



UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS DE SOBRAL
CURSO DE GRADUAÇÃO EM ENGENHARIA DE COMPUTAÇÃO

VINICIUS COSTA DOS SANTOS

**CLASSIFICAÇÃO DE CISTOS ODONTOLÓGICOS EM RADIOGRAFIAS
PANORÂMICAS USANDO ARQUITETURAS U-NET E TRANSFER LEARNING**

SOBRAL

2026

VINICIUS COSTA DOS SANTOS

CLASSIFICAÇÃO DE CISTOS ODONTOLÓGICOS EM RADIOGRAFIAS PANORÂMICAS
USANDO ARQUITETURAS U-NET E TRANSFER LEARNING

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: Dr. Iális Cavalcante de Paula Junior.

SOBRAL

2026

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)

S239c Santos, Vinicius Costa dos.
Classificação de Cistos Odontológicos em Radiografias Panorâmicas Usando Arquiteturas U-Net e Transfer Learning / Vinicius Costa dos Santos. – 2026.
58 f. : il.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Campus de Sobral, Curso de Engenharia da Computação, Sobral, 2026.
Orientação: Prof. Dr. Iális Cavalcante de Paula Junior.

1. Cistos Odontogênicos. 2. Radiografia Panorâmica. 3. CNN. 4. U-Net. I. Título.

CDD 621.39

VINICIUS COSTA DOS SANTOS

CLASSIFICAÇÃO DE CISTOS ODONTOLÓGICOS EM RADIOGRAFIAS PANORÂMICAS
USANDO ARQUITETURAS U-NET E TRANSFER LEARNING

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: 26/01/2026.

BANCA EXAMINADORA

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

Dr. Jarbas Joaci de Mesquita Sá Junior
Universidade Federal do Ceará (UFC)

Evecleison Albuquerque do Nascimento
Universidade Federal do Ceará (UFC)

A Deus.

AGRADECIMENTOS

A Deus.

Ao Estado Democrático de Direito do Brasil e à Educação Superior Pública Gratuita e de Qualidade (que recebi por meio da UFC).

Ao Prof. Dr. Iális Cavalcante de Paula Junior, por toda orientação e suporte durante minha passagem pela UFC.

Aos meus amigos do famoso "Uber", Jefferson Marques, Emanuel Valério, Emanuel Dêvid, Frank Willian, Vitor Hugo, João Vitor e Igor Pierre. Vocês têm uma grande importância em minha trajetória.

Aos servidores técnico-administrativos do *campus* Sobral, que sempre foram extremamente prestativos e gentis, mesmo além de suas obrigações formais.

Aos meus pais e à minha irmã.

À minha namorada que sempre esteve ao meu lado, Joviana.

Aos meus queridos amigos da vida que me acompanham desde o início no CVP. Isso mesmo, vocês sabem que estou falando de vocês. Em especial, ao meu melhor amigo, João Renato.

"O essencial é invisível aos olhos." (ANTOINE
DE SAINT-EXUPÉRY, 1943, *O Pequeno
Príncipe.*)

RESUMO

Os cistos odontogênicos são lesões benignas dos maxilares que podem provocar alterações estruturais e desconforto ao paciente, tornando relevante sua detecção precoce em exames de imagem. Entre as técnicas diagnósticas mais utilizadas na Odontologia, a radiografia panorâmica permite a visualização global da região bucomaxilofacial e subsidia a tomada de decisão clínica. Nesse contexto, métodos baseados em Inteligência Artificial (IA) podem atuar como ferramenta de apoio ao especialista, reduzindo subjetividades e ampliando a eficiência do processo diagnóstico. Assim, este trabalho investiga a metodologia mais eficaz para classificar radiografias panorâmicas quanto à presença ou ausência de cistos odontogênicos, por meio de arquiteturas baseadas em U-Net adaptadas para *classificação binária*. Foram utilizadas duas fontes de dados: uma base construída em colaboração com especialistas, composta por 604 radiografias panorâmicas, e a Tufts Dental Database, da qual foram selecionadas 24 imagens com regiões císticas de interesse. As regiões foram anotadas com o software LabelImg e, a partir das coordenadas, realizou-se a segmentação automática das áreas marcadas, totalizando 1219 sub-imagens císticas e 1212 sub-imagens não císticas.

Para aprimorar a qualidade das entradas, aplicaram-se técnicas de pré-processamento (filtro BlackHat e Correção Gamma). Em seguida, foram implementadas e avaliadas as versões U-Net v1, v2 e v3, com e sem *transfer learning* utilizando EfficientNetB0 como *backbone*. A validação foi conduzida por validação cruzada estratificada (Stratified K-Fold) com 4 *folds*, considerando as métricas de acurácia, sensibilidade, especificidade e F1-score, visando identificar a configuração mais robusta para o problema proposto.

Palavras-chave: cistos odontogênicos; radiografia panorâmica; U-Net; transfer learning; pré-processamento de imagens.

ABSTRACT

Odontogenic cysts are benign jaw lesions that may lead to structural changes and patient discomfort, making early detection clinically relevant. Panoramic radiography is widely used in Dentistry because it provides a global view of the maxillomandibular region and supports clinical decision-making. In this context, Artificial Intelligence (AI) methods can support specialists by reducing subjectivity and improving diagnostic efficiency.

This work investigates the most effective methodology for classifying panoramic radiographs according to the presence or absence of odontogenic cysts using U-Net-based architectures adapted for *binary classification*. Two data sources were considered: a dataset built in collaboration with experts, comprising 604 panoramic radiographs, and the Tufts Dental Database, from which 24 images containing cyst regions of interest were selected. Regions of interest were annotated with LabelImg and automatically cropped based on the recorded coordinates, yielding 1219 cystic and 1212 non-cystic sub-images.

To improve input quality, image pre-processing techniques were applied (BlackHat filter and Gamma Correction). Subsequently, U-Net v1, v2, and v3 variants were implemented and evaluated, with and without *transfer learning* using EfficientNetB0 as a pretrained *backbone*. Model validation was performed via stratified cross-validation (Stratified K-Fold) with 4 folds, using accuracy, sensitivity, specificity, and F1-score as evaluation metrics, aiming to identify the most robust configuration for the proposed task.

Keywords: odontogenic cysts; panoramic radiography; U-Net; transfer learning; image pre-processing.

LISTA DE FIGURAS

Figura 1 – Exemplos de radiografias intrabuciais: (a) oclusal; (b) interproximal; (c) periapical.	21
Figura 2 – Exemplo de radiografia panorâmica.	22
Figura 3 – Exemplo de lesão radiolúcida em radiografia panorâmica.	22
Figura 4 – Exemplo de operação de convolução.	24
Figura 5 – Exemplo ilustrativo da estrutura de uma <i>Convolutional Neural Network</i> (CNN).	26
Figura 6 – Visão geral do <i>pipeline</i> metodológico: aquisição, organização e anonimização das radiografias; rotulagem; segmentação em sub-imagens; realce e normalização; aumento de dados; treinamento/validação; e avaliação.	32
Figura 7 – Matrizes de confusão para os melhores resultados gerados em cada experimento.	47

LISTA DE TABELAS

Tabela 1 – Síntese dos principais pontos dos trabalhos relacionados	15
Tabela 2 – Arquitetura MBConv (síntese).	27
Tabela 3 – Arquitetura da EfficientNetB0 (síntese).	27
Tabela 4 – Acurácia média (%) e desvio padrão para diferentes variações da U-Net e filtros de realce, com validação cruzada estratificada em 4 <i> folds</i>	43
Tabela 5 – Métricas de desempenho da U-NetV1 com e sem <i> backbone</i> , sob diferentes filtros de realce.	44
Tabela 6 – Métricas de desempenho da U-NetV2 com e sem <i> backbone</i> , sob diferentes filtros de realce.	45
Tabela 7 – Métricas de desempenho da U-NetV3 com e sem <i> backbone</i> , sob diferentes filtros de realce.	46

LISTA DE ABREVIATURAS E SIGLAS

ATM	articulações temporomandibulares
CNN	<i>Convolutional Neural Network</i>
DL	<i>Deep Learning</i>
IA	Inteligência Artificial
LGPD	Lei Geral de Proteção de Dados Pessoais
ML	<i>Machine Learning</i>
PAN	radiografia panorâmica
UFC	Universidade Federal do Ceará
VC	Visão Computacional

SUMÁRIO

1	INTRODUÇÃO	14
1.1	Trabalhos Relacionados	15
1.2	Justificativa	17
1.2.1	<i>Adequação à legislação vigente</i>	18
1.3	Objetivo Geral	19
1.4	Objetivos Secundários	19
2	FUNDAMENTAÇÃO TEÓRICA	20
2.1	Cistos odontogênicos	20
2.2	Imagens odontológicas	21
2.3	Sistemas de Visão Computacional	23
2.3.1	<i>Pré-processamento</i>	23
2.3.2	<i>Filtros de imagens e convolução</i>	23
2.3.2.1	<i>Correção Gama</i>	24
2.3.2.2	<i>Morfologia e BlackHat</i>	25
2.4	Segmentação	25
2.5	<i>Machine Learning e Deep Learning</i>	25
2.5.1	<i>Redes Neurais Convolucionais</i>	26
2.6	EfficientNet	27
2.7	Função de perda	28
2.8	Otimizadores	28
2.9	Rede U-Net e sua adaptação para classificação	28
2.9.1	<i>Estrutura encoder-decoder e skip connections</i>	29
2.10	<i>Transfer learning e uso de backbones pré-treinados</i>	30
2.11	Classificação binária em imagens radiográficas	30
3	METODOLOGIA	32
3.1	Bases de imagens e organização inicial	32
3.1.1	<i>Anonimização e conformidade ética</i>	33
3.1.2	<i>Critérios de inclusão e exclusão</i>	33
3.2	Rotulagem supervisionada e construção do dataset por recortes	34
3.2.1	<i>Ferramenta e procedimento de anotação</i>	34

3.2.2	<i>Segmentação das radiografias em sub-imagens</i>	34
3.3	Pré-processamento e realce de imagens	35
3.3.1	<i>Padronização de tamanho e normalização de intensidades</i>	35
3.3.2	<i>Realce por operação morfológica Black-hat</i>	35
3.3.3	<i>Realce por correção Gamma</i>	36
3.3.4	<i>Cenários experimentais de realce</i>	36
3.4	Aumento de dados (Data Augmentation)	37
3.5	Arquiteturas propostas e adaptação da U-Net para classificação	37
3.5.1	<i>Adaptação para classificação</i>	37
3.5.2	<i>Variações de U-Net e backbone EfficientNetB0</i>	38
3.6	Configuração de treinamento e validação	38
3.6.1	<i>Divisão de dados e validação cruzada estratificada</i>	39
3.6.2	<i>Função de perda, otimizador e hiperparâmetros</i>	39
3.6.3	<i>Callbacks e controle de sobreajuste</i>	39
3.6.4	<i>Ambiente computacional e reprodutibilidade</i>	40
3.7	Métricas de avaliação e análise de desempenho	40
3.8	Síntese do protocolo experimental	41
4	RESULTADOS	42
4.1	Visão geral dos experimentos	42
4.2	Resultados com validação cruzada estratificada	42
4.3	Resultados das métricas finais por arquitetura	44
4.3.1	<i>U-NetV1 e U-NetV1B</i>	44
4.3.2	<i>U-NetV2 e U-NetV2B</i>	45
4.3.3	<i>U-NetV3 e U-NetV3B</i>	45
4.4	Matrizes de confusão dos melhores modelos	46
4.5	Análises comparativas e ablação	47
4.5.1	<i>Impacto do backbone pré-treinado</i>	47
4.5.2	<i>Impacto das técnicas de realce (BlackHat e Gamma)</i>	48
4.6	Discussão dos resultados	49
4.6.1	<i>Ameaças à validade e limitações do experimento</i>	49
4.7	Síntese dos principais achados	50
5	CONSIDERAÇÕES FINAIS	51

5.1	Contribuições do trabalho	52
5.2	Limitações	52
5.3	Trabalhos Futuros	53
	REFERÊNCIAS	55

1 INTRODUÇÃO

Na era científica atual, é senso comum que o eletromagnetismo dos raios X não beneficiou apenas a Medicina, mas também a Odontologia no uso da radiografia panorâmica (PAN). Tal técnica expõe o paciente à radiação uma única vez e fornece uma visualização global das estruturas que compõem o complexo maxilomandibular, isto é, os dentes e tecidos ósseos de suporte, além de anatomias adjacentes como a cavidade nasal, os seios maxilares e as articulações temporomandibulares (ATM) (FENYO-PEREIRA *et al.*, 2013).

