundo Galić *et al.* (2023), a base conceitual que sustenta esses avanços abrange diferentes paradigmas de aprendizado de máquina. O aprendizado supervisionado é amplamente utilizado para treinar modelos voltados à classificação e segmentação de imagens, enquanto o não supervisionado é voltado à identificação de padrões latentes e a detecção de anomalias em grandes volumes de dados. O aprendizado por reforço tem sido explorado na otimização de estratégias de aquisição e reconstrução de imagens.

A implementação desses paradigmas é viabilizada, em grande parte, por arquiteturas baseadas em redes neurais profundas. Nesse contexto, se destacam as Redes Neurais Convolucionais, cuja estrutura é particularmente adequada à análise de dados organizados em grade — como é o caso das imagens médicas. As Redes Geradoras Adversariais, do inglês *Generative Adversarial Networks* (GANs), também tem sido aplicadas com sucesso na síntese e reconstrução

de imagens médicas realistas, contribuindo para o aprimoramento de sistemas diagnósticos e a geração de dados sintéticos em contextos com escassez de amostras.

3.2.1 Pré-processamento das Imagens

O pré-processamento de imagens médicas busca garantir que os dados de entrada estejam compatíveis com os requisitos dos algoritmos de aprendizado profundo. Seu principal objetivo é padronizar o conteúdo visual, reduzir ruídos e realçar estruturas anatômicas relevantes, aumentando o potencial discriminativo dos modelos (RAYED *et al.*, 2024), como representado na Figura 5.

Figura 5 – Sequência ilustrativa das etapas iniciais de pré-processamento.



Fonte: Autoral

É prática comum submeter as imagens a etapas de redimensionamento, normalização dos valores de pixel e ajuste no espaço de cor em aplicações com modelos *deep learning*. Essas transformações devem adequar os dados de entrada às exigências das arquiteturas convolucionais, especialmente quando se utilizam modelos pré-treinados em grandes bases de dados, como o ImageNet (GE *et al.*, 2022; REINHOLD *et al.*, 2019).

Complementando esse processo, técnicas de realce — como filtros de gradiente e operadores morfológicos — podem ser aplicadas para destacar bordas e estruturas sutis, facilitando a identificação de regiões de interesse ainda nas primeiras camadas convolucionais (ANWAR *et al.*, 2018).

3.2.2 *imageNet*

O ImageNet é uma vasta base de dados de imagens anotadas, amplamente utilizada no treinamento e avaliação de modelos de redes neurais profundas, especialmente para tarefas de

classificação de imagens. A importância do ImageNet no campo de aprendizado profundo foi destacada quando se tornou um benchmark crucial para avaliar o desempenho de redes neurais convolucionais (*CNNs*) em tarefas de visão computacional.

O estudo de Recht *et al.* (2019) explora a capacidade de generalização dos classificadores treinados no ImageNet, investigando se esses modelos realmente conseguem generalizar para novas imagens e cenários fora do conjunto de dados de treinamento. Os autores discutem as limitações do ImageNet como um conjunto de dados representativo para o mundo real, questionando a eficácia dos classificadores treinados exclusivamente nele, especialmente em contextos que envolvem novas classes ou variações das imagens já presentes no conjunto de dados.

O ImageNet tem sido um pilar importante para o avanço de arquiteturas de redes neurais profundas, como o AlexNet, VGGNet e ResNet, que foram inicialmente avaliadas e comparadas com base no desempenho em sua competição anual, o *ImageNet Large Scale Visual Recognition Challenge*. O conjunto de dados contém milhões de imagens de mais de mil classes, e seu uso no treinamento de redes tem sido fundamental para o desenvolvimento de modelos de aprendizado profundo de alta precisão para tarefas de reconhecimento de imagens.

3.2.3 Técnicas de Realce de Imagem

Em imagens de ortopantomografia, o realce de contraste tem o objetivo de destacar estruturas sutis, como contornos dentários, canais radiculares e espaços pulparos. Essa operação pode ser realizada por filtros de borda baseados em gradiente, como Sobel, Prewitt, Roberts e Laplaciano, cada qual com propriedades específicas no que diz respeito à direção e intensidade da detecção (NASER *et al.*, 2024; ZANGANA *et al.*, 2024).

Enquanto Sobel e Prewitt enfatizam transições horizontais e verticais de intensidade, o filtro de Roberts é mais sensível a detalhes finos em direções diagonais, sendo útil em imagens com menor resolução. Por atuar com a segunda derivada, o filtro Laplaciano evidencia contornos com alta curvatura, embora demande cautela quanto à amplificação de ruídos.

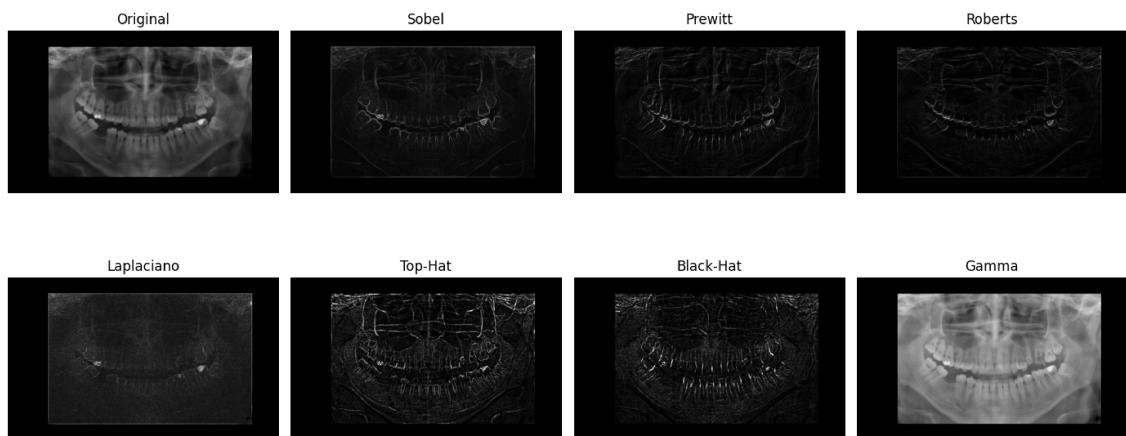
Transformações morfológicas, como *Top-Hat* e *Black-Hat*, também são amplamente utilizadas para ressaltar regiões mais claras ou mais escuras em relação ao seu entorno imediato. Essas técnicas operam com base em elementos estruturantes e têm aplicação consolidada na detecção de feições anatômicas específicas em exames radiográficos (KUSHOL *et al.*, 2019).

Outro recurso empregado é a correção gama, que realiza ajustes globais de contraste por meio de uma transformação não linear dos níveis de intensidade. Essa técnica é útil para

compensar exposições inadequadas ou evidenciar faixas específicas do histograma da imagem (SOMASUNDARAM; KALAVATHI, 2011).

A Figura 6 ilustra os diferentes efeitos visuais produzidos por cada filtro de pré-processamento aplicado às imagens radiográficas, evidenciando variações no realce de bordas e contraste.

Figura 6 – Visualização de uma radiografia panorâmica original e suas versões após aplicação dos filtros de pré-processamento considerados no estudo.



Fonte: Autoral

3.2.4 Aumento de Dados (Data Augmentation)

O aumento de dados é uma técnica empregada no treinamento de redes neurais convolucionais quando há escassez de amostras rotuladas, como é o caso da radiologia e da odontologia forense. Essa consiste em gerar variações sintéticas de imagens reais, mantendo sua semântica, mas introduzindo pequenas perturbações que ampliam o espaço amostral (WANG *et al.*, 2024).

Transformações como rotações, espelhamentos, deslocamentos e zooms têm o objetivo de simular variações naturais na aquisição das imagens — como diferenças de posicionamento ou de parâmetros técnicos do exame —, contribuindo para um modelo mais generalizável (KIM; BAE, 2020).

Nesse contexto, estratégias de aumento de dados, como rotações, deslocamentos e espelhamentos, têm se mostrado eficazes para ampliar o conjunto de treinamento e favorecer a capacidade de generalização dos modelos. Em exames clínicos, essas transformações também ajudam a simular variações comuns entre pacientes, como diferenças de posicionamento ou de parâmetros técnicos na aquisição das imagens (SALEHI *et al.*, 2023; TAYLOR; NITSCHKE,

2018).

Estratégias mais sofisticadas baseadas em redes geradoras, como as GANs, vêm sendo exploradas para a criação de imagens sintéticas realistas, ampliando ainda mais a diversidade do conjunto de treino. Estudos apontam que essas imagens podem complementar as técnicas tradicionais e melhorar o desempenho de modelos em tarefas complexas, como a detecção de lesões ou a classificação etária (FRID-ADAR *et al.*, 2018; PEREZ; WANG, 2017).

As transformações mais usuais para o aumento de dados estão dispostas na Figura 7. Ao centro se encontra radiografia panorâmica original, cercada por suas variações sintéticas. Para produzir cada variação ilustrada, foram realizadas na imagem original as operações de rotação, espelhamento horizontal, translações, zoom, alterações de brilho e contraste, além de deformações do tipo shear. Assim, as operações de *data augmentation* contribui para a diversidade do conjunto de treino sem comprometer a integridade da informação anatômica presente na imagem.

3.2.5 Avaliação do Impacto dos Processamentos

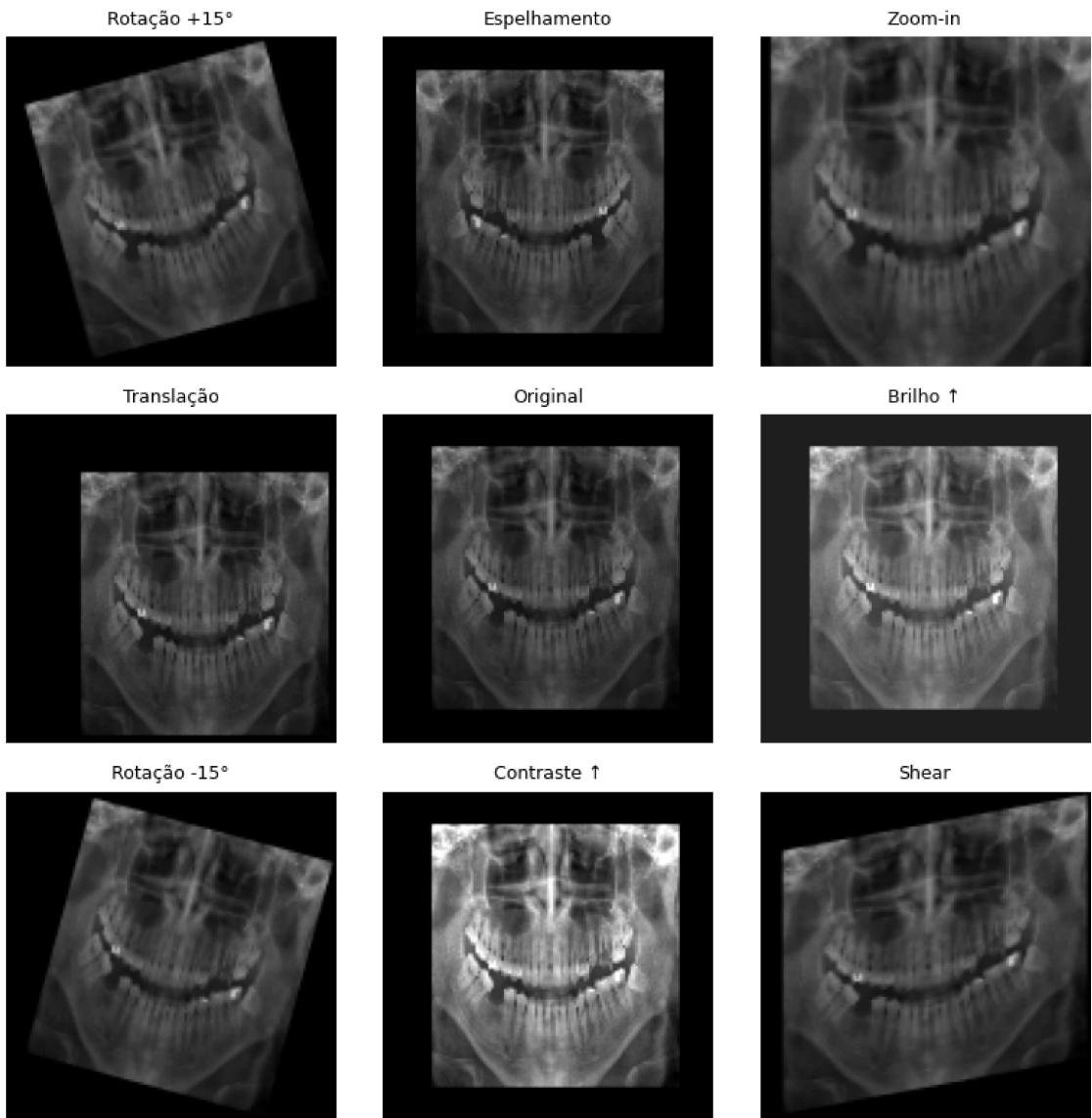
Diversas pesquisas vêm demonstrando que a qualidade visual e a padronização das imagens tem impacto no desempenho de modelos de classificação baseados em *deep learning*. Fatores como presença de ruído, contraste inadequado e heterogeneidade entre exames podem prejudicar a extração de características relevantes pelas redes convolucionais (LITJENS *et al.*, 2017).