A interpretação da PAN ajuda a reduzir os riscos cirúrgicos na região posterior da mandíbula - em anestésias, implantes, extrações dentárias e outros (SALVADOR *et al.*, 2010). Então, presume-se que análises errôneas de PANs geram mais demora no diagnóstico correto de algum problema de saúde bucal presente ou em prevenir um no futuro. Ademais, por se tratarem de representações bidimensionais de estruturas tridimensionais, as PANs podem apresentar variabilidade estrutural, ruídos e particularidades anatômicas que dificultam a análise clínica, sobretudo em situações em que há sobreposições e regiões de contraste limitado.

Tal situação pode ser mitigada, ou até mesmo prevenida, com a adoção de ferramentas analíticas mais eficientes, que aprendem com dados do passado, como, por exemplo, exames, laudos, prescrições médicas e outros. O desenvolvimento de uma ferramenta como esta auxilia o especialista humano a obter conclusões mais precisas com maior celeridade, bem como contribui na formação e no treinamento de futuros cirurgiões-dentistas.

Os avanços tecnológicos nas áreas de *Machine Learning* (ML), sua subárea *Deep Learning* (DL), e Visão Computacional (VC) têm impulsionado aplicações de Inteligência Artificial (IA) no cotidiano, inclusive nos campos médico e educacional, o que evidencia a aplicabilidade dessas técnicas também na Odontologia. Entre as diversas abordagens existentes, as redes neurais convolucionais (CNNs) têm se mostrado eficazes na identificação automática de padrões em imagens médicas, e arquiteturas consagradas podem ser adaptadas conforme a necessidade da tarefa.

Dentre essas arquiteturas, destaca-se a U-Net, originalmente desenvolvida para segmentação, pois utiliza uma estrutura *encoder-decoder* com conexões de atalho (*skip connections*) que preservam informações espaciais relevantes ao longo do processamento. Assim, ao adaptar variações da U-Net para um cenário de classificação binária, torna-se possível explorar sua capacidade de extração e refinamento de características, diferenciando regiões císticas de regiões não císticas, mesmo diante de ruídos e variações presentes em PANs.

Isto posto, este trabalho visa investigar, de forma sistemática, o desempenho de diferentes variações da arquitetura U-Net adaptadas para classificação binária de cistos odontogênicos em PANs. Além disso, avalia-se o impacto do *transfer learning* por meio do uso de um *backbone* pré-treinado (EfficientNetB0), bem como o efeito de técnicas de pré-processamento com filtros de realce (como BlackHat e Gamma Correction), buscando potencializar a capacidade de generalização dos modelos e contribuir para o aprimoramento de sistemas de apoio ao diagnóstico odontológico.

1.1 Trabalhos Relacionados

Na literatura científica, há pesquisas que aplicam técnicas de ML e DL para apoiar diagnósticos odontológicos, abrangendo desde a identificação de patologias até tarefas de classificação relacionadas a estruturas anatômicas. A Tabela 1 consolida os pontos centrais dos trabalhos discutidos a seguir, destacando tarefa, métodos empregados, conjunto de dados e melhores métricas reportadas.

Tabela 1 – Síntese dos principais pontos dos trabalhos relacionados

Autor(es)	Tarefa	Métodos	Base de dados	Melhores métricas
Oliveira et al., 2023	Identificar cistos dentígeros em terceiros molares	VGG16, Inception-V3, CNN1, CNN2; filtros e aumento de dados	333 radiografias odontológicas	Acc (81,99%); Prec (78,90%); Sens (86,17%); F1 (82,19%)
Deepak e Bhat, 2024	Classificar severidade de perda dentária para prótese	SqueezeNet, ResNet-50, EfficientNet-B0; aumento de dados	116 radiografias (segmentadas manualmente)	Acc (98%)
Ha et al., 2023	Detectar pseudocisto de retenção no seio maxilar	EfficientDet	213 radiografias (426 seios)	Precisão (81%)
Kumar et al., 2023	Distinguir cisto, tumor e abscesso por textura	SVM; GLCM, GLRLM, wavelet	172 radiografias odontológicas	Precisão (98%)

Fonte: Elaborada pelo autor.

Os autores de Oliveira *et al.* (2023) investigaram a identificação de cistos dentígeros associados a terceiros molares em imagens de radiografia odontológica por meio de redes convolucionais profundas. Além da comparação entre arquiteturas, o estudo deu ênfase ao papel do pré-processamento para tornar mais nítida a separação entre regiões de interesse e tecidos saudáveis. Para isso, os autores avaliaram diferentes estratégias de realce, incluindo ajustes de brilho/contraste, correção gamma (GONZALEZ; WOODS, 2008), operações morfológicas

(BIRDAL *et al.*, 2016) e *contrast limited adaptive histogram equalisation* (CLAHE) (PIZER *et al.*, 1987). A base utilizada continha 333 imagens, e o treinamento foi complementado por técnicas de aumento de dados (rotação, zoom, deslocamento e espelhamento), visando reduzir viés e melhorar generalização. No conjunto de modelos analisado (VGG16, Inception-V3 e duas CNNs projetadas pelos autores, CNN1 e CNN2), o melhor resultado foi obtido pela CNN2 em conjunto com correção gamma, registrando acurácia de 81,99%, precisão de 78,90%, sensibilidade de 86,17% e F1-score de 82,19%. Embora a Inception-V3 tenha se destacado em sensibilidade, o conjunto de métricas sugere que escolhas de pré-processamento podem ser determinantes mesmo quando se utilizam arquiteturas distintas.

Além disso, os autores de Deepak e Bhat (2024) propuseram uma abordagem de classificação para apoiar a avaliação clínica relacionada à necessidade de próteses dentárias. O estudo utilizou 116 radiografias odontológicas segmentadas manualmente e definiu três classes correspondentes ao nível de comprometimento da mandíbula (ideal, parcialmente danificada e totalmente danificada). Diferentemente de trabalhos que exploram filtros de realce, a ênfase aqui recaiu sobre o uso de *data augmentation* (como translação, rotação e mudança de escala) e sobre o aproveitamento de redes pré-treinadas. Foram comparadas SqueezeNet, ResNet-50 e EfficientNet-B0, e o melhor desempenho foi observado na EfficientNet-B0, com acurácia de 98%, superando ResNet-50 (92%) e SqueezeNet (91%). Assim, o estudo reforça a utilidade de *transfer learning* em bases pequenas, cenário recorrente em aplicações odontológicas.

Em outra perspectiva, Ha *et al.* (2023) trataram do problema de detecção de pseudocistos de retenção no seio maxilar em PANs, utilizando um modelo do tipo detector. A arquitetura adotada foi o EfficientDet, empregada para localizar e classificar os seios maxilares. Os autores consideraram 213 imagens, totalizando 426 seios, e avaliaram o desempenho em um conjunto de testes composto por casos de pseudocisto de retenção, seios saudáveis e casos rotulados como cisto ou tumor. No que se refere às métricas, o trabalho reportou precisão de 81% para a classe de pseudocisto de retenção, além de valores superiores para seios saudáveis e para a classe cisto/tumor. O resultado evidencia que detectores podem ser apropriados quando a tarefa exige a indicação espacial da região suspeita na imagem.

Por fim, Kumar *et al.* (2023) exploraram um caminho diferente ao investigar análise de textura para distinguir lesões do tipo cisto, tumor e abscesso em radiografias panorâmicas. A metodologia envolveu aprimoramento de imagem e segmentação manual da área de interesse, seguida de extração de características por descritores como matriz de coocorrência de níveis de

cinza (GLCM), matriz de comprimento de percurso em níveis de cinza (GLRLM) e wavelets. Com essas características, os autores empregaram uma máquina de vetores de suporte (SVM) para realizar a discriminação entre as classes, avaliando o desempenho por métricas como sensibilidade, especificidade e precisão. O estudo reportou que descritores baseados em GLCM atingiram precisão de 98% na separação entre as classes, enquanto wavelets e GLRLM também apresentaram resultados elevados. Dessa forma, o trabalho sugere que abordagens baseadas em características explícitas ainda podem ser competitivas em determinados cenários, principalmente quando as regiões de interesse são bem delimitadas.

Isto posto, observa-se que os trabalhos relacionados variam entre abordagens de detecção (com localização explícita), classificação direta e métodos baseados em extração de características. Ainda assim, é recorrente que artigos enfatizem os resultados finais e descrevam de forma limitada o trabalho bruto necessário para tornar um *software* reproduzível, incluindo escolhas de pré-processamento, padronização de recortes, balanceamento de classes e estratégia de validação. Portanto, há valor em consolidar essas etapas no limite de um Trabalho de Conclusão de Curso, especialmente quando se objetiva comparar arquiteturas sob um protocolo consistente e transparente.

1.2 Justificativa

No âmbito da análise mandibular e maxilomandibular por meio de radiografias PAN, vale mencionar que tais imagens, por representarem estruturas tridimensionais em um plano bidimensional, podem ser notoriamente difíceis de interpretar até mesmo por profissionais experientes. Além das limitações inerentes ao tipo de projeção, é comum que exames apresentem variações de contraste, ruídos, sobreposições de estruturas anatômicas e artefatos decorrentes do posicionamento do paciente e do próprio equipamento, fatores que podem dificultar a delimitação visual de regiões suspeitas e aumentar a subjetividade do julgamento clínico.

Desse modo, o ato de interpretar a imagem impacta diretamente o sucesso ou o fracasso de um diagnóstico em radiologia. Isso exige estudo e prática contínuos, sobretudo durante o processo formativo, para que o estudante consiga distinguir o que é normal de suas variações e, assim, reconhecer padrões patológicos. Ademais, determinadas estruturas anatômicas presentes na maxila e na mandíbula podem ser confundidas com lesões, o que torna ainda mais relevante o desenvolvimento de ferramentas que apoiem a análise, principalmente quando há suspeita de alterações que demandam investigação e acompanhamento.

Diante disso, justifica-se o presente trabalho em contribuir para o desenvolvimento de um *software* que possa auxiliar dentistas a obterem uma segunda opinião, ao mesmo tempo em que fortalece a formação de estudantes, oferecendo um suporte automatizado para triagem e interpretação de radiografias PAN. Essa proposta se insere em um contexto interdisciplinar, no qual métodos de ML e DL podem ser aplicados ao processamento de imagens odontológicas para reduzir variabilidade interobservador e aumentar a celeridade na identificação de padrões sugestivos de cistos odontogênicos.

Para isso, parte do caminho envolveu organizar e preparar bases de dados de radiografias PAN provenientes de duas fontes: uma base construída em colaboração com especialistas e uma base pública, totalizando centenas de radiografias que, após rotulação e segmentação das regiões de interesse, geraram um conjunto expressivo de subimagens císticas e não císticas. Além da construção do conjunto de dados, foram definidas etapas de pré-processamento para realçar estruturas relevantes (como filtros BlackHat e Correção Gamma), bem como estratégias de avaliação robusta por validação cruzada estratificada. Somado a isso, foram comparadas diferentes variações da arquitetura U-Net adaptadas para classificação binária, com e sem *transfer learning* por meio de *backbones* pré-treinados.

Tais esforços visam consolidar um protocolo reprodutível e transparente, que possa ser expandido em trabalhos futuros e contribuir tanto para o avanço do uso de IA na saúde oral quanto para o amadurecimento de pesquisas em visão computacional aplicada a radiografias panorâmicas. Assim, ao sistematizar etapas frequentemente tratadas de forma superficial em artigos (como seleção de recortes, padronização de entrada, pré-processamento e validação), este Trabalho de Conclusão de Curso busca entregar não apenas resultados, mas também um caminho metodológico consistente para estudos posteriores e potencial aplicação em cenários de apoio ao diagnóstico.

1.2.1 Adequação à legislação vigente

É digno de nota que as amostras de PANs cedidas pelo curso de Odontologia da Universidade Federal do Ceará (UFC) Sobral tinham identificação pessoal dos pacientes dos quais foram colhidas. Então, foi preciso operacionalizar um fluxo onde seus passos levem à devida anonimização de cada amostra.

Porém, devia-se garantir que certas características biológicas das mesmas fossem preservadas (por exemplo, a idade biológica) por serem úteis na obtenção de conclusões por

parte dos *experts* consultados. Por exemplo, com base no tempo decorrido de vida do paciente, a classificação de uma PAN poderia oscilar entre conter um cisto odontológico, ou apenas algum elemento anatômico normal para uma dada faixa etária (como dentes que ainda não eclodiram).

Neste caso, tal necessidade foi imposta pela Lei Nº 13.709/2018 - Lei Geral de Proteção de Dados Pessoais (LGPD) , mais especificamente no Art. 11, inciso II, item c), que diz que um dos usos possíveis de dados sensíveis é na "realização de estudos por órgão de pesquisa, garantida, sempre que possível, a anonimização dos dados pessoais sensíveis" (BRASIL, 2018).