Técnicas de pré-processamento como normalização, realce de bordas e correções de intensidade são incorporadas com o intuito de filtrar ruídos o conteúdo visual e enfatizar regiões de interesse anatômico (VU, 2024). Analisar o impacto dessas transformações permite fundamentar escolhas metodológicas mais adequadas e contribui para a construção de *pipelines* mais robustos, com maior capacidade de generalização e replicabilidade em contextos clínicos.

3.3 Redes Neurais Artificiais e Deep Learning

As Redes Neurais Artificiais (RNAs) são modelos computacionais inspirados no sistema nervoso biológico, nos quais neurônios artificiais interagem para processar informações de forma a aprender padrões a partir de dados empíricos, sem depender de regras explicitamente programadas. A Figura 8 ilustra, de forma esquemática, a comparação entre um neurônio

Figura 7 – Imagem original (ao centro) e variações geradas por técnicas de *data augmentation*, incluindo transformações geométricas e ajustes de intensidade.

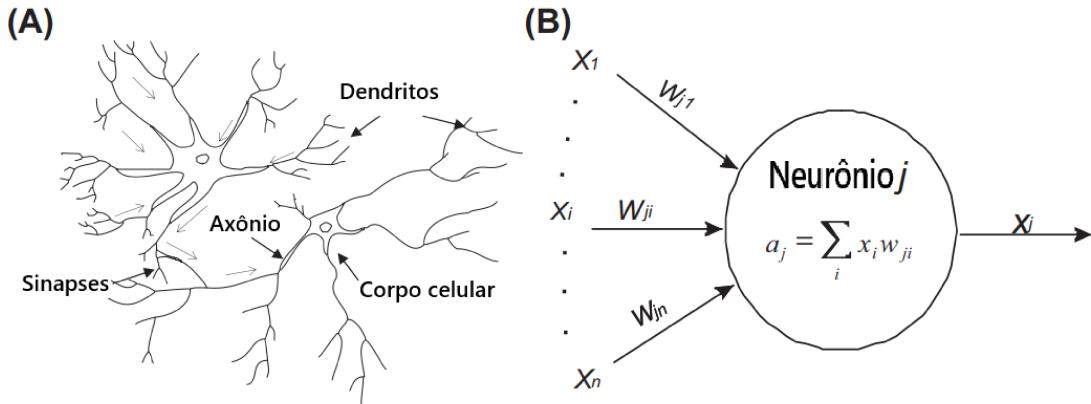


Fonte: Autoral

biológico e seu análogo artificial.

A Figura 8(B) representa um neurônio artificial do tipo McCulloch e Pitts (MP), amplamente empregado em redes neurais artificiais. Nessa representação, cada elemento possui uma função específica no processamento das informações. As entradas da rede, indicadas por x_i , correspondem a valores numéricos que podem ser provenientes tanto dos dados originais quanto de camadas anteriores da rede. Essas entradas são conectadas ao neurônio j por meio de pesos w_{ji} , que regulam a influência de cada entrada na ativação do neurônio. O valor a_j , obtido pela soma ponderada das entradas, segundo a equação $a_j = \sum_i x_i w_{ji}$, representa o potencial de ativação do neurônio antes da aplicação de uma função não linear. Em muitos modelos, também é

Figura 8 – Neurônio biológico (A) e neurônio artificial (B) como unidade de processamento.



Fonte: Park e Lek (2016) (adaptado)

adicionado um termo de viés b_j , que contribui para o deslocamento da função de ativação. Após essa etapa, o resultado da função de ativação gera a saída do neurônio x_j , que será transmitida para a próxima camada da rede. As setas presentes na figura ilustram o fluxo unidirecional de informação entre os neurônios, característica típica das redes do tipo *feedforward*.

A operação completa de saída do neurônio pode ser expressa pela equação:

$$x_j = \varphi \left(\sum_i x_i w_{ji} + b_j \right), \quad (3.2)$$

em que $\varphi(\cdot)$ representa a função de ativação, responsável por introduzir não linearidade ao modelo, e b_j corresponde ao viés adicionado à combinação linear das entradas.

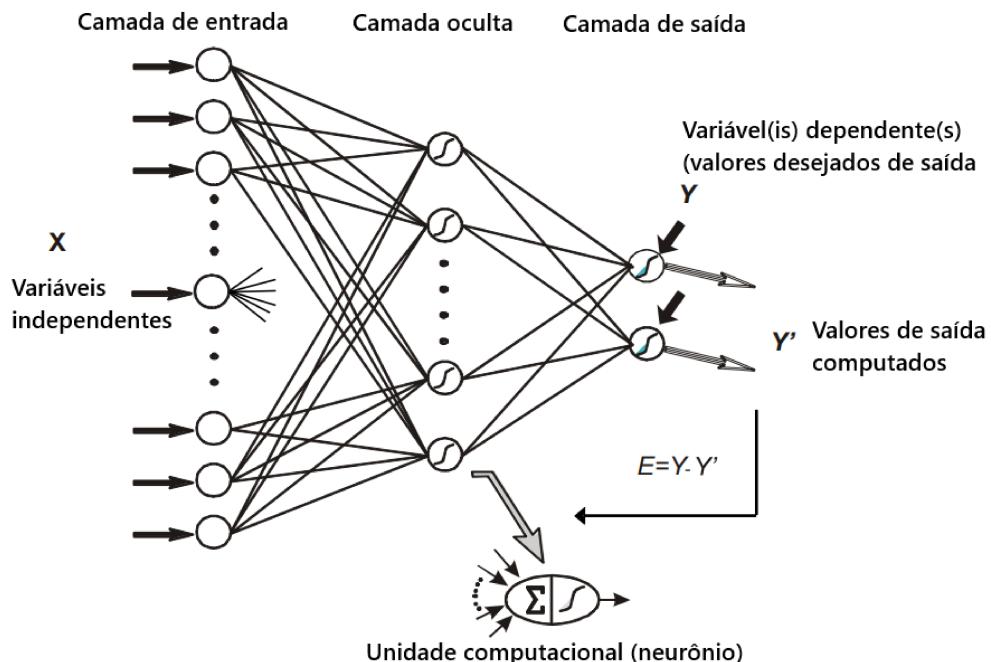
A unidade fundamental dessas redes é o neurônio artificial, também conhecido como unidade de processamento. Cada neurônio recebe múltiplos sinais de entrada, atribui um peso a cada um deles, soma os resultados e aplica uma função de ativação não linear, responsável por introduzir complexidade ao modelo. O valor final dessa função é transmitido à próxima camada da rede, formando um sistema de processamento em múltiplos estágios. Ainda que simplificado em relação ao neurônio biológico, esse modelo se mostrou eficaz para a representação de funções arbitrárias e aprendizado supervisionado (NIELSEN, 2015).

O primeiro modelo formal de redes neurais foi o Perceptron, proposto por Frank Rosenblatt em 1958. Trata-se de uma rede de camada única capaz de resolver apenas problemas linearmente separáveis, ou seja, aqueles em que uma linha ou hiperplano pode dividir as classes. A notoriedade do Perceptron, entretanto, entrou em declínio quando se revelou sua incapacidade de representar funções como o operador lógico XOR, limitação apontada por Minsky e Papert

em 1969 — o que desencadeou um período de ceticismo na área (SCHMIDHUBER, 2015).

O ressurgimento das redes neurais ocorreu com o desenvolvimento das chamadas Redes Neurais Multicamadas (ou *Multilayer Perceptrons* (MLPs)), compostas por uma ou mais camadas intermediárias entre as entradas e as saídas. A popularização do algoritmo de retropropagação do erro (*backpropagation*) foi determinante nesse processo, pois tornou viável o ajuste eficiente dos pesos sinápticos por meio do cálculo do gradiente do erro em relação à saída esperada (DENG *et al.*, 2014). Esse algoritmo baseia-se na aplicação da regra da cadeia do cálculo diferencial, permitindo que o erro da saída final seja propagado de volta através das camadas da rede. Com isso, é possível calcular a contribuição de cada peso no erro total e atualizá-los de forma a minimizar essa discrepância por meio de métodos como o gradiente descendente. A combinação entre o uso de múltiplas camadas, funções de ativação não lineares e o treinamento via *backpropagation* viabilizou o aprendizado de representações mais abstratas e complexas a partir dos dados.

Figura 9 – Arquitetura de rede neural do tipo *feedforward*.



Fonte: Park e Lek (2016) (adaptado)

A Figura 9 representa uma arquitetura típica de rede neural do tipo *feedforward*, composta por três tipos de camadas: entrada, ocultas e saída. A camada de entrada é responsável por receber as variáveis independentes do problema, representadas pelo vetor **X**, que alimenta os neurônios da primeira camada da rede. Em seguida, as camadas ocultas, formadas por

múltiplos neurônios interconectados, realizam transformações não lineares sobre os dados. O número de camadas ocultas pode variar conforme a complexidade do modelo, permitindo à rede extrair padrões progressivamente mais abstratos. Já a camada de saída fornece a predição final da rede, produzindo os valores computados Y' , os quais são comparados com os valores desejados Y durante o treinamento supervisionado. A diferença entre esses dois vetores gera o erro $E = Y - Y'$, que orienta o processo de ajuste dos pesos sinápticos da rede.

Cada neurônio dentro dessa estrutura é representado como uma unidade computacional que realiza duas operações fundamentais. Primeiramente, calcula-se o somatório ponderado das entradas, seguido da aplicação de uma função de ativação não linear, que introduz a capacidade da rede de modelar relações complexas nos dados. A escolha dessa função é essencial para o desempenho do modelo, pois influencia diretamente a forma como os padrões são aprendidos e propagados pelas camadas da rede.

Entre as funções de ativação mais utilizadas, destaca-se a tangente hiperbólica, definida no intervalo $[-1, 1]$, que possui a propriedade de centralizar os dados em torno de zero. Essa característica pode contribuir para uma convergência mais rápida durante o treinamento. A equação dessa função é dada por:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \quad (3.3)$$

Apesar dessas vantagens, a tangente hiperbólica pode sofrer com o problema dos gradientes pequenos, o que dificulta o ajuste dos pesos em redes profundas (FACELI *et al.*, 2011).

A função sigmoide logística, por sua vez, também é uma função suave e amplamente empregada, mapeando os valores de entrada para o intervalo $[0, 1]$. Sua equação é expressa por:

$$\sigma(x) = \frac{1}{1 + e^{-x}}. \quad (3.4)$$

Assim como a tangente hiperbólica, a função sigmoide também está sujeita ao problema dos gradientes desaparecendo, uma vez que seus valores de derivada se aproximam de zero nas regiões extremas, o que pode comprometer o treinamento de camadas mais profundas (BENGIO *et al.*, 2017).

Por outro lado, a função *Rectified Linear Unit* (ReLU) (*Rectified Linear Unit*) tornou-se a mais utilizada em redes neurais profundas devido à sua simplicidade computacional e eficácia na propagação do gradiente. Sua equação é dada por:

$$\text{ReLU}(x) = \begin{cases} 0, & \text{se } x \leq 0 \\ x, & \text{se } x > 0 \end{cases}. \quad (3.5)$$

A principal vantagem da ReLU está no fato de não saturar para valores positivos, o que evita o enfraquecimento do gradiente e acelera o processo de aprendizado. No entanto, essa função também possui limitações, como o fenômeno dos chamados "neurônios mortos", que ocorre quando determinadas unidades passam a retornar zero constantemente, impedindo sua atualização durante o treinamento (FERREIRA, 2024).

Essas funções impactam diretamente a dinâmica de propagação do gradiente e, portanto, o desempenho do processo de aprendizado em redes profundas.

Os modelos MLP são do tipo *feedforward*, nos quais o fluxo de informação ocorre em uma única direção, da entrada até a saída, sem ciclos ou retroalimentação. A inclusão de múltiplas camadas ocultas com funções de ativação não lineares, como a função ReLU, permite que a rede aprenda representações cada vez mais abstratas e não triviais dos dados, o que torna possível a resolução de tarefas complexas como reconhecimento de imagens, detecção de anomalias e análise de linguagem (LECUN *et al.*, 2015).

Esse princípio fundamenta o chamado aprendizado profundo (*deep learning*), caracterizado pelo uso de arquiteturas com muitas camadas (ou seja, “profundas”). Uma das maiores inovações dessa abordagem é a capacidade de realizar *representation learning*, ou aprendizado de representações: em vez de depender de atributos pré-definidos pelo ser humano, a rede aprende automaticamente quais características são mais relevantes para a tarefa (GOODFELLOW *et al.*, 2016).

O sucesso das redes profundas deve-se, também, ao avanço de outros fatores complementares: o aumento do poder computacional, especialmente com o uso de GPUs, a disponibilidade de grandes bases de dados rotuladas, e o aperfeiçoamento de algoritmos de otimização e funções de ativação. Com isso, o *deep learning* passou a liderar o estado da arte em áreas como visão computacional, processamento de linguagem natural e, mais recentemente, aplicações biomédicas (SHRESTHA; MAHMOOD, 2019).

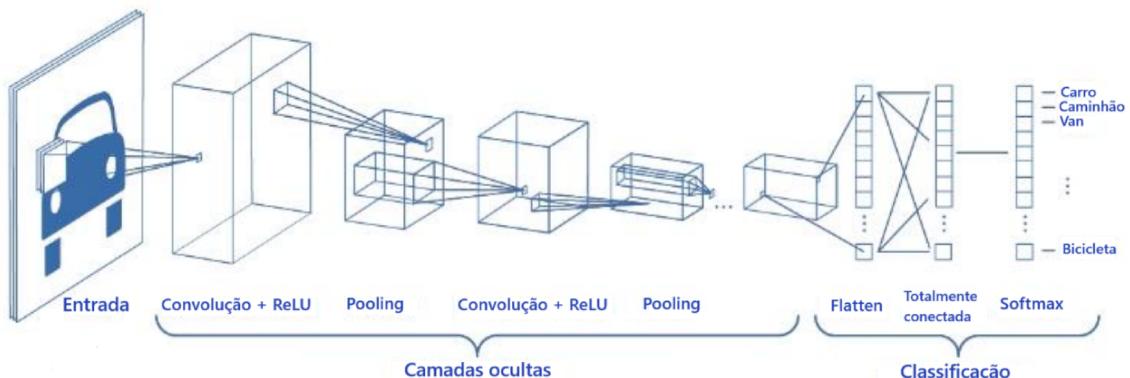
3.4 Redes Neurais Convolucionais

As redes convolucionais constituem uma classe de arquiteturas projetadas especificamente para o processamento de dados com estrutura de grade, como imagens bidimensionais, vídeos e sinais temporais. Inspiradas na organização do córtex visual de mamíferos, essas redes operam por meio de camadas hierárquicas que extraem padrões locais de forma progressiva e invariantes a pequenas variações espaciais. Em contraste com os modelos tradicionais, que tratam os dados de entrada de forma totalmente conectada, as CNNs exploram o princípio da localidade e utilizam o compartilhamento de pesos para reduzir drasticamente o número de parâmetros e aumentar a eficiência computacional (AREL *et al.*, 2010).

Esse tipo de arquitetura tem se consolidado como referência em tarefas de reconhecimento visual, incluindo classificação de imagens médicas, detecção de objetos e reconhecimento facial. Sua estrutura modular, composta por camadas convolucionais, funções de ativação, camadas de *pooling* e normalização, permite uma extração progressiva de características — desde bordas simples até composições morfológicas mais complexas —, tornando essas soluções particularmente eficazes para análises baseadas em imagem (TAYE, 2023).

A Figura 10 ilustra uma arquitetura típica de CNN, com a sua estrutura modular. A imagem de entrada é processada através de múltiplas camadas convolucionais com funções de ativação *ReLU*, seguidas por operações de *pooling* que reduzem a dimensionalidade espacial. Após essa etapa, os mapas de ativação são convertidos em vetores por meio de uma camada *flatten*, que os encaminha para camadas totalmente conectadas, responsáveis pela tomada de decisão final. A saída é, então, interpretada por uma função de ativação como a *softmax*, que atribui probabilidades às classes predefinidas.

Figura 10 – Arquitetura típica de uma Rede Neural Convolucional, composta por camadas de convolução, ReLU, pooling, flatten e classificadores totalmente conectados.



Fonte: Purwono *et al.* (2022) (adaptado)

3.4.1 Arquitetura Fundamental

A arquitetura de uma rede neural convolucional é composta por diferentes camadas organizadas de forma hierárquica, cada qual desempenhando uma função específica na extração, transformação e abstração de padrões visuais. O processamento tem início com as camadas convolucionais, que aplicam filtros (*kernels*) treináveis sobre a entrada, seja uma imagem ou os *feature maps* gerados por camadas anteriores. Cada filtro percorre a entrada por meio de janelas deslizantes, realizando operações de convolução que produzem mapas de ativação (*feature maps*), nos quais padrões locais, como bordas, texturas e formas geométricas são realçados (GOODFELLOW *et al.*, 2016).

Essa abordagem preserva a estrutura espacial da imagem, permitindo à rede reconhecer padrões mesmo quando deslocados, graças a mecanismos como *stride* e *padding* (RAWAT; WANG, 2017). Além de promover uma modelagem eficiente, esse tipo de camada reduz drasticamente o número de parâmetros, já que os pesos são compartilhados ao longo da imagem. A inspiração biológica dessa organização remete ao modelo Neocognitron, proposto por Fukushima (1980), no qual já se alternavam etapas de convolução e agrupamento hierárquico de padrões visuais.

Em sequência, os valores gerados pela convolução são processados por funções de ativação não lineares, que permitem à rede representar relações complexas entre os dados. (KRIZHEVSKY *et al.*, 2012). Variantes como a Leaky ReLU foram propostas para contornar limitações da versão padrão, mas a ReLU segue como escolha predominante por seu bom desempenho prático em redes profundas (GU *et al.*, 2018).

As camadas de *pooling*, que aparecem após as ativações, realizam uma forma de amostragem espacial dos mapas de ativação. A técnica mais comum é o *max-pooling*, que seleciona o valor máximo dentro de janelas fixas (como 2×2), reduzindo a dimensionalidade dos mapas e destacando as características mais salientes. Essa operação torna o modelo mais robusto a pequenas variações de posição, além de reduzir o custo computacional e o risco de sobreajuste. Arquiteturas pioneiras como a *LeNet-5* já exploravam com sucesso essa combinação entre convolução e *pooling* para reconhecimento de dígitos manuscritos (LECUN *et al.*, 2002), e essa estratégia permanece vigente em redes modernas.

Ao final da extração de características, os *feature maps* são achatados e repassados a camadas densamente conectadas (*fully connected layers*), que integram todas as ativações anteriores em representações globais. Cada neurônio dessas camadas recebe informações de

todos os neurônios da camada anterior, formando uma matriz de pesos completa. É nessa etapa que a rede realiza a tarefa de classificação propriamente dita, geralmente utilizando uma função *softmax* na última camada para converter os valores em probabilidades associadas às classes possíveis. Essa organização foi crucial para o sucesso de arquiteturas como a AlexNet, que marcou uma virada histórica em competições de classificação de imagens (KRIZHEVSKY *et al.*, 2012).

A combinação dessas camadas — convolucional, ativação, *pooling* e densamente conectada — constitui a base estrutural das CNNs modernas. Essa arquitetura permite capturar desde padrões visuais elementares até abstrações altamente complexas, tornando-a adequada a uma ampla variedade de tarefas, especialmente na análise de imagens médicas e forenses, onde a extração automática e hierárquica de características é particularmente vantajosa (AREL *et al.*, 2010).

4 METODOLOGIA

Este capítulo descreve os procedimentos metodológicos adotados para o desenvolvimento deste estudo. São apresentados, de forma estruturada, o conjunto de dados utilizado, as etapas de pré-processamento, as configurações experimentais, as estratégias de modelagem e treinamento, bem como as métricas e métodos empregados para extração e análise dos resultados. Além disso, detalham-se as arquiteturas convolucionais exploradas — VGG19, InceptionV3 e EfficientNetV2B0 — selecionadas com base em sua relevância na literatura e diversidade estrutural.

4.1 Conjunto de dados

O conjunto de dados utilizado neste trabalho é composto por radiografias panorâmicas odontológicas provenientes da Clínica de Odontologia da Universidade Federal do Ceará, campus Sobral. A coleta foi viabilizada por meio de parceria institucional entre os cursos de Engenharia da Computação e Odontologia, que possibilitou o acesso aos exames por meio do sistema clínico digital da unidade.

As imagens encontram-se originalmente em formato .jpg, totalizando 868 arquivos. Cada exame foi exportado individualmente com auxílio da equipe técnica, sendo posteriormente anonimizado: os identificadores pessoais foram removidos e substituídos por códigos alfanuméricos únicos. Os metadados relevantes, como idade e sexo dos pacientes, foram organizados em planilha auxiliar, permitindo a indexação automática dos dados.

Após triagem inicial, 69 imagens foram descartadas por não atenderem aos critérios estabelecidos (edentulismo total ou menos de seis dentes visíveis), resultando em uma base final composta por 799 radiografias válidas. Os arquivos foram renomeados seguindo a convenção Número_Idade_Sexo, o que facilitou tanto a rastreabilidade quanto a automação das etapas subsequentes.

4.1.1 Critérios de Inclusão e Exclusão

Os critérios adotados visaram assegurar a qualidade visual e clínica das imagens incluídas na base de dados:

- Presença de, no mínimo, seis dentes visíveis;
- Nitidez satisfatória da imagem, sem borrões ou artefatos de movimento;

- Enquadramento adequado da arcada dentária.

Imagens com edentulismo total, baixa resolução ou distorções anatômicas relevantes foram excluídas. Além disso, nomes ou informações embutidas diretamente nas radiografias foram removidos por meio de cortes localizados.

4.1.2 Limitações do Dataset

Embora adequado aos propósitos deste estudo, o *dataset* apresenta limitações. Por ser oriundo de uma única instituição, há risco de viés populacional. A distribuição entre faixas etárias e sexos é naturalmente desbalanceada, o que pode afetar o desempenho dos modelos em classes sub-representadas. Também não há informações sobre etnia, o que inviabiliza análises relacionadas a variações morfológicas associadas a esse fator.

Ainda assim, o banco apresenta volume razoável e homogeneidade visual suficiente para viabilizar a experimentação com arquiteturas convolucionais, sendo apropriado para estudos exploratórios na área de classificação etária forense.

4.2 Pré-processamento

As radiografias foram redimensionadas para 128×128 pixels com três canais (RGB), mesmo tratando-se de imagens essencialmente em tons de cinza. A codificação foi mantida em RGB para garantir compatibilidade com os filtros de realce e com os modelos pré-treinados utilizados.

Em seguida, aplicou-se uma máscara de região de interesse (ROI) para ocultar áreas irrelevantes e destacar a arcada dentária. A etapa buscou reduzir interferências visuais e direcionar o foco das redes às estruturas mais informativas.

Cada imagem foi submetida a oito variantes de pré-processamento, incluindo a versão original sem aplicação de filtros, além dos realces por Laplaciano, Sobel, Prewitt, Roberts, *Top-hat*, *Black-hat* e a correção gama com $\gamma = 3$.