1.3 Objetivo Geral

O objetivo deste trabalho é propor e avaliar uma metodologia reproduzível para classificar automaticamente radiografias PAN quanto à presença ou ausência de cistos odontológicos, investigando arquiteturas baseadas na U-Net adaptadas para classificação, com e sem *transfer learning* e técnicas de pré-processamento de imagens, visando contribuir como suporte ao diagnóstico em Odontologia.

1.4 Objetivos Secundários

- Preparar um banco de dados para a tarefa, segmentando automaticamente regiões de interesse a partir de radiografias PAN rotuladas, de modo a formar amostras císticas e não císticas para treinamento e testes;
- Desenvolver e avaliar variantes da arquitetura U-Net (v1, v2 e v3) adaptadas para classificação binária (cisto ou saudável);
- Investigar o impacto do *transfer learning* ao empregar a EfficientNetB0 como *backbone* nas arquiteturas propostas;
- Avaliar o efeito de técnicas de pré-processamento, em especial os filtros BlackHat e a Correção Gamma, no desempenho dos modelos;
- Comparar as abordagens desenvolvidas por meio de métricas estatísticas, incluindo acurácia, sensibilidade, especificidade e F1-score, de forma a identificar a estratégia mais eficaz para o problema em questão.

2 FUNDAMENTAÇÃO TEÓRICA

Nesta seção são apresentados os conceitos teóricos necessários para contextualizar o problema de detecção de cistos odontogênicos em radiografias PAN e para fundamentar as técnicas computacionais utilizadas ao longo deste trabalho. Assim, discute-se a caracterização clínica e radiográfica dos cistos odontogênicos, aspectos de imagens odontológicas, o fluxo típico de um sistema de visão computacional, técnicas de pré-processamento e filtragem, além de noções essenciais de ML e DL aplicadas à classificação de imagens.

2.1 Cistos odontogênicos

Os cistos odontogênicos consistem em cavidades patológicas revestidas por epitélio odontogênico e localizadas nos ossos gnáticos. Em geral, apresentam crescimento lento e podem permanecer assintomáticos por longos períodos, sobretudo quando não há infecção associada. Quando há progressão e/ou infecção, podem ocorrer inflamação, tumefação e dor, além de possível expansão e erosão da cortical, o que pode impactar a estrutura mandibular e gerar desconforto ao paciente (MALLYA; LAM, 2018).

Do ponto de vista radiográfico, essas lesões tendem a se apresentar como áreas radiolúcidas, frequentemente com limites bem definidos, podendo ser uniloculares ou multiloculares (WHITE; PHAROAH, 2013). A relação anatômica com elementos dentários e estruturas adjacentes é um fator central para o reconhecimento do padrão suspeito em exames de imagem. Além disso, conforme a origem, os cistos odontogênicos são comumente agrupados em inflamatórios e de desenvolvimento (NEVILLE *et al.*, 2009).

Os cistos inflamatórios surgem, em geral, a partir da proliferação epitelial em resposta a processos inflamatórios locais. Alterações pulpares e infecções dos canais radiculares, por exemplo, podem levar a lesões que envolvem reabsorção óssea, decorrente da tentativa do organismo de conter agentes infecciosos. Já os cistos de desenvolvimento podem ser descritos como cavidades patológicas compostas por cápsula e epitélio, com preenchimento líquido ou mole. Em termos clínicos, pode-se interpretá-los como formações semelhantes a bolsas cuja parede apresenta comportamento semipermeável, característica associada ao crescimento expansivo. Ademais, a literatura descreve diferentes subtipos dentro dessa categoria, conforme classificações consolidadas na área.

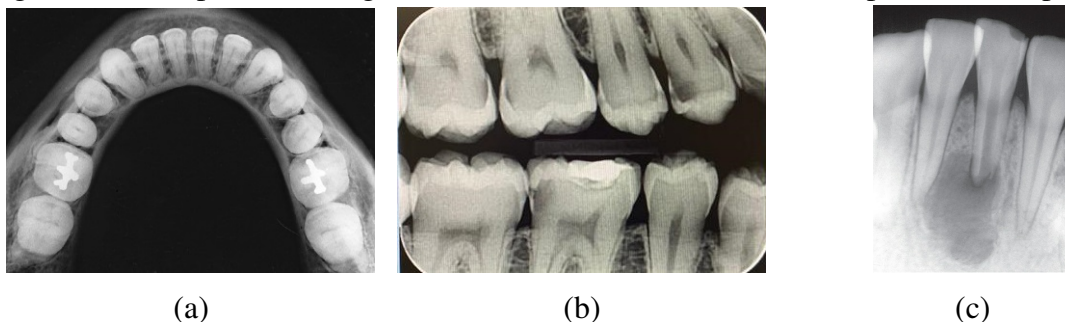
Quanto às manifestações clínicas, cistos odontogênicos podem estar associados a tumefação, descarga purulenta, parestesia, mobilidade dentária, dor e assimetria facial (VARINAUSKAS *et al.*, 2006). Assim, a detecção precoce e o encaminhamento adequado são relevantes, visto que a propagação de infecções pode resultar em complicações sistêmicas (SOUSA *et al.*, 2003). Além de potencialmente reduzir custos e simplificar o tratamento (NURTANIO *et al.*, 2011), o diagnóstico prévio pode evitar procedimentos desnecessários. Nesse cenário, radiografias são fundamentais por evidenciarem alterações em tecidos não visíveis em exames clínicos convencionais (LI *et al.*, 2007).

2.2 Imagens odontológicas

A radiografia é considerada um marco inicial das imagens médicas modernas, possibilitando a visualização bidimensional de estruturas internas. Historicamente, o uso dos raios X remonta ao final do século XIX, associado aos estudos de Wilhelm Röntgen, e rapidamente foi incorporado ao contexto odontológico (FEJERSKOV *et al.*, 2018). Em sua forma digital, a radiografia utiliza sensores capazes de converter a radiação em sinais elétricos digitalizados, permitindo armazenamento, processamento e análise computacional das imagens (GONZALEZ; WOODS, 2010).

As radiografias PAN bidimensionais podem ser classificadas em intrabucais e extrabucais. As intrabucais incluem, por exemplo, radiografias periapicais, interproximais e oclusais, que variam quanto ao campo de visão e às estruturas enfatizadas (ALVARES *et al.*, 1965). A Figura 1 ilustra exemplos desse grupo.

Figura 1 – Exemplos de radiografias intrabucais: (a) oclusal; (b) interproximal; (c) periapical.



Fonte: (a) Adaptada de (WHITE; PHAROAH, 2013); (b) Obtida em parceria com o curso de Odontologia da UFC; (c) Adaptada de (LOPES; JUNIOR, 1999).

Entre as extrabucais, destacam-se modalidades como telerradiografia lateral, PAN, radiografia posteroanterior, radiografia da articulação temporomandibular e tomografia computa-

dorizada de feixe cônico. Dentre essas, a PAN é amplamente empregada por fornecer uma visão global do complexo maxilomandibular. A Figura 2 apresenta um exemplo de PAN.

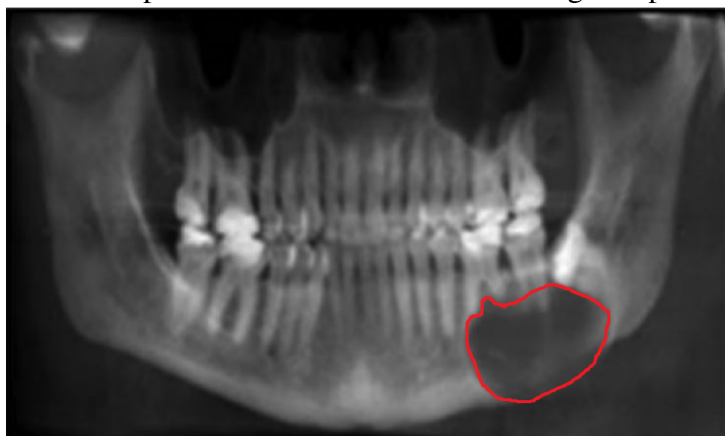
Figura 2 – Exemplo de radiografia panorâmica.



Fonte: Cedida pelo curso de Odontologia, UFC campus Sobral.

Em radiografias, cistos tendem a ser observados como lesões radiolúcidas, com frequência de formato oval ou arredondado, associadas a conteúdo fluido (COSTA *et al.*, 2016). Em casos de longa duração, pode ocorrer calcificação distrófica, alterando a aparência interna da lesão. Apesar do crescimento geralmente lento, essas regiões podem provocar deslocamento de estruturas, reabsorção dentária e expansão mandibular, formando um limite cortical fino nas áreas acometidas (SANTOS *et al.*, 2007). A Figura 3 apresenta um exemplo ilustrativo de lesões radiolúcidas em radiografias PAN.

Figura 3 – Exemplo de lesão radiolúcida em radiografia panorâmica.



Fonte: Adaptado de (SIVASUNDARAM; PANDIAN, 2021).

2.3 Sistemas de Visão Computacional

A VC é uma área que busca habilitar computadores a extrair informações de imagens e sensores visuais, apoiando a automatização de tarefas complexas por meio de algoritmos e ferramentas computacionais (QUEIROZ; GOMES, 2006). Definida classicamente por Ballard e Brown, essa ciência estuda métodos e tecnologias que permitem a máquinas “verem” e interpretar dados visuais (BALLARD; BROWN, 1982). Atualmente, é um campo interdisciplinar que integra processamento de imagens, inteligência artificial, reconhecimento de padrões e elementos de neurociência, com aplicações que incluem monitoramento, robótica, veículos autônomos e diagnóstico médico (NEVES *et al.*, 2012; SILVA *et al.*, 2019; VAZ, 2023; JUNIOR, 2023).

Em termos práticos, sistemas de VC costumam envolver: aquisição de imagem, pré-processamento, realce e/ou remoção de ruído, segmentação de regiões de interesse, extração de características (implícita ou explícita) e, por fim, classificação/decisão. Em particular, quando se utiliza DL, várias dessas etapas podem ser incorporadas ao próprio modelo, mas o pré-processamento continua sendo relevante em cenários com variabilidade de qualidade de aquisição.

2.3.1 Pré-processamento

O pré-processamento de imagens busca realçar aspectos estruturais e reduzir elementos indesejados, preparando os dados para etapas posteriores, como segmentação, reconhecimento de padrões e ML. Conforme a necessidade do problema, podem ser aplicadas operações aritméticas, ajustes de contraste e tratamento de ruído, além de transformações geométricas como inversão, rotação, translação, ajuste de escala e correções de perspectiva (GONZALES; WINTZ, 1987). Assim, o conjunto de técnicas selecionado deve ser coerente com o objetivo do sistema e com as características do domínio (por exemplo, radiografias com contraste limitado e presença de artefatos).

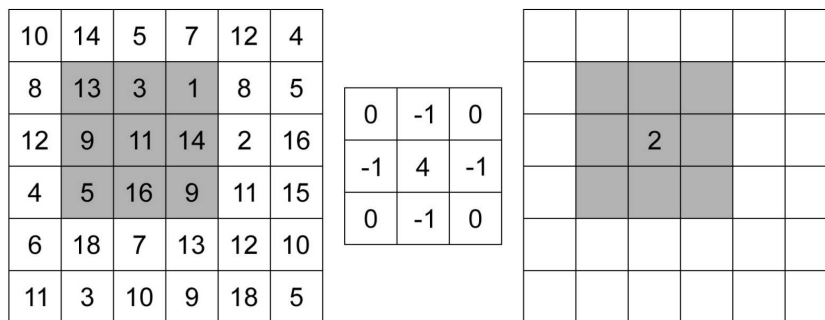
2.3.2 Filtros de imagens e convolução

A filtragem é um conjunto de técnicas amplamente utilizado em processamento digital de imagens. Em termos gerais, pode-se dividir os métodos em filtragem no domínio espacial e no domínio da frequência. No domínio espacial, operações são aplicadas diretamente à matriz de pixels, frequentemente por meio de convolução com máscaras (*kernels*) (ABDURRAZZAQ

et al., 2020). Já no domínio da frequência, modificações são realizadas sobre representações como a transformada de Fourier.

Em uma filtragem espacial típica, a máscara percorre a imagem e modifica valores de pixels conforme uma regra local. Na convolução, realiza-se um produto elemento a elemento entre a vizinhança do pixel-alvo e o *kernel*, acumulando-se o resultado para produzir o novo valor (KIM; CASPER, 2013). A Figura 4 exemplifica o processo.