Além de sua aplicação rotineira em tarefas gerais de segmentação e detecção de bordas, os filtros utilizados neste estudo foram selecionados por sua capacidade de evidenciar características anatômicas relevantes nas radiografias panorâmicas. Os operadores de borda, como Sobel, Prewitt, Roberts e Laplaciano, foram empregados por sua eficiência em realçar transições de intensidade entre tecidos, o que pode favorecer a identificação de contornos

dentários e limites ósseos.

Já os filtros morfológicos *Top-hat* e *Black-hat* foram escolhidos por sua habilidade em destacar estruturas mais claras ou mais escuras em relação ao fundo, respectivamente, o que pode auxiliar na diferenciação de regiões como raízes, espaços periodontais e trabeculado ósseo. A correção gama ($\gamma = 3$), foi aplicada com o intuito de melhorar a visibilidade de regiões hipodensas, como a dentina e o canal radicular, sem saturar áreas mais claras da imagem. A imagem original (sem aplicação de filtros) também foi incluída como controle, permitindo avaliar se o processamento adicional de realce de contraste traria ganhos efetivos no desempenho dos modelos.

Para o aumento de dados, foi adotado um conjunto de transformações geométricas com o objetivo de simular variações naturais que podem ocorrer durante a aquisição das imagens, sem alterar a informação semântica essencial para a classificação etária. As rotações de até 1 grau permitem que o modelo seja mais robusto a pequenas mudanças de angulação durante o exame, algo comum no contexto clínico, onde a posição do paciente pode variar. Deslocamentos horizontais e verticais de até 8% e zoom de até 8% são úteis para simular variações na posição e proximidade do paciente ao sistema de captura, o que pode ocorrer em imagens radiográficas. O cisalhamento de até 30% serve para introduzir distorções horizontais e verticais, ajudando o modelo a generalizar melhor em situações de variações na geometria das imagens. Como etapa complementar, o espelhamento horizontal e vertical foi utilizado para simular possíveis reflexos ou inversões de imagem, contribuindo para aumentar a diversidade do conjunto de dados e melhorar a generalização do modelo, tornando-o mais robusto a distorções que não afetam as características anatômicas essenciais.

Essas transformações foram aplicadas apenas ao conjunto de treino, com a validação e o teste permanecendo inalterados, de modo a garantir que a avaliação do modelo refletisse seu desempenho em dados não modificados. Com a aplicação dessas técnicas, o conjunto de treino foi expandido em 48.5%, o que significa que o número de imagens foi multiplicado por 486, proporcionando uma maior variedade de dados e, assim, melhorando a robustez do modelo frente a diferentes variações nas imagens.

4.3 Ambiente Experimental

Os experimentos foram conduzidos na plataforma Google Colab Pro, com suporte de GPU NVIDIA Tesla T4 e memória RAM expandida. A implementação foi realizada em Python

3 com o framework TensorFlow/Keras 2.18.0.

As bibliotecas auxiliares incluíram:

- **OpenCV**: aplicação dos filtros e realces;
- **NumPy e Pandas**: manipulação de dados e metadados;
- **Scikit-learn**: validação cruzada e cálculo de métricas;
- **Matplotlib e Seaborn**: geração de gráficos.

As redes InceptionV3, EfficientNetV2B0 e VGG19 foram carregadas com pesos pré-treinados do ImageNet, e adaptadas para classificação em três classes.

4.4 Arquiteturas de Redes Neurais Convolucionais

As constantes inovações no design de redes convolucionais têm levado ao desenvolvimento de arquiteturas cada vez mais eficientes, profundas e adaptadas a diferentes demandas computacionais. No contexto da classificação etária forense a partir de radiografias odontológicas, a escolha da arquitetura da rede desempenha uma função estabelecida no desempenho do modelo, influenciando diretamente sua capacidade de extrair padrões morfológicos sutis e relevantes. Dentre as diversas opções disponíveis, este trabalho selecionou três arquiteturas de destaque e com abordagens distintas: VGG19, InceptionV3 e EfficientNetV2B0.

A VGG19 representa uma das primeiras tentativas bem-sucedidas de aprofundar redes convolucionais utilizando um padrão estrutural consistente e simples. Seu diferencial está na aplicação sistemática de filtros convolucionais 3×3 , empilhados em sequência ao longo de 19 camadas com pesos, sendo 16 convolucionais e 3 totalmente conectadas. Essa estratégia possibilita aumentar a profundidade da rede sem inflar excessivamente o número de parâmetros, além de incorporar múltiplas ativações ReLU que contribuem para a não linearidade e capacidade expressiva do modelo. Apesar de sua estrutura direta, a VGG19 apresentou desempenho competitivo em benchmarks como o *ImageNet*, consolidando-se como um marco importante na evolução das CNNs. No entanto, seu custo computacional elevado, tanto em memória quanto em tempo de treinamento, pode limitar seu uso em sistemas com recursos restritos (SIMONYAN; ZISSERMAN, 2014).

Em contraste ao desenho sequencial da VGG19, a InceptionV3 adota uma arquitetura mais sofisticada baseada em módulos Inception, que operam múltiplas convoluções de diferentes tamanhos de kernel em paralelo. Essa característica permite à rede capturar informações em diversas escalas, favorecendo uma extração de características mais rica sem comprometer a

eficiência computacional. Além disso, técnicas como a fatoração de convoluções maiores em sequências de convoluções menores (por exemplo, substituir uma operação 5×5 por duas 3×3), o uso de convoluções 1×1 para redução de dimensionalidade, e práticas otimizadas de normalização e regularização tornam essa arquitetura robusta e eficiente. Tais características posicionam a InceptionV3 como uma opção versátil e precisa para tarefas de classificação em imagens complexas (SZEGEDY *et al.*, 2016).

O conjunto de arquiteturas EfficientNet, traz uma abordagem distinta de escalonamento de redes neurais convolucionais, introduzindo o conceito de *compound scaling*. Em vez de expandir aleatoriamente a profundidade, largura ou resolução da rede, esse método ajusta os três fatores de forma equilibrada, com base em princípios empíricos. A variante EfficientNetV2B0, utilizada neste trabalho, é uma das versões mais leves da série e foi projetada com ênfase em velocidade de treinamento e economia de recursos, mantendo alta acurácia. Dentre suas inovações destacam-se os blocos Fused-MBConv, que integram convoluções, ativações e normalizações eficientemente, além do uso de técnicas de regularização modernas que aceleram a convergência. Por essas características, a EfficientNetV2B0 mostra-se particularmente indicada para aplicações em ambientes com restrições computacionais, sem prejuízo de desempenho (TAN; LE, 2021)

4.5 Modelagem e Treinamento

Cada arquitetura foi adaptada à tarefa de classificação multiclasse por meio da substituição de sua camada final original por uma camada totalmente conectada (Dense), composta por três neurônios e função de ativação softmax. Essa configuração permite que o modelo associe a imagem de entrada a uma das três faixas etárias previamente definidas, produzindo uma distribuição de probabilidade sobre as classes.

A divisão inicial do conjunto de dados foi estratificada, reservando-se 80% das imagens para as etapas de treinamento e validação, e os 20% restantes para o teste final. A estratificação visou preservar a proporção entre as classes etárias em todas as partições, evitando viés na distribuição dos exemplos. Na fase de treinamento, foi aplicada validação cruzada com quatro subdivisões (*Stratified K-Fold*, $k = 4$), assegurando que cada partição refletisse adequadamente a diversidade da base e permitindo a obtenção de estimativas mais robustas de desempenho.

Com o objetivo de mitigar o sobreajuste, foram incorporadas técnicas de regularização específicas para cada arquitetura. As redes InceptionV3 e EfficientNetV2B0 receberam

camadas Dropout com taxa de 0,2, enquanto a VGG19 utilizou uma taxa mais elevada (0,5), dado seu maior número de parâmetros treináveis e maior suscetibilidade ao overfitting. Complementarmente, foram empregados três mecanismos de controle durante o treinamento: o EarlyStopping, com paciência de 50 épocas, que interrompe a execução caso não haja melhora no desempenho de validação por um número consecutivo de ciclos; o ReduceLROnPlateau, com fator de redução de 0,05, que ajusta dinamicamente a taxa de aprendizado ao detectar estagnação; e o ModelCheckpoint, responsável por armazenar apenas os modelos com menor perda em validação.

O treinamento foi conduzido com o otimizador *Adam*, amplamente adotado em aplicações de aprendizado profundo por sua capacidade de adaptar individualmente os coeficientes de atualização para cada peso da rede. A taxa de aprendizado inicial foi fixada em 10^{-4} , valor que favorece uma convergência mais estável. A função de perda selecionada foi a *categorical_crossentropy*, apropriada para tarefas de classificação multiclasse com saídas mutuamente exclusivas.

4.6 Métricas de Avaliação

As métricas utilizadas para aferir o desempenho dos modelos foram Acurácia, *Precision* macro, *Recall* macro e *F1-Score* macro.

A acurácia corresponde à proporção de acertos globais sobre o total de amostras, sendo útil para uma visão geral do desempenho do modelo. Sua formulação é dada por:

$$\text{Acurácia} = \frac{\sum_{i=1}^n \mathbf{1}(y_i = \hat{y}_i)}{n} \quad (4.1)$$

em que y_i representa o rótulo verdadeiro da i -ésima amostra, \hat{y}_i o rótulo previsto, $\mathbf{1}(\cdot)$ é a função indicadora, e n o número total de amostras.

No entanto, a acurácia pode ser limitada em contextos de desequilíbrio entre classes, como ocorre na distribuição etária deste estudo. Por essa razão, foram incorporadas métricas complementares mais sensíveis ao desempenho por classe, calculadas segundo a média macro — isto é, com igual peso para cada classe, independentemente do número de amostras.

A *Precision* macro, que avalia a proporção de previsões corretas dentre aquelas atribuídas a cada classe, é definida por:

$$Precision_{macro} = \frac{1}{C} \sum_{c=1}^C \frac{TP_c}{TP_c + FP_c} \quad (4.2)$$

O *Recall* macro, por sua vez, mede a capacidade do modelo de recuperar corretamente os exemplos de cada classe no total de ocorrências. Sua equação é:

$$Recall_{macro} = \frac{1}{C} \sum_{c=1}^C \frac{TP_c}{TP_c + FN_c} \quad (4.3)$$

O *F1-Score* macro combina as duas métricas anteriores por meio da média harmônica, buscando equilíbrio entre *Precision* e *Recall* em cada classe. Sua formulação é expressa por:

$$F1\text{-}Score_{macro} = \frac{1}{C} \sum_{c=1}^C \frac{2 \cdot Precision_c \cdot Recall_c}{Precision_c + Recall_c} \quad (4.4)$$

onde C representa o número total de classes, TP_c os verdadeiros positivos, FP_c os falsos positivos, e FN_c os falsos negativos da classe c .

Essa abordagem se justifica pela necessidade de avaliar o comportamento do modelo de forma equitativa, evitando que o desempenho sobre classes majoritárias domine a média geral e oculte falhas nas demais.

Todas as métricas foram calculadas com base no conjunto de teste, considerando o melhor modelo obtido em cada *fold* da validação cruzada.

4.7 Extração e Visualização dos Resultados

Durante o treinamento e avaliação dos modelos, as métricas de desempenho foram calculadas para cada *fold* da validação cruzada e salvas em arquivos (.csv). Esses arquivos foram organizados com o auxílio da biblioteca Pandas, permitindo a estruturação dos dados em tabelas e sua manipulação para posterior análise.

Para facilitar a comparação entre filtros de pré-processamento e arquiteturas convolucionais, foram gerados gráficos de barras representando a acurácia média por combinação, acompanhados de barras de erro correspondentes ao desvio-padrão. A escolha desse tipo de visualização teve por objetivo não apenas apresentar os valores médios, mas também evidenciar a variabilidade entre os *folds*. Ao incluir o desvio-padrão, é possível avaliar a consistência dos modelos ao longo das repetições, o que fornece uma dimensão adicional sobre a confiabilidade dos resultados, aspecto relevante em contextos com bases de dados limitadas ou desbalanceadas.