Figura 4 – Exemplo de operação de convolução.



Fonte: Obtido de (BARELLI, 2018).

2.3.2.1 Correção Gama

A correção gama é uma técnica de ajuste não linear de brilho e contraste, aplicada por uma transformação sobre os valores de intensidade. Em termos práticos, a operação eleva os valores normalizados a uma potência controlada por um parâmetro γ , permitindo realçar detalhes em regiões muito escuras ou atenuar áreas muito claras. Uma forma comum de expressar essa transformação é:

$$g(x) = \left(\frac{f(x)}{255} \right)^{\frac{1}{\gamma}} \times 255 \quad (2.1)$$

onde $f(x)$ representa a intensidade original do pixel e $g(x)$ a intensidade após a correção. Valores de $\gamma < 1$ tendem a clarear a imagem e destacar regiões escuras; valores de $\gamma > 1$ tendem a escurecer a imagem, podendo reduzir saturação em regiões claras. Por isso, a escolha de γ deve ser guiada pela característica do dado e pelo objetivo do realce.

2.3.2.2 *Morfologia e BlackHat*

Operações morfológicas buscam alterar a forma e a estrutura de objetos em uma imagem, com uso de um elemento estruturante (*kernel*) que percorre a matriz de pixels (GONZALEZ; WOODS, 2010). São frequentemente aplicadas em imagens binárias, mas também podem ser adaptadas para imagens em tons de cinza, visando reduzir ruído e realçar regiões de interesse.

Entre operações básicas, destacam-se erosão (reduz objetos, “corroendo” bordas) e dilatação (expande objetos). A abertura consiste em erosão seguida de dilatação, geralmente útil para remoção de ruídos pequenos; o fechamento aplica dilatação seguida de erosão, comumente empregado para preencher pequenas lacunas. Ademais, existem operações derivadas, como o BlackHat, definido como a diferença entre o fechamento da imagem (com um elemento estruturante) e a imagem original, sendo útil para evidenciar estruturas mais escuras em relação ao entorno.

2.4 Segmentação

A segmentação é uma etapa fundamental em muitos sistemas de visão computacional, pois busca isolar o objeto de interesse do restante da imagem. Em termos conceituais, considera-se que os pixels pertencentes ao objeto formam o primeiro plano, enquanto o restante corresponde ao fundo, que pode ser descartado para facilitar análises posteriores (KAUR; KAUR, 2014). Em aplicações odontológicas, segmentar regiões suspeitas pode reduzir interferências de estruturas não relevantes e contribuir para a robustez da classificação.

2.5 *Machine Learning e Deep Learning*

Algoritmos de ML podem ser classificados conforme seu modo de treinamento e o grau de supervisão disponível. De modo geral, há algoritmos treinados em lote (*batch*) ou de forma incremental (*online*), além de paradigmas supervisionado, não supervisionado e semi-supervisionado (GÉRON, 2019). No aprendizado supervisionado, utiliza-se um conjunto de exemplos rotulados para que o modelo aprenda uma função capaz de generalizar para dados novos. Já no não supervisionado, busca-se extrair estruturas ou padrões sem rótulos; no semi-supervisionado, combina-se um conjunto reduzido de rótulos com grande volume de dados não rotulados.

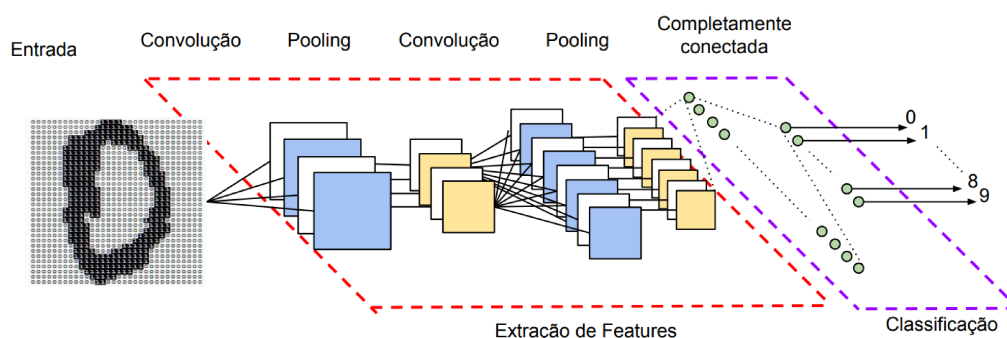
No contexto de imagens, técnicas baseadas em CNNs destacam-se por aprenderem representações diretamente dos pixels, reduzindo a necessidade de engenharia manual de características. Ainda assim, o desempenho do modelo depende da qualidade do conjunto de dados, do balanceamento das classes e do protocolo de validação.

2.5.1 Redes Neurais Convolucionais

Redes neurais convencionais recebem entradas vetorizadas e unidimensionais. Em problemas de imagem, tal representação tende a ser inadequada, pois ignora relações espaciais e pode tornar a dimensionalidade inviável. Para contornar esse cenário, CNNs foram desenvolvidas como mecanismos para reconhecimento de padrões visuais (FUKUSHIMA; MIYAKE, 1982). Desde então, tornaram-se base para reconhecimento e classificação de imagens em diversas aplicações.

De forma resumida, CNNs aprendem filtros que atribuem pesos e vieses a padrões relevantes, construindo mapas de características por camadas convolucionais (FERLITSCH, 2021). Em geral, uma arquitetura inclui camadas de convolução, funções de ativação e camadas de *pooling*, além de camadas totalmente conectadas na fase final (dependendo do desenho do modelo). As camadas de *pooling* sintetizam mapas de características e ajudam a reduzir número de parâmetros e risco de sobreajuste (SUN *et al.*, 2017). Funções de ativação (como ReLU e *Softmax*) introduzem não-linearidade e permitem que o modelo represente relações complexas (STEVENS *et al.*, 2020). Além disso, camadas como *Dropout* são frequentemente utilizadas para regularização. A Figura 5 apresenta um exemplo ilustrativo de uma arquitetura típica de CNN, destacando o fluxo entre as etapas de convolução, *pooling* e classificação.

Figura 5 – Exemplo ilustrativo da estrutura de uma CNN.



Fonte: Obtido de (VARGAS *et al.*, 2016).

2.6 EfficientNet

À medida que problemas de classificação se tornam mais complexos, é comum aumentar a largura (número de canais) e a profundidade (número de camadas) de redes convolucionais, o que normalmente exige ajustes empíricos até atingir uma arquitetura adequada. Para estruturar esse processo, a família EfficientNet propõe um dimensionamento uniforme e sistemático, escalando largura, profundidade e resolução de entrada de forma coordenada (TAN; LE, 2019).

Além disso, EfficientNet utiliza blocos baseados em convoluções separáveis e em ideias difundidas por arquiteturas eficientes, como os blocos residuais de gargalo invertido (inspirados na MobileNetV2) (SANDLER *et al.*, 2018). Esses blocos (frequentemente referenciados como MBConv) expandem canais por convolução 1×1 , aplicam convolução *depthwise* por canal e, por fim, comprimem a representação novamente por convolução 1×1 . Assim, é possível reduzir custo computacional mantendo desempenho competitivo. A Tabela 2 sintetiza as principais operações do bloco MBConv, indicando a transformação de dimensões ao longo das etapas. Em seguida, a Tabela 3 apresenta uma visão geral da arquitetura da EfficientNetB0, destacando as etapas, resoluções e número de canais.

Tabela 2 – Arquitetura MBConv (síntese).

Entrada	Operação	Saída
$h \times w \times k$	Conv2d, $k = 1 \times 1$, ReLU6	$h \times w \times (tk)$
$h \times w \times tk$	Dwise, $k = 3 \times 3$, $s=s$, ReLU6	$\frac{h}{s} \times \frac{w}{s} \times (tk)$
$\frac{h}{s} \times \frac{w}{s} \times (tk)$	Conv2d, $k=1 \times 1$	$\frac{h}{s} \times \frac{w}{s} \times K'$

Fonte: Baseado em (SANDLER *et al.*, 2018).

Tabela 3 – Arquitetura da EfficientNetB0 (síntese).

Etapa	Operação	Resolução	Canais	Camadas
1	Conv2d $k=3 \times 3$	224×224	32	1
2	MBConv1 $k=3 \times 3$	112×112	16	1
3	MBConv6 $k=3 \times 3$	112×112	24	2
4	MBConv6 $k=5 \times 5$	56×56	40	2
5	MBConv6 $k=3 \times 3$	28×28	80	3
6	MBConv6 $k=5 \times 5$	14×14	112	3
7	MBConv6 $k=5 \times 5$	14×14	192	4
8	MBConv6 $k=3 \times 3$	7×7	320	1
9	Conv2d $k=1 \times 1$, Pooling, FC	7×7	1280	1

Fonte: Baseado em (TAN; LE, 2019).

2.7 Função de perda

Funções de perda quantificam o erro cometido pelo modelo durante o treinamento e orientam o ajuste dos parâmetros para reduzir esse erro ao longo das épocas. Em problemas de classificação, uma função comumente empregada é a entropia cruzada, que compara a distribuição dos rótulos verdadeiros com a distribuição de saída produzida pelo modelo. Uma forma usual de representar a entropia cruzada (CE) é:

$$CE = \sum_i^C y_i \log(s_i) \quad (2.2)$$

onde C é o número de classes, y representa a distribuição de rótulos verdadeiros e s a distribuição de probabilidades produzida na camada de saída.

2.8 Otimizadores

Para minimizar a função de perda, os pesos e hiperparâmetros do modelo são atualizados por algoritmos de otimização, buscando acelerar a convergência e melhorar a estabilidade do treinamento. Um método amplamente utilizado é o Adam, que estima momentos adaptativos e ajusta taxas de aprendizado por parâmetro (KINGMA; BA, 2014). Em linhas gerais, o Adam utiliza termos para a média exponencialmente decrescente dos gradientes e para a média exponencialmente decrescente dos quadrados dos gradientes, o que o torna prático em cenários com grande número de parâmetros. Uma discussão mais abrangente sobre métodos de otimização pode ser encontrada em (RUDER, 2016).

2.9 Rede U-Net e sua adaptação para classificação

Em aplicações biomédicas de VC, é comum que as regiões relevantes na imagem sejam pequenas, ruidosas e cercadas por tecidos com textura semelhante, o que exige modelos com boa capacidade de preservar informação espacial ao longo do processamento. Nesse cenário, a arquitetura *U-Net* tornou-se uma referência por combinar uma estrutura do tipo *encoder-decoder* com conexões de atalho (*skip connections*), permitindo que detalhes de baixa profundidade (bordas, contornos e padrões locais) sejam reaproveitados em níveis mais altos de abstração (RONNEBERGER *et al.*, 2015).

A U-Net foi originalmente proposta para segmentação, isto é, para produzir uma máscara que indica, pixel a pixel, a região associada ao objeto de interesse. No entanto, suas propriedades arquiteturais (extração hierárquica de características e preservação de informação espacial via *skips*) também podem ser exploradas em cenários onde o objetivo final não é segmentar, mas decidir se a imagem pertence a uma classe clínica ou não. Em outras palavras, apesar de ser uma rede tradicionalmente aplicada em tarefas de segmentação em imagem médica, sua capacidade de extrair e refinar *features* em diferentes escalas pode ser reaproveitada em tarefas de classificação (RONNEBERGER *et al.*, 2015).

No contexto deste trabalho, a U-Net é adaptada para classificação binária ao substituir a saída típica de segmentação por uma cabeça de classificação ao final do *decoder*. Em termos de implementação, isso pode ser realizado conectando o último bloco do *decoder* a um operador de agregação global (por exemplo, *Global Average Pooling*) e, em seguida, a camadas densas, produzindo uma decisão final para as classes *cisto* ou *saudável*. Essa adaptação permite que a rede mantenha sua habilidade de representar padrões espaciais relevantes, enquanto devolve uma predição no nível da imagem, em vez de uma máscara por pixel.

2.9.1 Estrutura *encoder-decoder* e *skip connections*

O *encoder* realiza uma sequência de convoluções e operações de redução de resolução (por exemplo, *pooling*), com o objetivo de transformar a imagem em um conjunto de representações latentes progressivamente mais abstratas. Já o *decoder* executa o processo inverso, recuperando resolução por meio de *upsampling* (por exemplo, convoluções transpostas) e refinando mapas de ativação. As *skip connections* unem níveis correspondentes do *encoder* e do *decoder*, preservando detalhes espaciais que tenderiam a se perder apenas com informação de alto nível (RONNEBERGER *et al.*, 2015).