Além dos gráficos comparativos, curvas de aprendizado foram geradas para cada experimento, com foco na evolução da acurácia e da perda ao longo das épocas de treinamento. As curvas foram produzidas com `Matplotlib`, utilizando o `Seaborn` para maior legibilidade. A inclusão de faixas de desvio-padrão em torno das linhas principais buscou fornecer uma visualização mais robusta da estabilidade do modelo durante o treinamento. Ao observar essas variações, é possível identificar padrões de oscilação, sobreajuste precoce ou instabilidades específicas que não seriam perceptíveis apenas pelos valores finais de acurácia.

Todas as figuras foram salvas em alta resolução (.png) e organizadas de forma sistemática, garantindo uma leitura clara dos efeitos das diferentes abordagens testadas.

5 RESULTADOS E DISCUSSÕES

Os resultados experimentais não só permitem identificar as melhores combinações de arquitetura e filtro de pré-processamento, como também compreender melhor o comportamento de cada rede diante de diferentes abordagens de realce e transformação das imagens.

5.1 Apresentação dos Resultados

Nesta seção, são apresentados os resultados quantitativos obtidos a partir da validação cruzada estratificada em 4 *folds*, contemplando as métricas de desempenho (Acurácia, *Precision*, *Recall* e *F1-Score*) para diferentes combinações de arquitetura e técnica de pré-processamento. Além desses resultados, são exibidas curvas de aprendizado (acurácia e função de perda), para avaliação da dinâmica do treinamento e eventuais indícios de sobreajuste ou subajuste. Por fim, apresenta-se um gráfico comparativo que sintetiza o desempenho médio dos modelos em relação aos diferentes filtros aplicados.

5.1.1 InceptionV3

Os resultados obtidos com essa arquitetura são apresentados na Tabela 1. O melhor desempenho foi com o pré-processamento do filtro *Top-hat* (modo add), resultando em um *F1-Score* de **$83,81\% \pm 1,91$** , seguido pela utilização da imagem original e pela correção gama, ambas com *F1-Scores* muito próximos — $82,18\% \pm 1,30$ e $81,99\% \pm 2,62$, respectivamente. Esses resultados sugerem que a aplicação de técnicas de realce, especialmente aquelas voltadas à amplificação de estruturas claras sobre fundo escuro, favorece que a extração de características realizadas pela rede seja mais discriminativas. Embora haja variações entre os filtros, o InceptionV3 manteve desempenho estável, podendo estar associado à sua capacidade de captar padrões em múltiplas escalas por meio dos módulos Inception. Logo, a arquitetura é robusta o suficiente para lidar com diferentes formas de pré-processamento.

Tabela 1 – Desempenho médio do modelo InceptionV3 com diferentes técnicas de pré-processamento.

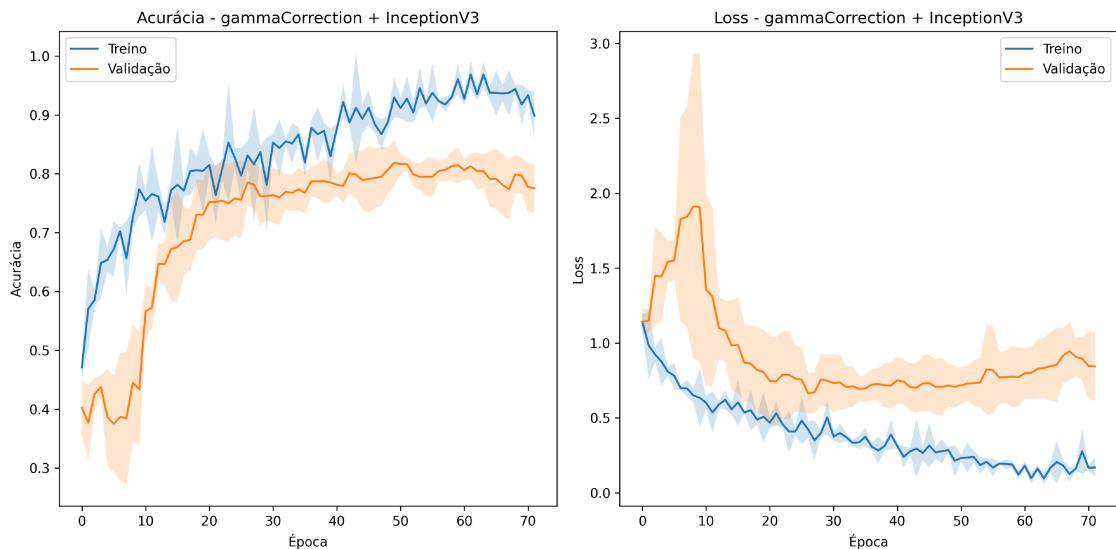
Filtro	Acurácia	Precision	Recall	FI-Score
Imagen original	81.97 ± 1.30	83.42 ± 1.94	81.97 ± 1.30	82.18 ± 1.30
Laplaciano	80.27 ± 4.87	82.16 ± 3.00	80.27 ± 4.87	79.74 ± 5.58
Sobel	79.76 ± 0.34	82.31 ± 0.91	79.76 ± 0.34	80.05 ± 0.32
Prewitt	79.93 ± 5.57	82.00 ± 4.70	79.93 ± 5.57	79.88 ± 5.80
Roberts	80.44 ± 3.16	81.96 ± 3.83	80.44 ± 3.16	80.60 ± 3.27
Black-hat	80.27 ± 5.80	84.04 ± 2.35	82.27 ± 5.80	80.35 ± 5.96
Gamma	81.97 ± 2.63	82.94 ± 2.51	81.97 ± 2.63	81.99 ± 2.62
Top-hat (add)	83.67 ± 2.04	85.08 ± 1.64	83.67 ± 2.04	83.81 ± 1.91

Fonte: Autoral

As curvas de aprendizado para o modelo InceptionV3 com os filtros Gamma Correction e Top-hat (modo add) são apresentadas nas Figuras 11 e 12, respectivamente.

A acurácia de validação do modelo com o filtro Gamma Correction alcança patamares entre 78% e 80%, mantendo-se estável nas últimas épocas — desempenho superior ao de diversos outros filtros, como *Black-hat*, Prewitt e Laplaciano. A curva de perda de validação apresenta comportamento decrescente e consistente, sem aumento nas iterações finais, o que indica boa generalização e baixo risco de sobreajuste. Além disso, a coerência entre a acurácia crescente e o *loss* decrescente, aliada à baixa variabilidade entre os *folds*, evidencia a estabilidade do modelo com esse tipo de pré-processamento.

Figura 11 – Curvas de aprendizado do modelo InceptionV3 com o filtro Gamma Correction.



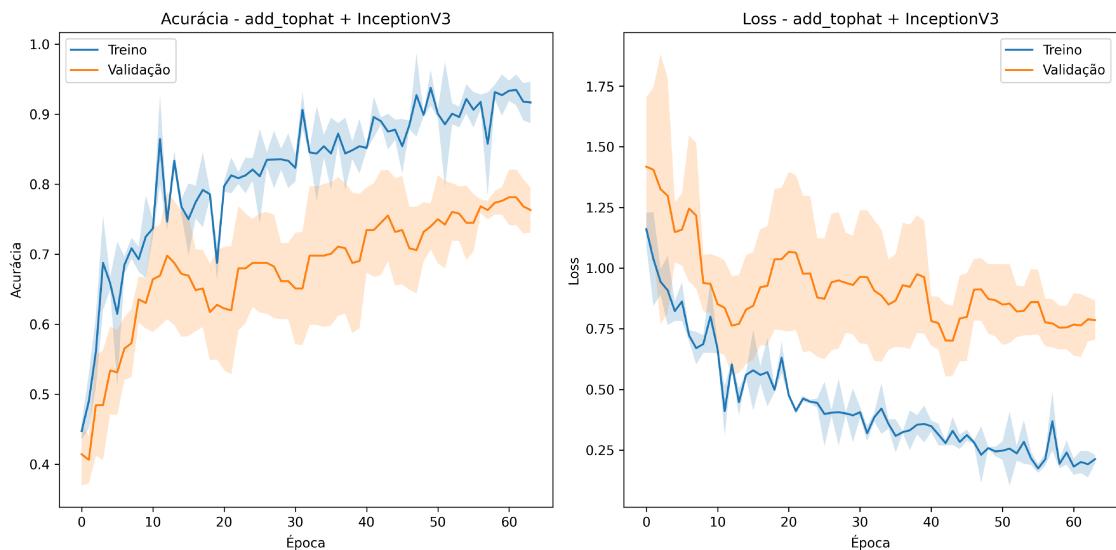
Fonte: Autoral

Em contraste, o comportamento do modelo com o filtro Top-hat (modo add), apresen-

tou o maior *F1-Score* médio com o InceptionV3 ($83,81\% \pm 1,91$). Entretanto, as curvas revelam maior oscilação entre os *folds* e maior distância entre as curvas de treino e validação — indícios de um aprendizado mais instável. A acurácia de validação apresenta flutuações, especialmente após a época 30, e a curva de perda mantém-se acima de 0,8 nas últimas iterações. Esses aspectos sugerem que o bom desempenho do filtro pode estar concentrado em determinados *folds*, reforçando a importância de considerar tanto as métricas finais quanto o comportamento do treinamento na avaliação dos modelos.

Nos gráficos de treinamento, com as curvas de acurácia (à esquerda) e função de perda (à direita), as faixas sombreadas indicam o desvio-padrão entre os *folds* da validação cruzada.

Figura 12 – Curvas de aprendizado do modelo InceptionV3 com o filtro Top-hat (modo add).



Fonte: Autoral

5.1.2 EfficientNET

A Tabela 2 apresenta os resultados obtidos pela arquitetura EfficientNetV2B0, que demonstrou o melhor desempenho global entre os modelos avaliados. O destaque foi o modelo com filtro Gamma Correction, que alcançou o maior *F1-Score* ($84,69\% \pm 0,02$), acompanhado de acurácia de $84,52\% \pm 0,02$ e Recall igualmente elevado, indicando alta sensibilidade na classificação correta das amostras. A imagem original, mesmo sem pré-processamento adicional, também proporcionou resultados robustos (*F1-Score* de $82,11\% \pm 0,32$), enquanto o filtro Black-hat apresentou desempenho ligeiramente superior ($82,32\% \pm 0,02$). Por outro lado, o filtro

Top-hat, que destacou-se em outras arquiteturas, obteve o pior desempenho nesta configuração. A consistência dos resultados e os baixos desvios-padrão observados reforçam a estabilidade da EfficientNetV2B0, cuja estrutura balanceada entre profundidade, largura e resolução favorece a aprendizagem eficiente, mesmo em cenários com diferentes transformações de entrada.

Tabela 2 – Desempenho médio do modelo EfficientNetV2B0 com diferentes técnicas de pré-processamento.

Filtro	Acurácia	Precision	Recall	F1-Score
Imagen original	81.97 ± 0.32	83.33 ± 0.42	81.97 ± 0.32	82.11 ± 0.32
Laplaciano	80.27 ± 0.02	80.65 ± 0.02	80.27 ± 0.02	79.74 ± 0.02
Sobel	80.27 ± 0.01	81.53 ± 0.01	80.27 ± 0.01	80.05 ± 0.01
Prewitt	81.29 ± 0.02	82.00 ± 0.02	81.29 ± 0.02	81.92 ± 0.02
Roberts	81.97 ± 0.02	83.50 ± 0.02	81.97 ± 0.02	82.06 ± 0.02
Black-hat	82.14 ± 0.02	83.63 ± 0.02	82.14 ± 0.02	82.32 ± 0.02
Gamma	84.52 ± 0.02	85.47 ± 0.02	84.52 ± 0.02	84.69 ± 0.02
Top-hat (add)	79.76 ± 0.04	80.43 ± 0.04	79.76 ± 0.04	79.70 ± 0.04

Fonte: Autoral

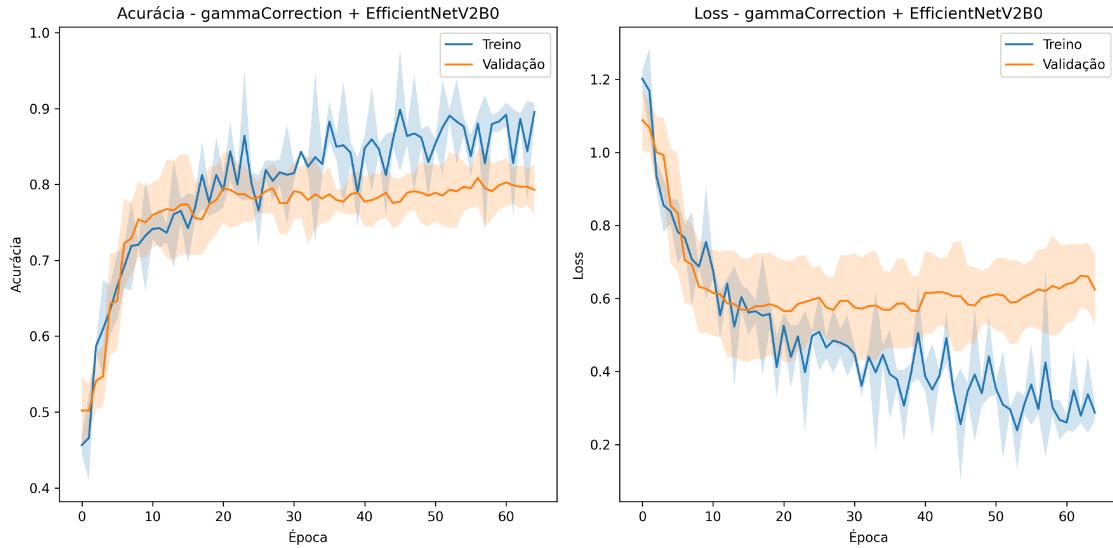
As curvas de aprendizado para o modelo EfficientNetV2B0 com os filtros Gamma Correction e Subtract Prewitt são apresentadas nas Figuras 13 e 14, respectivamente.

Na Figura 13, observa-se que a acurácia de validação atinge patamares superiores a 80% e apresenta crescimento estável até cerca da época 50, mantendo-se constante até o final do treinamento. A curva de perda de validação apresenta padrão decrescente, com oscilações discretas, indicando boa generalização. A distância entre as curvas de treino e validação é moderada, mas persistente, sugerindo leve sobreajuste. Ainda assim, o desempenho global obtido com esse filtro foi o melhor entre todas as combinações avaliadas, justificando sua posição de destaque na Tabela 2.

Já a Figura 14 apresenta o comportamento com o filtro Subtract Prewitt, que, embora não tenha atingido os maiores valores absolutos de desempenho, apresentou curvas com menor dispersão entre os *folds* e comportamento de perda mais estável. A acurácia de validação estabiliza-se acima de 78%, e a curva de perda mantém trajetória suavemente decrescente, sem flutuações expressivas. O pequeno distanciamento entre as curvas de treino e validação reforça a regularidade do modelo com este filtro, sugerindo um treinamento mais controlado e sem sinais evidentes de sobreajuste.

As curvas representam a acurácia (à esquerda) e a função de perda (à direita) em treino e validação ao longo das épocas. As faixas sombreadas indicam o desvio-padrão entre os

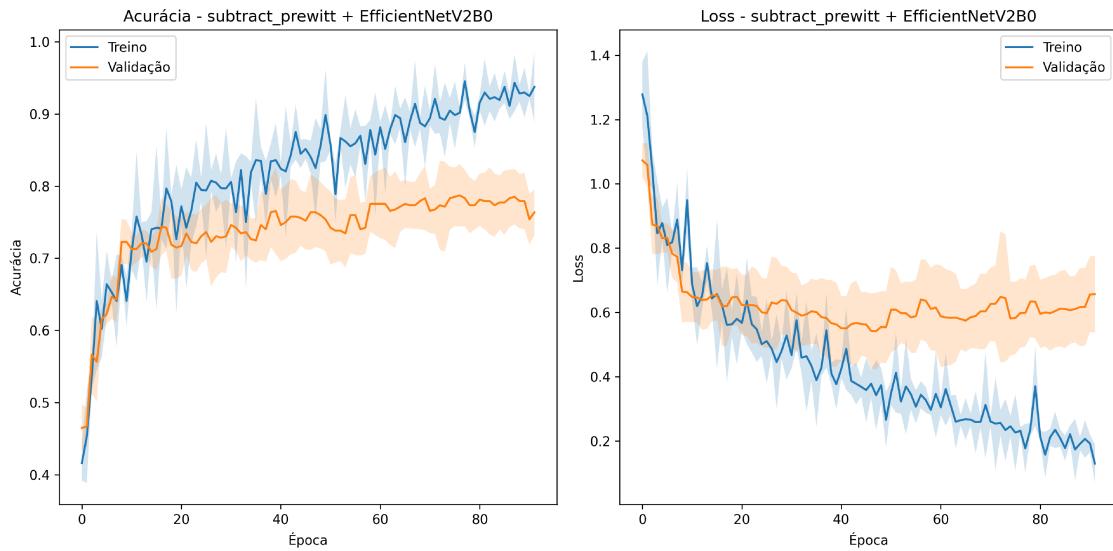
Figura 13 – Curvas de aprendizado do modelo EfficientNetV2B0 com o filtro Gamma Correction.



Fonte: Autoral

folds da validação cruzada.

Figura 14 – Curvas de aprendizado do modelo EfficientNetV2B0 com o filtro Subtract Prewitt.



Fonte: Autoral

5.1.3 VGG19

A Tabela 3 apresenta o desempenho da arquitetura VGG19 com diferentes filtros de pré-processamento. De modo geral, esse modelo obteve resultados sólidos, com destaque para a imagem original, que apresentou o maior *F1-Score* médio ($82,70\% \pm 1,37$), além de

Acurácia e *Precision* superiores às obtidas com os demais filtros. Em seguida, o filtro Prewitt apresentou desempenho bastante próximo (*F1-Score* de $82,61\% \pm 2,45$), seguido por Black-hat ($81,82\% \pm 2,59$) e Roberts ($81,76\% \pm 1,48$). Por outro lado, o filtro Gamma apresentou o pior desempenho, com *F1-Score* de $74,21\% \pm 18,52$ e alta variabilidade entre os *folds* — um indicativo de instabilidade. Esses resultados sugerem que o VGG19 é mais sensível às variações introduzidas pelas transformações na imagem, possivelmente devido à sua arquitetura sequencial profunda, que tende a amplificar padrões indesejados quando exposta a alterações visuais excessivas.

Tal sensibilidade pode ser atribuída ao fato de a arquitetura da VGG19 ser composta por blocos convolucionais lineares e empilhados, sem mecanismos internos de adaptação multiescala, como os encontrados em arquiteturas mais recentes. Nesse contexto, pequenas perturbações nas entradas — como bordas intensificadas ou regiões morfológicamente alteradas — podem se propagar de forma cumulativa ao longo da rede, afetando negativamente a extração de características e o desempenho geral.

Tabela 3 – Desempenho médio do modelo VGG19 com diferentes técnicas de pré-processamento.

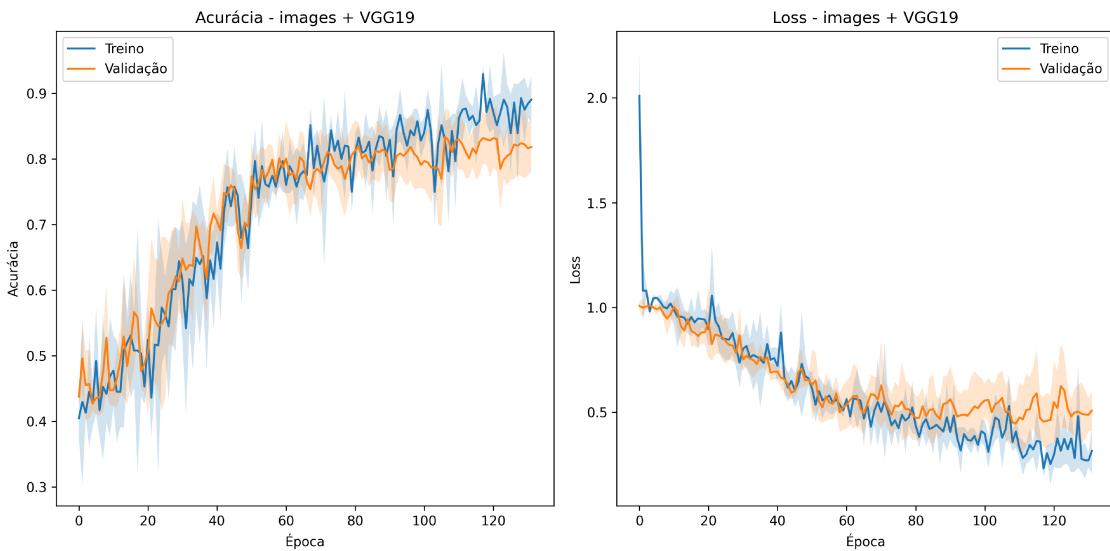
Filtro	Acurácia	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Imagen original	82.48 ± 1.51	85.13 ± 2.05	82.48 ± 1.51	82.70 ± 1.37
Laplaciano	80.78 ± 2.25	82.68 ± 3.64	80.78 ± 2.25	80.65 ± 2.47
Sobel	79.59 ± 2.42	81.00 ± 1.41	79.59 ± 2.42	79.67 ± 2.32
Prewitt	82.48 ± 2.44	83.88 ± 2.10	82.48 ± 2.44	82.61 ± 2.45
Roberts	81.29 ± 1.30	81.90 ± 0.81	81.29 ± 1.30	81.76 ± 1.48
Black-hat	81.29 ± 2.51	83.33 ± 2.35	81.29 ± 2.51	81.82 ± 2.59
Gamma	76.02 ± 15.12	74.59 ± 19.17	76.02 ± 15.12	74.21 ± 18.52
Top-hat	81.80 ± 3.78	83.34 ± 3.72	81.80 ± 3.78	81.44 ± 4.39

Fonte: Autoral

A Figura 15 apresenta as curvas de aprendizado do modelo VGG19 com a imagem original (sem aplicação de filtros). Observa-se uma evolução contínua e estável da acurácia em treino e validação ao longo de mais de 120 épocas, com a acurácia de validação ultrapassando 85% nas iterações finais. A proximidade entre as curvas de treino e validação indica capacidade de generalização e ausência de sobreajuste. A curva de perda (*loss*) em validação apresenta declínio suave e consistente, sem oscilações abruptas — um sinal adicional de estabilidade. Além disso, a baixa dispersão entre os *folds*, representada pelas faixas sombreadas estreitas, reforça a consistência dos resultados. Em contraste, filtros como Gamma Correction, Laplacian e Roberts exibiram maior variabilidade e maior distância entre as curvas, evidenciando desempenho inferior.

Esses aspectos tornam a imagem original, sem pré-processamento mais eficaz para a arquitetura VGG19.

Figura 15 – Curvas de aprendizado do modelo VGG19 com imagem original (sem filtros de pré-processamento).



Fonte: Autoral

A coerência entre os resultados quantitativos e a análise visual reforça que, preservar as características originais da imagem tende a ser mais vantajoso do que aplicar filtros morfológicos para a arquitetura VGG19, pois esse pré-processamento compromete a estabilidade e a capacidade de generalização do modelo.

5.2 Análise Comparativa

Esta seção realiza uma análise comparativa entre as três arquiteturas avaliadas — InceptionV3, EfficientNetV2B0 e VGG19 — com base no desempenho médio obtido para cada técnica de pré-processamento. O objetivo é identificar quais modelos demonstraram maior robustez frente às transformações aplicadas às imagens e quais filtros contribuíram de forma mais expressiva para o desempenho de cada arquitetura.

A Tabela 4 apresenta um resumo dos melhores filtros para cada modelo em cada uma das métricas avaliadas, destacando em negrito os maiores valores absolutos entre todos os modelos. Essa visualização permite comparar não apenas os efeitos dos diferentes filtros em cada arquitetura, mas também quais modelos se mostraram mais eficazes para a tarefa de classificação etária forense com base nas radiografias panorâmicas avaliadas.