Em aplicações em radiografias, esse desenho é particularmente útil porque as alterações de interesse podem possuir contraste discreto e contornos pouco definidos, sobretudo quando há sobreposições anatômicas ou ruídos inerentes ao exame. Dessa forma, reter detalhes locais ao longo da rede tende a ser benéfico para destacar padrões associados a regiões patológicas.

2.10 *Transfer learning* e uso de *backbones* pré-treinados

Outro aspecto recorrente em projetos de DL em saúde é a limitação de dados rotulados em larga escala. Uma estratégia amplamente utilizada para contornar esse problema é o *transfer learning*, no qual uma rede pré-treinada em um grande conjunto de imagens (como a ImageNet) é reaproveitada como extrator de características para um novo domínio (YOSINSKI *et al.*, 2014). Em geral, a motivação é que camadas iniciais aprendem filtros mais genéricos (bordas, gradientes e texturas), que são úteis em múltiplas tarefas; já camadas finais tendem a especializar-se e podem ser ajustadas ao problema-alvo por *fine-tuning* (YOSINSKI *et al.*, 2014).

No presente trabalho, versões da U-Net são avaliadas com e sem *transfer learning*, utilizando o EfficientNetB0 como *backbone* no *encoder* (TAN; LE, 2019). A família EfficientNet busca equilibrar desempenho e custo computacional por meio de um escalonamento composto (*compound scaling*) de profundidade, largura e resolução, o que a torna atrativa em cenários com restrição de recursos e com interesse em boa capacidade de generalização (TAN; LE, 2019). Nessa configuração, camadas intermediárias do *backbone* podem ser reutilizadas como pontos de *skip connection*, mantendo o formato em “U” e aproveitando representações já aprendidas.

Do ponto de vista experimental, esse desenho permite comparar dois comportamentos esperados: (i) treinar uma U-Net do zero pode exigir mais dados e maior tempo para estabilizar, e (ii) usar um *backbone* pré-treinado tende a acelerar a convergência e a fornecer uma base mais robusta de características, especialmente quando o conjunto rotulado não é massivo (YOSINSKI *et al.*, 2014).

2.11 Classificação binária em imagens radiográficas

Em tarefas de classificação binária, o modelo deve estimar a qual de duas classes uma amostra pertence. No contexto deste estudo, o objetivo é decidir se uma sub-imagem derivada da PAN contém ou não uma região cística, caracterizando as classes *cisto* e *saudável*. Uma prática comum é produzir a saída como uma probabilidade única por meio de *sigmoid* (ou, alternativamente, duas probabilidades via *softmax*) e empregar funções de perda apropriadas, como a entropia cruzada, amplamente utilizada em classificação supervisionada (GOODFELLOW *et al.*, 2016).

Entretanto, em aplicações clínicas, métricas agregadas como Acurácia podem não capturar sozinhas o comportamento desejado do sistema, sobretudo quando os custos de erro são

assimétricos (por exemplo, quando falsos negativos são mais críticos do que falsos positivos). Assim, é frequente complementar a avaliação com métricas como sensibilidade (capacidade de identificar positivos), especificidade (capacidade de identificar negativos) e F1-score (equilíbrio entre precisão e sensibilidade), pois elas oferecem uma visão mais informativa do desempenho em cenários de interesse diagnóstico (SOKOLOVA; LAPALME, 2009).

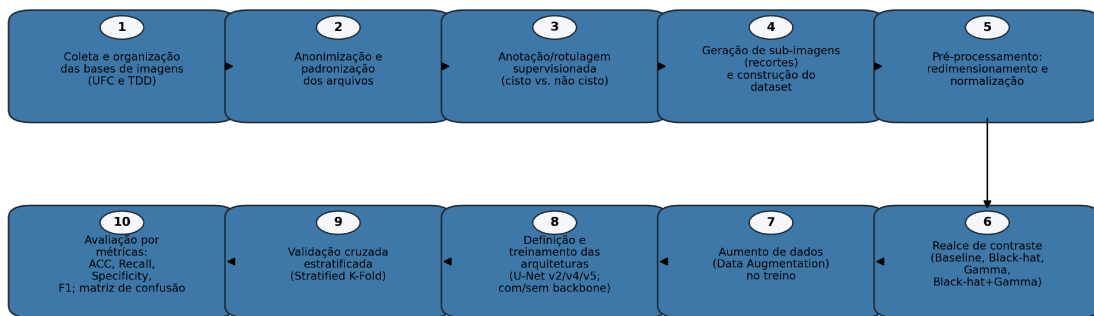
Por fim, é importante notar que a etapa de pré-processamento pode impactar diretamente a separabilidade entre classes, ao realçar características relevantes e reduzir ruídos. Dessa forma, técnicas de realce e transformações de contraste podem ser empregadas para evidenciar estruturas sutis em radiografias, contribuindo para que os modelos aprendam representações mais discriminativas (GONZALEZ; WOODS, 2008).

3 METODOLOGIA

Este capítulo descreve, de forma detalhada, a metodologia adotada para detecção automatizada de cistos odontológicos em radiografias PAN por meio de modelos baseados em U-Net. A metodologia foi planejada para garantir (i) rastreabilidade de todas as etapas do *pipeline*, (ii) padronização do preparo do conjunto de dados, (iii) comparabilidade entre diferentes variações de arquitetura e estratégias de realce de imagem, e (iv) reprodutibilidade dos experimentos.

De modo geral, o fluxo de trabalho do estudo é composto por: coleta e organização das bases de imagens; anonimização e padronização dos arquivos; anotação/rotulagem supervisionada; segmentação das regiões de interesse (recortes) para construção do *dataset*; pré-processamento e realce de contraste; aumento de dados; definição e treinamento das arquiteturas; validação cruzada estratificada; e avaliação por métricas quantitativas. A Figura 6 ilustra o encadeamento dessas etapas.

Figura 6 – Visão geral do *pipeline* metodológico: aquisição, organização e anonimização das radiografias; rotulagem; segmentação em sub-imagens; realce e normalização; aumento de dados; treinamento/validação; e avaliação.



Fonte: Elaborada pelo autor.

3.1 Bases de imagens e organização inicial

As radiografias PAN utilizadas neste estudo foram obtidas a partir de duas fontes: (i) uma base institucional de radiografias PAN provenientes de uma universidade (UFC) e (ii) um subconjunto público oriundo do *Tufts Dental Database* (TDD). No total, foram consideradas 628 radiografias PAN, sendo 604 imagens da UFC e 24 imagens do TDD.

Antes de qualquer processamento, as imagens foram organizadas em uma estrutura de diretórios padronizada, de modo a facilitar a automação das etapas subsequentes e evitar inconsistências de nomenclatura. Para garantir integridade e rastreabilidade, cada imagem foi associada a um identificador interno (ID) único, e todas as operações realizadas ao longo do *pipeline* foram registradas (por exemplo, em arquivo de log e/ou planilha de controle), incluindo data de processamento, versão do *script* e parâmetros relevantes.

3.1.1 Anonimização e conformidade ética

Por se tratar de imagens clínicas reais, foi adotado um protocolo de anonimização com o objetivo de impedir a identificação direta ou indireta dos pacientes. Quando necessário, informações pessoais presentes no cabeçalho/rodapé da radiografia (nome, número de prontuário, data de nascimento, etc.) foram ocultadas antes da inserção no fluxo experimental. Em seguida, os arquivos foram renomeados de forma padronizada (por exemplo, IMG_000001.png), mantendo-se uma tabela de correspondência sob guarda do pesquisador responsável, quando aplicável. Esse procedimento assegura que o estudo utilize apenas dados desidentificados durante as etapas de treinamento e avaliação, reduzindo riscos relacionados a privacidade e promovendo aderência às boas práticas de pesquisa.

3.1.2 Critérios de inclusão e exclusão

Com o objetivo de reduzir ruídos experimentais e manter um padrão mínimo de qualidade visual, foram priorizadas radiografias com:

- nitidez suficiente para visualização das estruturas anatômicas relevantes;
- ausência de cortes severos que comprometessem a região maxilomandibular;
- contraste global compatível com leitura clínica.

Imagens com artefatos extremos (por exemplo, superexposição generalizada ou perdas severas de informação) foram evitadas, quando identificadas, para não introduzir vieses de treinamento. Ressalta-se que pequenas variações de brilho/contraste são esperadas entre equipamentos e condições de aquisição, razão pela qual o *pipeline* inclui etapas de normalização e realce descritas adiante.

3.2 Rotulagem supervisionada e construção do *dataset* por recortes

A tarefa abordada neste trabalho é de **classificação binária** (cisto vs. não cisto), porém baseada em recortes (sub-imagens) extraídos das radiografias PAN. Essa decisão metodológica foi tomada para (i) aumentar a quantidade de amostras disponíveis para aprendizado supervisionado, (ii) reduzir a complexidade da cena, uma vez que a PAN contém múltiplas estruturas sobrepostas e regiões com variação significativa de contraste, e (iii) direcionar o modelo para padrões locais relevantes associados à presença de lesões.

3.2.1 Ferramenta e procedimento de anotação

A etapa de rotulagem foi conduzida por anotadores treinados, utilizando ferramenta de anotação de imagens (por exemplo, *LabelImg* ou equivalente). Para cada PAN:

1. foram identificadas regiões compatíveis com lesões císticas (quando presentes), delimitando-se uma caixa delimitadora (*bounding box*) envolvendo a região de interesse;
2. foram selecionadas regiões *não císticas* (controle) em áreas sem evidências de lesão, de modo a representar variações anatômicas e texturais normais;
3. as anotações foram exportadas em formato estruturado (por exemplo, XML/JSON), contendo coordenadas do retângulo e rótulo de classe.

A qualidade da anotação é uma variável crítica para o desempenho de modelos supervisionados. Por isso, adotou-se uma etapa de inspeção visual amostral para reduzir erros típicos (caixas deslocadas, classes trocadas, coordenadas inconsistentes), bem como padronizar decisões de contorno (por exemplo, incluir margens mínimas ao redor do cisto para preservar contexto).

3.2.2 Segmentação das radiografias em sub-imagens

A partir das caixas delimitadoras anotadas, foram gerados recortes retangulares correspondentes às regiões de interesse. Cada recorte foi salvo como uma imagem independente e rotulado conforme a classe associada (cística ou não cística). Ao final desse processo, obteve-se um *dataset* segmentado com 2.431 sub-imagens, sendo 1.219 sub-imagens císticas e 1.212 sub-imagens não císticas, caracterizando um conjunto aproximadamente balanceado.

Para a extração dos recortes, foi implementado um *script* automatizado que:

- lê a imagem original e as coordenadas anotadas;

- aplica o recorte, garantindo que as coordenadas estejam dentro do domínio da imagem (tratando limites);
- opcionalmente aplica uma margem (*padding*) ao redor do retângulo (quando desejado), a fim de preservar contexto anatômico e minimizar cortes excessivos;
- grava o recorte em diretório específico da classe, mantendo um ID consistente.

Esse procedimento permite transformar um conjunto relativamente pequeno de radiografias em um conjunto maior de amostras supervisionadas, facilitando o treinamento de redes profundas e o uso de validação cruzada.

3.3 Pré-processamento e realce de imagens

Radiografias PAN podem apresentar grande variabilidade de contraste, ruído e intensidade de cinza, influenciada por fatores como equipamento, parâmetros de aquisição e condições clínicas. Para reduzir o impacto dessa variabilidade e facilitar a aprendizagem, foi definido um *pipeline* de pré-processamento composto por etapas de padronização e realce.

3.3.1 Padronização de tamanho e normalização de intensidades

Após a extração dos recortes, cada sub-imagem foi convertida para o formato esperado pela rede (matriz de intensidades em escala de cinza ou 3 canais, conforme implementação). Em seguida, foi aplicado redimensionamento para uma dimensão fixa de entrada. Neste trabalho, adotou-se **128×128 pixels** como tamanho padrão, de modo a reduzir custo computacional e viabilizar o treinamento com múltiplas configurações.

Além do redimensionamento, as intensidades foram normalizadas para um intervalo controlado (por exemplo, $[0, 1]$), garantindo estabilidade numérica e evitando que variações absolutas de brilho dominassem o gradiente durante o treinamento. Quando necessário, foram aplicadas conversões de tipo (por exemplo, `uint8` \rightarrow `float32`) e padronizações adicionais (por exemplo, centralização por média e desvio padrão do conjunto de treino), sempre calculadas **apenas no treino** para evitar vazamento de informação.

3.3.2 Realce por operação morfológica *Black-hat*

Uma das estratégias de realce avaliadas foi a operação morfológica **Black-hat**. Essa transformação é definida como a diferença entre o fechamento morfológico da imagem e a

imagem original, destacando estruturas escuras menores do que o elemento estruturante sobre um fundo mais claro. Em radiografias, essa operação pode realçar bordas e regiões de baixa intensidade associadas a lesões ou variações ósseas, dependendo do tamanho do elemento estruturante e da escala do padrão anatômico.