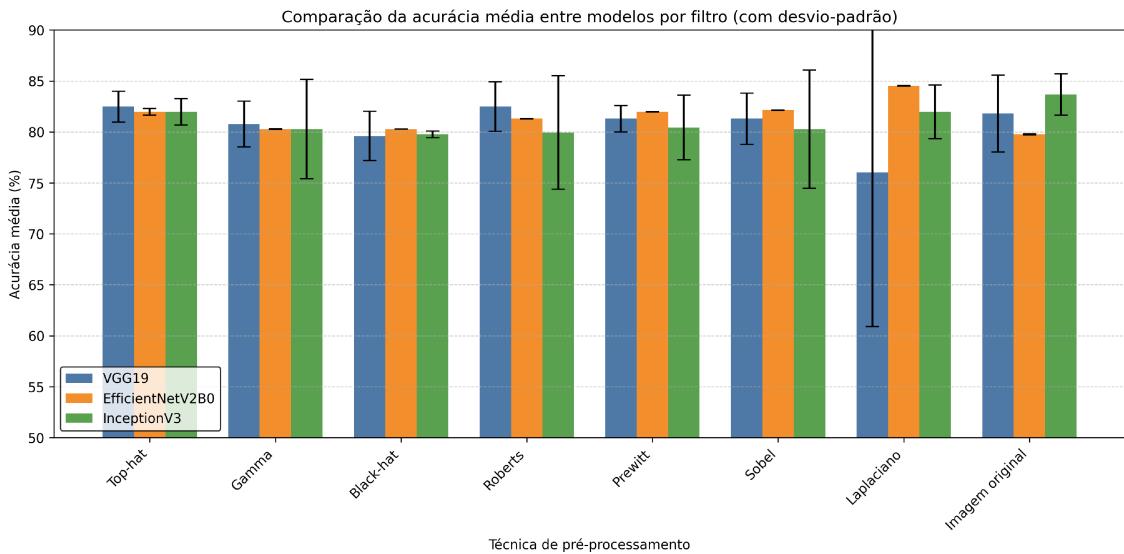
Tabela 4 – Filtros de pré-processamento que obtiveram o melhor desempenho em cada métrica por modelo. Os destaque em negrito representam os melhores resultados globais.

Modelo	Acurácia	Precision	Recall	F1-score
InceptionV3	Top-hat	Top-hat	Top-hat	Top-hat
EfficientNetV2B0	Gamma	Gamma	Gamma	Gamma
VGG19	Sem filtro	Sem filtro	Sem filtro	Sem filtro

A Tabela 4 sintetiza os melhores filtros de pré-processamento para cada modelo em cada métrica avaliada. Como indicado na tabela, o modelo VGG19 obteve os melhores resultados utilizando a imagem original, sem aplicação de filtros, enquanto o InceptionV3 apresentou o melhor desempenho com o filtro Top-hat. Já a EfficientNetV2B0 se destacou com o filtro Gamma, obtendo os melhores resultados em todas as métricas avaliadas, incluindo Acurácia, *Precision*, *Recall* e *F1-Score*. Esses resultados indicam que a EfficientNetV2B0 foi o modelo mais robusto e consistente frente às transformações aplicadas às imagens.

Além da síntese tabular, a Figura 16 ilustra a comparação visual das acurácias médias obtidas por cada modelo, em função das técnicas de pré-processamento avaliadas, e inclui as respectivas barras de erro (desvios-padrão). O gráfico facilita a observação de tendências de desempenho e proporciona uma dimensão adicional sobre a estabilidade dos modelos. A inclusão das barras de erro permite avaliar a consistência dos resultados ao longo dos *folds*, o que é particularmente relevante em contextos com amostras limitadas ou classes desbalanceadas — como é o caso da classificação etária forense.

Figura 16 – acurácia média dos modelos em função das técnicas de pré-processamento. As barras de erro representam o desvio-padrão obtido na validação cruzada (4-fold).



Fonte: Autoral

O filtro Gamma Correction se destacou na arquitetura EfficientNetV2B0, com os maiores valores de Acurácia, *Precision*, *Recall* e *F1-Score* entre todas as combinações testadas. Isso reforça a eficácia da técnica na amplificação de contrastes locais, favorecendo redes mais recentes e otimizadas, que parecem beneficiar-se de ajustes sutis na intensidade da imagem para extrair representações mais discriminativas.

Já no modelo InceptionV3, o melhor desempenho foi o Top-hat (modo add), que realça regiões claras sobre fundos escuros. Essa técnica pode ter favorecido a detecção de contornos dentários e estruturas de interesse ao interagir positivamente com a arquitetura modular da rede, que combina convoluções em diferentes escalas (1×1 , 3×3 , 5×5) e operações de *pooling* nos módulos Inception. Tal estrutura permite a extração simultânea de padrões locais e globais, tornando o modelo mais adaptável às variações introduzidas por esse tipo de realce.

Já o VGG19 apresentou seus melhores resultados com a imagem original, sem aplicação de filtros. Isso sugere que a arquitetura, composta por blocos convolucionais sequenciais e de profundidade fixa, é mais sensível a perturbações nas imagens e se beneficia da preservação dos padrões visuais originais, desde que os dados estejam bem normalizados e padronizados.

Também se observa que a EfficientNetV2B0 apresentou os menores desvios-padrão ao longo dos *folds* de validação, o que indica maior estabilidade e confiabilidade nos resultados. Essa característica é especialmente desejável em aplicações forenses, onde a consistência do modelo frente à variabilidade das imagens é um critério importante para sua adoção prática.

Além da análise do desempenho das arquiteturas, é relevante destacar o poder de processamento e o tempo de resposta de cada modelo. No que diz respeito ao tempo de resposta, o modelo EfficientNetV2B0 se mostrou o mais eficiente, com o menor tempo de processamento total, equivalente a 89,01 minutos. Isso pode ser atribuído à sua arquitetura otimizada, que é projetada para maximizar a eficiência computacional sem sacrificar a precisão, o que é especialmente valioso em contextos de análise em larga escala, como a classificação forense. Por outro lado, a arquitetura VGG19, apesar de apresentar resultados impressionantes em termos de precisão, demandou o maior poder de processamento, com um tempo total de 162,66 minutos, refletindo a maior complexidade de suas camadas convolucionais e a necessidade de mais recursos computacionais para lidar com seu maior número de parâmetros. Isso implica que, embora o VGG19 seja eficaz para capturar padrões visuais complexos, sua implementação em tempo real pode ser desafiadora, especialmente em cenários com limitações de hardware.

6 CONCLUSÕES E TRABALHOS FUTUROS

Este trabalho avaliou o desempenho de diferentes arquiteturas de Redes Neurais Convolucionais na tarefa de classificação etária forense a partir de radiografias panorâmicas odontológicas. Foram conduzidos experimentos com três arquiteturas amplamente utilizadas na literatura — VGG19, InceptionV3 e EfficientNetV2B0 — combinadas a oito variações de pré-processamento, incluindo filtros clássicos de realce (como Sobel, Laplaciano e Top-hat), correção gama e imagens sem filtro.

A avaliação foi realizada por meio de validação cruzada estratificada (4-fold), com análise de métricas como Acurácia, *Precision*, *Recall* e *F1-Score*. Os resultados indicaram que o desempenho dos modelos está diretamente associado à interação entre arquitetura e técnica de pré-processamento. A combinação entre InceptionV3 e o filtro *Top-hat* apresentou os melhores resultados médios, alcançando altos níveis de *F1-score* e menor variação entre os folds. A EfficientNetV2B0 também se destacou em alguns cenários, especialmente com a aplicação da correção gama. Já a VGG19 teve desempenho mais consistente ao ser treinada com imagens sem qualquer aplicação de filtro, evidenciando maior estabilidade das curvas de aprendizado.

Dentre os resultados, destaca-se o impacto negativo da correção gama sobre a VGG19, cuja performance caiu significativamente, com aumento no desvio-padrão das métricas entre os folds. Esse comportamento sugere que ajustes não lineares de brilho podem prejudicar a capacidade de redes com arquitetura sequencial de extrair padrões sutis associados à estimativa etária.

De forma geral, os experimentos realizados comprovaram que não há uma configuração universalmente superior, e que o desempenho depende do alinhamento entre o tipo de filtro de pré-processamento e a estrutura interna da rede utilizada. A análise conjunta dos resultados permitiu identificar combinações mais eficazes dentro do escopo do estudo, além de reforçar a necessidade de considerar as particularidades do domínio médico-forense no planejamento de modelos baseados em CNNs.

6.1 Trabalhos Futuros

Algumas direções podem ser exploradas em estudos futuros, com o intuito de aprimorar a acurácia, robustez e interpretabilidade dos modelos de classificação etária forense a partir de radiografias panorâmicas. Dentre as possibilidades, destacam-se:

1. **Aplicar técnicas de segmentação de imagem:** Isolar regiões anatômicas específicas, como mandíbula ou áreas dentárias, pode reduzir ruídos e melhorar a qualidade das informações repassadas ao modelo, favorecendo a aprendizagem de padrões mais relevantes.
2. **Testar novas arquiteturas e abordagens:** Incluir modelos mais recentes, como *ResNet*, *DenseNet* ou redes baseadas em Transformers (como o *Vision Transformer*), pode contribuir para ganhos em desempenho, especialmente em conjuntos de dados com maior complexidade.
3. **Expandir e diversificar o conjunto de dados:** A inclusão de imagens oriundas de diferentes instituições e faixas etárias pode melhorar a capacidade de generalização dos modelos e permitir avaliações mais abrangentes.
4. **Utilizar métodos de interpretabilidade:** Ferramentas como *Grad-CAM* ou *Layer-wise Relevance Propagation (LRP)* podem auxiliar na visualização das regiões da imagem que mais influenciam a decisão da rede, o que é especialmente relevante em aplicações de caráter forense.
5. **Refinar o pipeline de pré-processamento:** Avaliar o impacto de outras técnicas de realce, como CLAHE ou normalização adaptativa, bem como ajustar parâmetros dos filtros já utilizados, pode otimizar ainda mais o desempenho dos modelos.
6. **Realizar validação cruzada com maior número de folds e repetições:** Aumentar a quantidade de particionamentos ou introduzir validações repetidas pode reduzir o viés amostral e conferir maior robustez estatística às análises.
7. **Avaliar combinações das técnicas de pré-processamento:** Uma linha de investigação seria combinar diferentes técnicas de pré-processamento, como filtros de borda e realce de contraste, para melhorar a robustez e a generalização do modelo, evitando overfitting.

REFERÊNCIAS

- ALAM, S. S.; RASHID, N.; FAIZA, T. A.; AHMED, S.; HASSAN, R. A.; DUDLEY, J.; FAROOK, T. H. Estimating age and sex from dental panoramic radiographs using neural networks and vision–language models. **Oral**, MDPI, v. 5, n. 1, p. 3, 2025.
- ALQAHTANI, S. J.; HECTOR, M. P.; LIVERSIDGE, H. M. Brief communication: The london atlas of human tooth development and eruption. **American Journal of physical anthropology**, Wiley Online Library, v. 142, n. 3, p. 481–490, 2010.
- ANWAR, S. M.; MAJID, M.; QAYYUM, A.; AWAIS, M.; ALNOWAMI, M.; KHAN, M. K. Medical image analysis using convolutional neural networks: a review. **Journal of medical systems**, Springer, v. 42, p. 1–13, 2018.
- APAYDIN, B.; YASAR, F. Accuracy of the demirjian, willem and cameriere methods of estimating dental age on turkish children. **Nigerian journal of clinical practice**, Medknow, v. 21, n. 3, p. 257–263, 2018.
- AREL, I.; ROSE, D. C.; KARNOWSKI, T. P. Deep machine learning-a new frontier in artificial intelligence research [research frontier]. **IEEE computational intelligence magazine**, IEEE, v. 5, n. 4, p. 13–18, 2010.
- 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, 2023.
- BENGIO, Y.; GOODFELLOW, I.; COURVILLE, A. *et al.* **Deep learning**. [S.l.]: MIT press Cambridge, MA, USA, 2017. v. 1.
- BIRCHLER, F. A.; KILIARIDIS, S.; COMBESCURE, C.; JULKU, J.; PIRTINIEMI, P. M.; VAZQUEZ, L. Dental age assessment on panoramic radiographs: comparison between two generations of young finnish subjects. **Journal of International Medical Research**, SAGE Publications Sage UK: London, England, v. 47, n. 1, p. 311–324, 2019.
- BIZJAK, Ž.; ROBIĆ, T. Dentage: Deep learning for automated age prediction using panoramic dental x-ray images. **Journal of forensic sciences**, Wiley Online Library, v. 69, n. 6, p. 2069–2074, 2024.
- DENG, L.; YU, D. *et al.* Deep learning: methods and applications. **Foundations and trends® in signal processing**, Now Publishers, Inc., v. 7, n. 3–4, p. 197–387, 2014.
- DRUSINI, A. The coronal pulp cavity index: A forensic tool for age determination in human adults. **Cuad Med Forense**, v. 14, n. 53-54, p. 235–249, 2008.
- FACELI, K.; LORENA, A. C.; GAMA, J.; CARVALHO, A. C. d. **Inteligência Artificial: uma abordagem de aprendizado de máquina**. Rio de Janeiro: Editora LTC, 2011.
- FERREIRA, J. M. S. Aproximação da cinemática inversa de um robô manipulador didático através de algoritmos de aprendizado de máquina. 2024.
- 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.

- FUKUSHIMA, K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. **Biological cybernetics**, Springer, v. 36, n. 4, p. 193–202, 1980.
- GALIĆ, I.; HABIJAN, M.; LEVENTIĆ, H.; ROMIĆ, K. Machine learning empowering personalized medicine: A comprehensive review of medical image analysis methods. **Electronics**, MDPI, v. 12, n. 21, p. 4411, 2023.
- GE, Y.; ZHANG, Q.; SUN, Y.; SHEN, Y.; WANG, X. Grayscale medical image segmentation method based on 2d&3d object detection with deep learning. **BMC Medical Imaging**, Springer, v. 22, n. 1, p. 33, 2022.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A.; BENGIO, Y. **Deep learning**. [S.l.]: MIT press Cambridge, 2016. v. 1.
- GU, J.; WANG, Z.; KUEN, J.; MA, L.; SHAHROUDY, A.; SHUAI, B.; LIU, T.; WANG, X.; WANG, G.; CAI, J. *et al.* Recent advances in convolutional neural networks. **Pattern recognition**, Elsevier, v. 77, p. 354–377, 2018.
- HIYARI, M. H.; PASIC, M.; ZUKIC, S. Application of convolutional neural networks for determining gender and age in forensic dentistry. **Cureus**, Cureus, v. 16, n. 11, 2024.
- IGLOVIKOV, V. I.; RAKHLIN, A.; KALININ, A. A.; SHVETS, A. A. Paediatric bone age assessment using deep convolutional neural networks. In: SPRINGER. **Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings** 4. [S.l.], 2018. p. 300–308.
- ISHWARKUMAR, S.; PILLAY, P.; CHETTY, M.; SATYAPAL, K. S. The application of the cameriere's methodologies for dental age estimation in a select kwazulu-natal population of south africa. **Dentistry Journal**, MDPI, v. 10, n. 7, p. 130, 2022.
- JAYARAMAN, J.; ROBERTS, G. J.; WONG, H. M.; KING, N. M. Dental age estimation in southern chinese population using panoramic radiographs: validation of three population specific reference datasets. **BMC medical imaging**, Springer, v. 18, p. 1–8, 2018.
- KIM, M.; BAE, H.-J. Data augmentation techniques for deep learning-based medical image analyses. **Journal of the Korean Society of Radiology**, The Korean Society of Radiology, v. 81, n. 6, p. 1290–1304, 2020.
- 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.
- KOCH, R. M.; MENTZEL, H.-J.; HEINRICH, A. Deep learning for forensic age estimation using orthopantomograms in children, adolescents, and young adults. **European Radiology**, Springer, p. 1–12, 2025.
- KOKOMOTO, K.; KARIYA, R.; MURANAKA, A.; OKAWA, R.; NAKANO, K.; NOZAKI, K. Automatic dental age calculation from panoramic radiographs using deep learning: a two-stage approach with object detection and image classification. **BMC Oral Health**, Springer, v. 24, n. 1, p. 143, 2024.

- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. **Advances in neural information processing systems**, v. 25, 2012.
- KUSHOL, R.; RAIHAN, M. N.; SALEKIN, M. S.; RAHMAN, A. Contrast enhancement of medical x-ray image using morphological operators with optimal structuring element. **arXiv preprint arXiv:1905.08545**, 2019.
- 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.
- 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, 2002.
- LEE, Y.-H.; AUH, Q.-S.; CHUN, Y.-H.; AN, J.-S. Age-related radiomorphometric changes on panoramic radiographs. **Clinical and Experimental Dental Research**, Wiley Online Library, v. 7, n. 4, p. 539–551, 2021.
- LITJENS, G.; KOOI, T.; BEJNORDI, B. E.; SETIO, A. A. A.; CIOMPI, F.; GHAFOORIAN, M.; LAAK, J. A. V. D.; GINNEKEN, B. V.; SÁNCHEZ, C. I. A survey on deep learning in medical image analysis. **Medical image analysis**, Elsevier, v. 42, p. 60–88, 2017.
- MILOŠEVIĆ, D.; VODANOVIĆ, M.; GALIĆ, I.; SUBAŠIĆ, M. Automated estimation of chronological age from panoramic dental x-ray images using deep learning. **Expert systems with applications**, Elsevier, v. 189, p. 116038, 2022.
- NARIN, A.; KAYA, C.; PAMUK, Z. Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. **Pattern Analysis and Applications**, Springer, v. 24, n. 3, p. 1207–1220, 2021.
- NASER, W. A. K.; KADHIM, A. A.; ABBAS, S. H. Subject review: Image edge detection techniques. **Global J Eng Tech Adv**, v. 19, n. 1, p. 001–005, 2024.
- NIELSEN, M. A. **Neural networks and deep learning**. [S.l.]: Determination press San Francisco, CA, USA, 2015. v. 25.
- 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.
- 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.
- PARK, Y.-S.; LEK, S. Artificial neural networks: Multilayer perceptron for ecological modeling. In: **Developments in environmental modelling**. [S.l.]: Elsevier, 2016. v. 28, p. 123–140.
- PÉREZ, I. M. M.; BOURHIM, S.; PÉREZ, S. E. M. Artificial intelligence-based models for automated bone age assessment from posteroanterior wrist x-rays: A systematic review. **Applied Sciences**, MDPI, v. 15, n. 11, p. 5978, 2025.

- PEREZ, L.; WANG, J. The effectiveness of data augmentation in image classification using deep learning. **arXiv preprint arXiv:1712.04621**, 2017.
- PRETTY, I.; SWEET, D. A look at forensic dentistry—part 1: The role of teeth in the determination of human identity. **British dental journal**, Nature Publishing Group, v. 190, n. 7, p. 359–366, 2001.
- PURWONO, P.; MA’ARIF, A.; RAHMANIAR, W.; FATHURRAHMAN, H. I. K.; FRISKY, A. Z. K.; HAQ, Q. M. ul. Understanding of convolutional neural network (cnn): A review. **International Journal of Robotics and Control Systems**, v. 2, n. 4, p. 739–748, 2022.
- RAWAT, W.; WANG, Z. Deep convolutional neural networks for image classification: A comprehensive review. **Neural computation**, MIT Press, v. 29, n. 9, p. 2352–2449, 2017.
- RAYED, M. E.; ISLAM, S. S.; NIHA, S. I.; JIM, J. R.; KABIR, M. M.; MRIDHA, M. Deep learning for medical image segmentation: State-of-the-art advancements and challenges. **Informatics in Medicine Unlocked**, Elsevier, p. 101504, 2024.
- RECHT, B.; ROELOFS, R.; SCHMIDT, L.; SHANKAR, V. Do imagenet classifiers generalize to imagenet? In: PMLR. **International conference on machine learning**. [S.l.], 2019. p. 5389–5400.
- REINHOLD, J. C.; DEWEY, B. E.; CARASS, A.; PRINCE, J. L. Evaluating the impact of intensity normalization on mr image synthesis. In: **Proceedings of SPIE—the International Society for Optical Engineering**. [S.l.: s.n.], 2019. v. 10949, p. 109493H.
- SALEHI, A. W.; KHAN, S.; GUPTA, G.; ALABDULLAH, B. I.; ALMJALLY, A.; ALSOLAI, H.; SIDDIQUI, T.; MELLIT, A. A study of cnn and transfer learning in medical imaging: Advantages, challenges, future scope. **Sustainability**, MDPI, v. 15, n. 7, p. 5930, 2023.
- SCHMELING, A.; DETTMAYER, R.; RUDOLF, E.; VIETH, V.; GESERICK, G. Forensic age estimation: methods, certainty, and the law. **Deutsches Ärzteblatt International**, v. 113, n. 4, p. 44, 2016.
- SCHMIDHUBER, J. Deep learning in neural networks: An overview. **Neural networks**, Elsevier, v. 61, p. 85–117, 2015.
- SHEN, S.; GUO, Y.; HAN, J.; SUI, M.; ZHOU, Z.; TAO, J. Predicting chronological age of 14 or 18 in adolescents: integrating dental assessments with machine learning. **BMC pediatrics**, Springer, v. 24, n. 1, p. 248, 2024.
- SHEN, S.; LIU, Z.; WANG, J.; FAN, L.; JI, F.; TAO, J. Machine learning assisted cameriere method for dental age estimation. **BMC Oral Health**, Springer, v. 21, n. 1, p. 641, 2021.
- SHRESTHA, A.; MAHMOOD, A. Review of deep learning algorithms and architectures. **IEEE access**, IEEE, v. 7, p. 53040–53065, 2019.
- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. **arXiv preprint arXiv:1409.1556**, 2014.
- SOMASUNDARAM, K.; KALAVATHI, P. Medical image contrast enhancement based on gamma correction. **Int J Knowl Manag e-learning**, v. 3, n. 1, p. 15–18, 2011.

- SZEGEDY, C.; VANHOUCKE, 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. Efficientnetv2: Smaller models and faster training. In: PMLR. **International conference on machine learning**. [S.l.], 2021. p. 10096–10106.
- TAYE, M. M. Theoretical understanding of convolutional neural network: Concepts, architectures, applications, future directions. **Computation**, MDPI, v. 11, n. 3, p. 52, 2023.
- TAYLOR, L.; NITSCHKE, G. Improving deep learning with generic data augmentation. In: IEEE. **2018 IEEE symposium series on computational intelligence (SSCI)**. [S.l.], 2018. p. 1542–1547.
- TIMME, M.; TIMME, W.; OLZE, A.; OTTOW, C.; RIBBECKE, S.; PFEIFFER, H.; DETTMAYER, R.; SCHMELING, A. The chronology of the radiographic visibility of the periodontal ligament and the root pulp in the lower third molars. **Science & Justice**, Elsevier, v. 57, n. 4, p. 257–261, 2017.
- VILA-BLANCO, N.; CARREIRA, M. J.; VARAS-QUINTANA, P.; BALSA-CASTRO, C.; TOMAS, I. Deep neural networks for chronological age estimation from opg images. **IEEE transactions on medical imaging**, IEEE, v. 39, n. 7, p. 2374–2384, 2020.
- VILA-BLANCO, N.; VARAS-QUINTANA, P.; TOMAS, I.; CARREIRA, M. J. A systematic overview of dental methods for age assessment in living individuals: from traditional to artificial intelligence-based approaches. **International Journal of Legal Medicine**, Springer, v. 137, n. 4, p. 1117–1146, 2023.
- VU, H. A. Integrating preprocessing methods and convolutional neural networks for effective tumor detection in medical imaging. **arXiv preprint arXiv:2402.16221**, 2024.
- 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.
- WANG, Z.; WANG, P.; LIU, K.; WANG, P.; FU, Y.; LU, C.-T.; AGGARWAL, C. C.; PEI, J.; ZHOU, Y. A comprehensive survey on data augmentation. **arXiv preprint arXiv:2405.09591**, 2024.
- YAVUZ, B. S.; EKMEKCIOLLU, O.; ANKARALI, H. Comparison of different dental age estimation methods with deep learning: Willems, cameriere-european, london atlas. **International Journal of Legal Medicine**, Springer, v. 139, n. 4, p. 1661–1672, 2025.
- 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.
- ZANGANA, H. M.; MOHAMMED, A. K.; MUSTAFA, F. M. Advancements in edge detection techniques for image enhancement: A comprehensive review. **International Journal of Artificial Intelligence & Robotics (IJAIR)**, v. 6, n. 1, p. 29–39, 2024.