O procedimento foi implementado com operações morfológicas clássicas, definindo-se:

- um elemento estruturante (por exemplo, elíptico ou retangular);
- um tamanho do elemento estruturante selecionado empiricamente para enfatizar estruturas relevantes sem introduzir artefatos excessivos;
- aplicação do fechamento e cálculo da diferença (Black-hat).

3.3.3 *Realce por correção Gamma*

Outra estratégia de realce utilizada foi a **correção Gamma**, que consiste em uma transformação não linear da intensidade de pixels com o objetivo de aumentar ou reduzir contraste em regiões específicas do histograma. De forma geral, a transformação pode ser expressa por:

$$I_{\text{out}} = I_{\text{in}}^{\gamma}, \quad (3.1)$$

em que I_{in} é a intensidade normalizada e γ controla a curvatura da transformação. Valores de $\gamma < 1$ tendem a clarear regiões escuras, enquanto valores $\gamma > 1$ tendem a escurecer regiões claras, alterando o contraste percebido.

No contexto deste estudo, a correção Gamma foi avaliada como forma de evidenciar detalhes em regiões de interesse potencialmente pouco contrastadas, mantendo consistência entre amostras provenientes de diferentes fontes.

3.3.4 *Cenários experimentais de realce*

Para analisar o impacto do realce na performance de classificação, foram definidos cenários comparativos, aplicados de forma idêntica em todo o conjunto segmentado:

1. **Baseline**: apenas padronização (redimensionamento e normalização);
2. **Black-hat**: padronização + transformação Black-hat;
3. **Gamma**: padronização + correção Gamma;
4. **Black-hat + Gamma**: combinação das duas abordagens (ordem definida e mantida fixa).

Essa estrutura possibilita comparar o ganho (ou perda) de desempenho associado ao realce, isolando o efeito de cada transformação.

3.4 Aumento de dados (Data Augmentation)

Mesmo com a estratégia de recortes, o *dataset* permanece relativamente limitado para padrões de redes profundas, sobretudo quando se considera a diversidade de condições clínicas e aquisição. Para mitigar sobreajuste e promover invariâncias desejáveis (rotação, pequenas translações, zoom), foi aplicado **aumento de dados** durante o treinamento.

O aumento de dados foi implementado com geradores de imagem (por exemplo, *ImageDataGenerator* do Keras), aplicando transformações estocásticas em tempo de treino. A configuração utilizada inclui, entre outras, rotações leves, deslocamentos horizontais e verticais, cisalhamento (*shear*), zoom e espelhamentos horizontal/vertical, mantendo-se as classes e a distribuição do conjunto. Uma parametrização típica empregada inclui valores como `rotation_range=1`, `width_shift_range=0.08`, `height_shift_range=0.08`, `shear_range=0.3`, `zoom_range=0.08`, `horizontal_flip=True` e `vertical_flip=True`, além de um tamanho de lote (*batch*) definido para o processo de treino.

É importante ressaltar que o aumento de dados foi aplicado **somente no conjunto de treinamento** em cada partição da validação cruzada, preservando a integridade dos conjuntos de validação e teste para estimar generalização.

3.5 Arquiteturas propostas e adaptação da U-Net para classificação

A U-Net é tradicionalmente empregada em tarefas de segmentação semântica, devido ao seu formato *encoder-decoder* e às conexões de atalho (*skip connections*) que preservam informação espacial. Neste trabalho, a U-Net foi adaptada para **classificação**, mantendo-se a lógica de extração hierárquica de características, mas substituindo-se a saída segmentadora por um cabeçalho classificatório.

3.5.1 Adaptação para classificação

Ao invés de produzir um mapa de classes pixel a pixel, a rede foi configurada para produzir uma predição global por recorte. Para isso, após o bloco final do *decoder*, foi aplicado um mecanismo de agregação (por exemplo, *GlobalAveragePooling2D*) para reduzir a

dimensionalidade espacial, seguido por camadas densas e uma camada de saída para classificação binária (por exemplo, *softmax* com 2 neurônios ou *sigmoid* com 1 neurônio, conforme codificação das classes). Essa abordagem permite que a rede utilize a capacidade de modelar padrões locais e contextuais, mas retorne uma decisão única por sub-imagem.

3.5.2 Variações de U-Net e backbone EfficientNetB0

Foram avaliadas variações progressivas da U-Net (denominadas v1, v2 e v3) com o objetivo de investigar o impacto de alterações arquiteturais no desempenho. Em particular, foi explorado o uso de *transfer learning* por meio do backbone **EfficientNetB0** no encoder, aproveitando pesos pré-treinados em grande escala para acelerar convergência e melhorar generalização.

De forma resumida:

- **U-Net v1**: arquitetura *encoder-decoder* padrão, com blocos convolucionais (Conv2D + normalização + ativação) e conexões de atalho diretas, seguida por *pooling* global e camadas densas para classificação.
- **U-Net v2**: variação com ajustes de camadas/níveis de extração do *backbone* e foco em melhor alinhamento dimensional dos tensores do *decoder* com os *skip connections*, preservando informação contextual e compatibilidade espacial.
- **U-Net v3**: variação que incorpora **Attention Gates** nas conexões do *decoder*, com a finalidade de filtrar e ponderar as características vindas do *encoder*, realçando regiões relevantes e reduzindo ruído. Após o *decoder*, aplica *pooling* global e camadas densas antes da classificação.

A inserção de módulos de atenção é particularmente relevante em imagens médicas, pois tende a aumentar a sensibilidade do modelo a regiões informativas (por exemplo, bordas de lesões e áreas de rarefação óssea), reduzindo a influência de regiões anatômicas não relacionadas ao alvo.

3.6 Configuração de treinamento e validação

Esta seção descreve o protocolo experimental adotado para treinamento e validação dos modelos, incluindo a estratégia de particionamento dos dados e o procedimento de validação cruzada.

3.6.1 Divisão de dados e validação cruzada estratificada

Para estimar desempenho de generalização de forma robusta e reduzir variância associada a uma única partição de treino/validação, foi adotada **validação cruzada estratificada** (*Stratified K-Fold*). Inicialmente, definiu-se uma separação com **20%** do conjunto para teste (`test_size=0.2`), preservando proporcionalidade entre classes. Em seguida, sobre o conjunto de treino (80%), aplicou-se **StratifiedKFold com 4 partições** (`n_splits=4`), assegurando distribuição de classes semelhante em cada *fold*.

Essa estratégia permite reportar desempenho médio e dispersão entre *folds*, fornecendo uma visão mais confiável do comportamento do modelo sob variações do conjunto de treino.

3.6.2 Função de perda, otimizador e hiperparâmetros

Os modelos foram treinados sob uma função de perda apropriada à classificação, adotando codificação compatível com a camada de saída. Para as implementações com *softmax* em duas classes, foi utilizada a perda *categorical cross-entropy*. O otimizador adotado foi o **Adam**, com taxa de aprendizado inicial de **0.0001**. O treinamento foi configurado para um máximo de **400 épocas**, com *batch size* definido (por exemplo, 32) e embaralhamento de amostras no treino.

3.6.3 Callbacks e controle de sobreajuste

Para aumentar estabilidade, evitar treinamento desnecessariamente longo e reduzir sobreajuste, foram utilizados *callbacks* durante o treinamento:

- **EarlyStopping**: interrompe o treinamento quando não há melhora na métrica de validação por um número fixo de épocas (*patience*), restaurando os melhores pesos;
- **ReduceLRonPlateau**: reduz a taxa de aprendizado quando a perda de validação estagna, favorecendo refinamento em regiões próximas a mínimos locais;
- **ModelCheckpoint**: salva automaticamente o modelo com melhor desempenho no conjunto de validação (por exemplo, menor *val_loss*).

A parametrização desses *callbacks* (por exemplo, *patience* e fator de redução) foi mantida constante entre os experimentos para permitir comparação justa entre arquiteturas e cenários de realce.

3.6.4 Ambiente computacional e reprodutibilidade

A execução dos experimentos ocorreu em dois ambientes: (a) um notebook Dell G3, com processador Intel Core i5, 16 GB de RAM e GPU NVIDIA GeForce RTX 3050; e (b) o Google Colab Pro, em servidor em nuvem. O uso do Colab Pro foi relevante para reduzir o tempo total de execução e facilitar a repetibilidade no ajuste de hiperparâmetros e na comparação entre arquiteturas.

A implementação foi realizada em Python, utilizando bibliotecas consolidadas para VC e DL (por exemplo, OpenCV para pré-processamento e TensorFlow/Keras para modelagem). Para reprodutibilidade:

- foram definidos *seeds* (quando aplicável) nos módulos de aleatoriedade;
- os parâmetros de pré-processamento e aumento de dados foram fixados e versionados;
- a estrutura de diretórios e nomes de arquivos foi padronizada;
- versões de dependências e configurações de hardware (CPU/GPU, memória) foram registradas em apêndice/ambiente.

3.7 Métricas de avaliação e análise de desempenho

A avaliação dos modelos foi conduzida por métricas complementares, uma vez que, em tarefas médicas, Acurácia isolada pode ser insuficiente para refletir desempenho clínico. Assim, além da acurácia (ACC), foram computadas:

- **Precisão** (*Precision*): proporção de predições positivas corretas entre todas as predições positivas;
- **Revocação/Sensibilidade** (*Recall*): proporção de positivos corretamente identificados entre todos os positivos reais;
- **F1-Score**: média harmônica entre precisão e revocação, equilibrando ambos os critérios;
- **Matriz de confusão**: contabilizando VP (verdadeiros positivos), VN (verdadeiros negativos), FP (falsos positivos) e FN (falsos negativos).

As métricas foram calculadas por *fold* e, posteriormente, agregadas (média e desvio padrão), permitindo comparação entre: (i) arquiteturas (v1, v2, v3), e (ii) cenários de realce (*baseline*, Black-hat, Gamma e combinação). Adicionalmente, foram analisados padrões de erro (FP e FN) para discutir limitações e possíveis causas, como baixa nitidez, variações anatômicas e padrões de intensidade semelhantes entre lesões e estruturas normais.

3.8 Síntese do protocolo experimental

Por fim, o protocolo experimental completo pode ser sintetizado nos seguintes passos:

1. Organizar e anonimizar as radiografias das bases UFC e TDD;
2. Anotar regiões císticas e não císticas e exportar coordenadas;
3. Gerar sub-imagens (recortes) a partir das anotações e balancear classes;
4. Aplicar padronização (128×128) e normalização;
5. Executar cenários de realce (*baseline*, Black-hat, Gamma, Black-hat+Gamma);
6. Aplicar aumento de dados durante o treino;
7. Treinar U-Nets ($v1$, $v2$, $v3$) com e sem *backbone* pré-treinado, conforme configuração;
8. Validar com Stratified K-Fold (4 folds), mantendo conjunto de teste isolado;
9. Computar métricas (ACC, Precisão, Revocação, F1 e matriz de confusão) e comparar resultados.

Esse desenho metodológico assegura que as conclusões do estudo se baseiem em comparações controladas, com avaliação robusta e documentação suficiente para reprodução dos experimentos.

4 RESULTADOS

Neste capítulo são apresentados e discutidos os resultados obtidos a partir do protocolo experimental descrito no Capítulo 3. Os resultados são organizados por configuração experimental, comparando as variações de U-Net avaliadas e discutindo o impacto das estratégias de pré-processamento/realce, com base nas métricas de classificação e na análise por matriz de confusão.

De forma geral, buscou-se responder às seguintes perguntas centrais:

1. Qual variação de U-Net apresenta melhor desempenho na classificação binária (cisto vs. não cisto), considerando o mesmo *pipeline* de treinamento e validação?
2. Em que medida as técnicas de realce (Black-hat e correção Gamma) e o uso de *backbone* pré-treinado influenciam o desempenho de classificação?

4.1 Visão geral dos experimentos

Os resultados aqui apresentados abrangem três grupos de experimentos, correspondentes às variações arquiteturais avaliadas:

- **U-NetV1 e U-NetV1B**: versão base e versão com *backbone* (*B*) pré-treinado;
- **U-NetV2 e U-NetV2B**: versão intermediária e sua variante com *backbone*;
- **U-NetV3 e U-NetV3B**: versão mais avançada e sua variante com *backbone*.

Para cada arquitetura, foram comparados três cenários de entrada:

- **Sem realce**: apenas padronização/normalização descrita na metodologia;
- **BlackHat**: realce morfológico do tipo Black-hat aplicado nas sub-imagens;
- **Gamma**: correção Gamma aplicada nas sub-imagens.

A avaliação foi conduzida em dois níveis complementares:

1. **Acurácia média por *fold*** (validação cruzada estratificada com 4 *fold*s), de forma a medir robustez e variabilidade entre partições;
2. **Métricas finais do melhor modelo por experimento**, incluindo Acurácia, Sensibilidade, Especificidade e F1-Score, além da inspeção por matriz de confusão.

4.2 Resultados com validação cruzada estratificada

A Tabela 4 apresenta a acurácia obtida em cada um dos quatro *fold*s para todas as combinações de arquitetura e técnica de realce, juntamente com a média e o desvio padrão

(DP). Essa tabela é fundamental para avaliar não apenas o desempenho médio, mas também a estabilidade do modelo: valores altos com DP baixo indicam desempenho consistente sob diferentes partições de treino/validação.

Tabela 4 – Acurácia média (%) e desvio padrão para diferentes variações da U-Net e filtros de realce, com validação cruzada estratificada em 4 *folds*.

CNN	Realce	K=1	K=2	K=3	K=4	Média ± DP
U-NetV1	Sem realce	70,89%	73,20%	70,25%	71,77%	71,53% ± 1,11
U-NetV1	BlackHat	71,50%	72,48%	67,44%	73,83%	71,31% ± 2,38
U-NetV1	Gamma	69,13%	72,12%	70,74%	68,46%	70,11% ± 1,42
U-NetV1B	Sem realce	77,65%	80,96%	78,77%	77,92%	78,82% ± 1,30
U-NetV1B	BlackHat	77,11%	78,46%	79,87%	76,41%	77,96% ± 1,33
U-NetV1B	Gamma	76,05%	77,92%	79,55%	78,61%	78,03% ± 1,28
U-NetV2	Sem realce	69,35%	73,25%	69,07%	68,10%	69,94% ± 1,97
U-NetV2	BlackHat	67,19%	72,25%	68,50%	73,00%	70,24% ± 2,45
U-NetV2	Gamma	65,20%	71,97%	58,41%	65,52%	65,28% ± 4,80
U-NetV2B	Sem realce	78,05%	77,78%	80,90%	78,49%	78,80% ± 1,24
U-NetV2B	BlackHat	77,71%	79,94%	79,50%	74,70%	77,96% ± 2,06
U-NetV2B	Gamma	77,95%	78,64%	80,66%	77,19%	78,61% ± 1,29
U-NetV3	Sem realce	71,26%	67,01%	68,04%	69,28%	68,90% ± 1,58
U-NetV3	BlackHat	72,33%	72,25%	72,08%	72,76%	72,35% ± 0,25
U-NetV3	Gamma	64,95%	71,52%	59,66%	72,18%	67,08% ± 5,13
U-NetV3B	Sem realce	79,43%	80,50%	81,68%	82,17%	80,95% ± 1,06
U-NetV3B	BlackHat	82,17%	79,96%	78,75%	76,50%	79,34% ± 2,05
U-NetV3B	Gamma	76,88%	79,46%	79,16%	79,30%	78,70% ± 1,06

Fonte: Elaborada pelo autor.

A partir da Tabela 4, destacam-se três observações principais:

1. **Efeito do *backbone*:** em termos de acurácia média, as variantes com *backbone* (U-NetV1B, U-NetV2B e U-NetV3B) apresentam desempenho superior às versões sem *backbone* em praticamente todos os cenários. Isso sugere que a transferência de aprendizado contribui para uma extração de características mais eficiente, mesmo em um domínio específico como radiografias.
2. **Melhor média global:** a configuração com melhor acurácia média foi a **U-NetV3B sem realce** (80,95% ± 1,06), com desempenho estável e variação reduzida entre *folds*.
3. **Realce como fator não monotônico:** o impacto de Gamma e BlackHat não foi uniforme. Em algumas arquiteturas, houve melhora (por exemplo, U-NetV3 com BlackHat teve média 72,35% com DP 0,25), enquanto em outras houve queda e aumento de instabilidade (por exemplo, U-NetV3 com Gamma apresentou DP 5,13 e menor média).

Em outras palavras, os resultados de acurácia por *fold* já indicam que **arquitetura e pré-processamento** devem ser interpretados de forma conjunta: nem todo realce beneficia toda rede, e o ganho do *backbone* tende a ser mais consistente do que o ganho do realce.

4.3 Resultados das métricas finais por arquitetura

Além da acurácia média em validação cruzada, foi realizada a avaliação do melhor modelo obtido em cada experimento, com cálculo das métricas clássicas para classificação binária: Acurácia, Sensibilidade (Recall da classe cística), Especificidade (Recall da classe não cística) e F1-Score. Esta análise é importante porque, em aplicações médicas, falsos positivos e falsos negativos possuem implicações distintas; assim, duas configurações com acurácia semelhante podem apresentar comportamentos clínicos muito diferentes.

4.3.1 U-NetV1 e U-NetV1B

A Tabela 5 apresenta as métricas finais para U-NetV1 e U-NetV1B em três cenários de realce. Observa-se que a variante com *backbone* (U-NetV1B) supera a U-NetV1 em acurácia e F1-Score em todos os cenários avaliados, reforçando a contribuição do aprendizado transferido.

Tabela 5 – Métricas de desempenho da U-NetV1 com e sem *backbone*, sob diferentes filtros de realce.

CNN	Realce	Acurácia	Sensibilidade	Especificidade	F1 Score
U-NetV1	Sem realce	84,59%	80,56%	88,75%	84,14%
U-NetV1	BlackHat	83,16%	85,42%	80,83%	83,73%
U-NetV1	Gamma	83,36%	81,78%	85,00%	83,29%
U-NetV1B	Sem realce	85,42%	83,80%	87,08%	85,36%
U-NetV1B	BlackHat	85,83%	89,87%	81,66%	86,54%
U-NetV1B	Gamma	85,62%	86,23%	85,00%	85,88%

Fonte: Elaborada pelo autor.

Entre as configurações acima, dois comportamentos merecem destaque:

- **U-NetV1B + BlackHat**: apresentou a maior sensibilidade (89,87%), sugerindo maior tendência a identificar regiões císticas corretamente. Entretanto, essa configuração reduziu a especificidade (81,66%), o que implica aumento de falsos positivos, isto é, maior chance de sinalizar cisto onde não existe.

- **U-NetV1B + Gamma**: apresentou o melhor equilíbrio geral (F1-Score 85,88%), com sensibilidade e especificidade próximas, indicando comportamento mais estável na distinção entre classes.

4.3.2 U-NetV2 e U-NetV2B

A Tabela 6 apresenta as métricas finais para U-NetV2 e U-NetV2B. Diferentemente do observado na família V1, aqui a versão sem realce (U-NetV2) obteve o melhor resultado de acurácia e F1-Score, o que sugere que, para esta arquitetura específica, a informação discriminativa já é extraída adequadamente sem necessidade de realce adicional.

Tabela 6 – Métricas de desempenho da U-NetV2 com e sem *backbone*, sob diferentes filtros de realce.

CNN	Realce	Acurácia	Sensibilidade	Especificidade	F1 Score
U-NetV2	Sem realce	87,67%	90,28%	85,00%	88,14%
U-NetV2	BlackHat	85,21%	87,44%	82,91%	85,71%
U-NetV2	Gamma	79,67%	78,54%	80,83%	79,67%
U-NetV2B	Sem realce	87,47%	88,66%	86,25%	87,77%
U-NetV2B	BlackHat	84,18%	87,04%	81,25%	84,81%
U-NetV2B	Gamma	86,03%	89,47%	82,50%	86,66%

Fonte: Elaborada pelo autor.

O pior desempenho do bloco V2 foi observado na combinação **U-NetV2 + Gamma**, com acurácia e F1-Score de 79,67%, além de sensibilidade inferior às demais configurações. Uma interpretação plausível é que a correção Gamma, para certos parâmetros e para essa arquitetura, altere o contraste de maneira a reduzir a separabilidade entre padrões de textura associados às classes, tornando a aprendizagem menos efetiva. Esse comportamento também é coerente com a instabilidade observada em acurácia por *fold* para configurações com Gamma em algumas arquiteturas.

4.3.3 U-NetV3 e U-NetV3B

A Tabela 7 apresenta as métricas finais para U-NetV3 e U-NetV3B. Neste grupo, a inclusão do *backbone* foi decisiva para elevar todas as métricas, e a combinação **U-NetV3B + Gamma** obteve o melhor desempenho global entre todas as combinações avaliadas, atingindo acurácia de 90,34% e F1-Score de 90,54%.

Tabela 7 – Métricas de desempenho da U-NetV3 com e sem *backbone*, sob diferentes filtros de realce.

CNN	Realce	Acurácia	Sensibilidade	Especificidade	F1 Score
U-NetV3	Sem realce	81,10%	83,40%	78,75%	81,74%
U-NetV3	BlackHat	80,28%	80,16%	80,41%	80,48%
U-NetV3	Gamma	83,16%	81,37%	85,00%	83,05%
U-NetV3B	Sem realce	86,44%	89,47%	83,33%	87,00%
U-NetV3B	BlackHat	85,83%	84,21%	87,50%	85,77%
U-NetV3B	Gamma	90,34%	91,09%	89,58%	90,54%

Fonte: Elaborada pelo autor.

Do ponto de vista de aplicação, esse resultado é relevante porque a combinação vencedora (U-NetV3B + Gamma) não apenas maximiza a acurácia, mas também apresenta **equilíbrio** entre sensibilidade (91,09%) e especificidade (89,58%). Em cenários clínicos, essa característica é desejável, pois reduz simultaneamente a probabilidade de deixar passar um caso positivo (FN) e de gerar alarmes indevidos (FP).

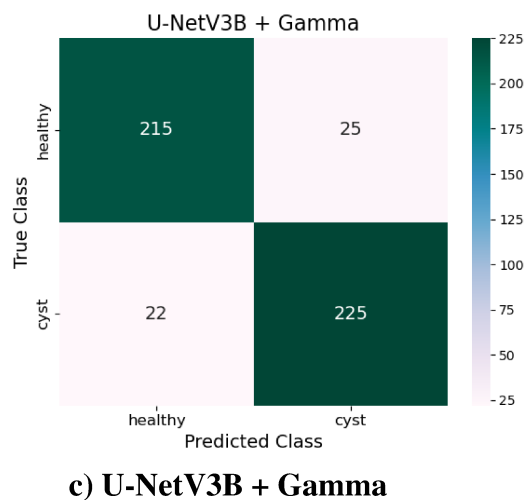
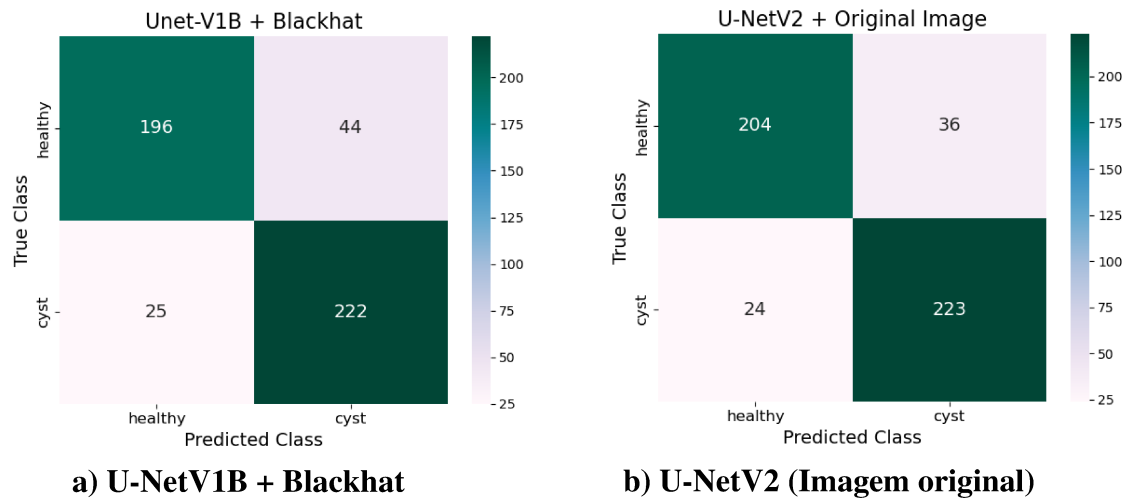
4.4 Matrizes de confusão dos melhores modelos

A análise por matriz de confusão permite observar diretamente a distribuição de erros por classe. A Figura 7 apresenta as matrizes de confusão para os melhores modelos selecionados em cada família de arquitetura, fornecendo uma visão concreta de VP, VN, FP e FN.

A partir das matrizes, observa-se o seguinte:

- **U-NetV1B + BlackHat (Figura 7a):** classificou corretamente 196 imagens não císticas e 222 imagens císticas, mas produziu 44 falsos positivos e 25 falsos negativos. O volume de falsos positivos sugere uma tendência a superestimar a presença de cisto, o que pode ser interpretado como um comportamento mais “sensível”, porém menos específico.
- **U-NetV2 sem realce (Figura 7b):** apresentou melhora em relação ao caso anterior, com 204 verdadeiros negativos e 223 verdadeiros positivos, além de reduzir falsos positivos para 36 e manter falsos negativos em 24. Em termos práticos, esse resultado indica uma redução de alarmes indevidos mantendo uma boa taxa de detecção.
- **U-NetV3B + Gamma (Figura 7c):** obteve o melhor equilíbrio entre erros. Foram 215 verdadeiros negativos e 225 verdadeiros positivos, com apenas 25 falsos positivos e 22 falsos negativos. Essa redução simultânea de FP e FN evidencia maior robustez e sugere que o realce Gamma, quando combinado com uma arquitetura mais forte e *backbone* pré-treinado, favorece uma separação mais clara entre padrões associados às duas classes.

Figura 7 – Matrizes de confusão para os melhores resultados gerados em cada experimento.



Fonte: Elaborada pelo autor.

4.5 Análises comparativas e ablação

Esta seção apresenta análises comparativas entre as configurações avaliadas, com foco em identificar o impacto de escolhas de arquitetura e pré-processamento no desempenho de classificação.

4.5.1 Impacto do backbone pré-treinado

Comparando-se, dentro de cada versão (V1, V2 e V3), as redes com e sem *backbone*, observa-se um padrão: o *backbone* tende a elevar a acurácia e, frequentemente, a estabilizar as métricas (reduzindo flutuação entre sensibilidade e especificidade). Esse efeito é coerente com a hipótese de que pesos pré-treinados oferecem um ponto de partida mais informativo para

filtros convolucionais iniciais, permitindo que o treinamento com *dataset* odontológico refine características já úteis para bordas, texturas e contrastes.

Esse ganho, porém, não é absolutamente uniforme em todos os cenários de realce. Por exemplo, em determinadas combinações, o realce pode reduzir a compatibilidade entre a estatística das imagens de entrada e aquilo que o *backbone* “espera” como distribuição de intensidades/texturas, exigindo maior adaptação do modelo. Ainda assim, no agregado, a presença do *backbone* foi um fator consistente de melhoria.

4.5.2 *Impacto das técnicas de realce (BlackHat e Gamma)*

Os resultados mostram que BlackHat e Gamma não devem ser tratados como melhorias garantidas, mas como transformações cuja eficácia depende da arquitetura e do modo como as características são extraídas internamente.

- **BlackHat:** em alguns casos, aumentou sensibilidade (ex.: U-NetV1B + BlackHat), sugerindo que o realce pode evidenciar padrões associados à classe positiva. Entretanto, a contrapartida observada foi, em certos cenários, a redução de especificidade, indicando que estruturas anatômicas normais podem passar a se assemelhar aos padrões enfatizados pelo filtro.
- **Gamma:** mostrou-se particularmente eficaz quando associado à U-NetV3B, resultando no melhor desempenho global. Isso sugere que a correção Gamma, ao uniformizar contrastes e evidenciar detalhes em determinadas faixas de intensidade, pode facilitar a identificação de contornos e regiões de baixa densidade que compõem padrões compatíveis com lesões císticas. Em contrapartida, Gamma deteriorou o desempenho em algumas configurações sem *backbone* (ex.: U-NetV2 + Gamma), possivelmente por alterar excessivamente a separabilidade dos padrões ou introduzir variações que o modelo não consegue acomodar bem com capacidade limitada.

Portanto, a principal conclusão desta análise é que **o realce deve ser selecionado em conjunto com a arquitetura**. Estratégias de pré-processamento não são universalmente benéficas e podem, inclusive, amplificar ruídos relevantes para a rede.

4.6 Discussão dos resultados

A análise conjunta das métricas e das matrizes de confusão revela uma progressão consistente de desempenho vinculada à complexidade das arquiteturas propostas. O modelo de maior destaque foi a **U-NetV3B com correção Gamma**, que atingiu os melhores índices globais: **90,34% de acurácia, 91,09% de sensibilidade, 89,58% de especificidade e 90,54% de F1-Score**.

O sucesso desta configuração específica pode ser atribuído a uma sinergia técnica entre o pré-processamento e a arquitetura: enquanto a **correção Gamma** atua no realce de detalhes e bordas em regiões de baixa densidade óssea, os **Attention Gates** da U-NetV3 permitem que a rede direcione seu foco justamente para essas características espaciais mais informativas, filtrando ruídos irrelevantes. Esse comportamento demonstra que a combinação de uma arquitetura expressiva com um *backbone* pré-treinado maximiza a capacidade do modelo em diferenciar padrões texturais sutis que caracterizam as lesões císticas.

Quanto à robustez dos experimentos, a validação cruzada indicou que o desempenho médio é estável em diferentes partições dos dados. No entanto, observou-se uma instabilidade significativa (elevado desvio padrão) em cenários como o da U-NetV3 sem *backbone* utilizando Gamma. Essa flutuação sugere que, sem o benefício da transferência de aprendizado, redes complexas com filtros de realce tornam-se excessivamente sensíveis à composição do conjunto de treinamento, o que reforça a importância do *transfer learning* para a generalização em bases de dados médicas limitadas.

Por fim, a redução simultânea de falsos positivos (FP) e falsos negativos (FN) observada nas matrizes de confusão do melhor modelo é um indicador de viabilidade clínica. Em um sistema de apoio ao diagnóstico, o equilíbrio alcançado pela U-NetV3B mitiga tanto a fadiga de triagem, causada por alarmes indevidos, quanto o risco crítico de subdiagnóstico, proporcionando uma ferramenta de segunda opinião mais confiável para o especialista.

4.6.1 Ameaças à validade e limitações do experimento

Os resultados apresentados devem ser interpretados considerando o desenho experimental adotado, baseado na classificação de sub-imagens recortadas a partir de radiografias PAN, bem como a heterogeneidade inerente às condições de aquisição (por exemplo, variações de contraste, ruído, posicionamento do paciente e configurações do equipamento). Além disso,

o uso de técnicas de realce pode introduzir alterações no perfil de intensidades da imagem que beneficiam determinadas arquiteturas, mas não necessariamente generalizam para outros cenários ou bases.

Dessa forma, esta seção limita-se a registrar as principais ameaças à validade interna e externa do experimento, enquanto uma discussão mais ampla sobre limitações, aplicabilidade, implicações do uso em ambiente clínico e possíveis extensões do *pipeline* é apresentada no capítulo de Considerações Finais.

4.7 Síntese dos principais achados

Em síntese, os principais achados deste capítulo são:

1. As variantes com *backbone* (V1B, V2B e V3B) apresentaram desempenho superior, de modo consistente, quando comparadas às respectivas versões sem *backbone*, sugerindo benefício da transferência de aprendizado no contexto avaliado.
2. As técnicas de realce não resultaram em ganhos universais: o efeito observado depende da arquitetura e da presença (ou ausência) de *backbone*, reforçando a necessidade de avaliação conjunta do pré-processamento e do modelo.
3. Entre as configurações testadas, a combinação **U-NetV3B + correção Gamma** obteve os melhores resultados globais e maior equilíbrio entre sensibilidade e especificidade, conforme as métricas reportadas neste capítulo.
4. As matrizes de confusão do melhor modelo indicam redução simultânea de falsos positivos e falsos negativos em relação às demais configurações destacadas, corroborando maior consistência na discriminação entre as classes consideradas.

5 CONSIDERAÇÕES FINAIS

A detecção automatizada de alterações patológicas em radiografias PAN representa um desafio relevante em VC aplicada à saúde, sobretudo devido à variabilidade intrínseca das imagens, à presença de estruturas anatômicas complexas e à dependência de contraste e textura para a diferenciação entre tecidos e regiões de interesse. Neste trabalho, investigou-se a classificação binária de sub-imagens extraídas de radiografias PAN em duas classes (cística e não cística), com o objetivo de analisar como escolhas de pré-processamento e variações arquiteturais impactam o desempenho de modelos baseados em U-Net adaptadas para classificação.

A metodologia adotada permitiu construir um protocolo experimental reproduzível, partindo da composição do *dataset* e segmentação em sub-imagens, passando pela aplicação de técnicas de realce (BlackHat e correção Gamma), e culminando no treinamento e avaliação de diferentes versões da arquitetura U-Net com e sem *backbones* pré-treinados. A validação cruzada estratificada em 4 *folds* e a análise por métricas de desempenho (acurácia, sensibilidade, especificidade e F1-Score), complementadas pela inspeção de matrizes de confusão, viabilizaram uma avaliação robusta do comportamento de generalização dos modelos.

Os resultados obtidos evidenciam, em primeiro lugar, que a evolução arquitetural das variantes de U-Net influencia diretamente a capacidade discriminativa do modelo. As versões mais avançadas, com maior poder de extração e refinamento de características, apresentaram desempenho superior na distinção entre regiões císticas e não císticas. Em segundo lugar, observou-se que o uso de *backbones* pré-treinados é consistentemente vantajoso. A incorporação de pesos oriundos de treinamento em bases extensas fornece um ponto de partida mais informativo para os estágios iniciais da rede, potencializando a aprendizagem mesmo em um domínio visual especializado como o odontológico e reduzindo a dependência de grandes volumes de dados rotulados.

Adicionalmente, as técnicas de realce aplicadas às imagens mostraram impacto significativo, porém não monotônico. Embora filtros como BlackHat e correção Gamma possam aumentar a separabilidade entre classes ao evidenciar contrastes e padrões de textura, os experimentos demonstraram que a eficácia do realce depende da arquitetura e do uso (ou não) de *backbone*. Em outras palavras, o pré-processamento não deve ser tratado como uma etapa universalmente benéfica, mas como um componente que precisa ser calibrado e selecionado conjuntamente com o modelo, sob risco de amplificar ruídos ou reduzir a coerência estatística da distribuição de entrada.

Em termos quantitativos, a configuração que apresentou melhor desempenho global foi a **U-NetV3B com aplicação do filtro Gamma Correction**, alcançando **acurácia de 90,34%**, **sensibilidade de 91,09%**, **especificidade de 89,58%** e **F1-Score de 90,54%**. Esses resultados indicam um equilíbrio desejável entre detecção de casos positivos e controle de falsos alarmes, característica particularmente importante em aplicações de apoio ao diagnóstico, nas quais falsos negativos podem representar perda de oportunidade de investigação clínica e falsos positivos podem gerar sobrecarga de triagem.

Assim, este trabalho reforça que abordagens baseadas em CNNs profundas, quando combinadas a uma estratégia adequada de pré-processamento e ao uso de *transfer learning*, possuem elevado potencial para apoiar a análise de radiografias PAN. Embora não substituam a avaliação especializada, tais modelos podem contribuir como ferramenta complementar, auxiliando na triagem e na priorização de casos, além de estimular o avanço científico ao propor protocolos de avaliação sistemáticos e comparáveis.

5.1 Contribuições do trabalho

Como contribuições principais desta pesquisa, destacam-se:

- a construção de um *pipeline* completo para classificação binária de sub-imagens oriundas de radiografias PAN, incluindo segmentação, realce e treinamento supervisionado;
- a comparação sistemática de diferentes variantes da U-Net adaptadas para classificação, com e sem *backbones* pré-treinados;
- a análise do efeito de técnicas de realce (BlackHat e correção Gamma) sob diferentes configurações arquiteturais;
- a validação por *k-fold* estratificado e avaliação por métricas de classificação, oferecendo uma base sólida para comparação futura com novas abordagens.

5.2 Limitações

Apesar dos resultados promissores, algumas limitações devem ser consideradas:

- **Dependência de recortes/segmentação prévia:** a abordagem assume a disponibilidade de sub-imagens rotuladas (regiões de interesse). Para uso real em ambiente clínico, seria necessário integrar um módulo de detecção/segmentação automática capaz de localizar regiões candidatas em imagens panorâmicas completas.

- **Generalização para diferentes protocolos de aquisição:** radiografias PAN podem variar quanto a equipamento, resolução, contraste e presença de artefatos. Ainda que o realce e a normalização reduzam parte dessas diferenças, a generalização para cenários externos pode exigir estratégias adicionais de adaptação de domínio.
- **Interpretação e explicabilidade:** embora métricas globais indiquem desempenho elevado, aplicações médicas demandam transparência no processo decisório. A ausência de análise de interpretabilidade limita a compreensão sobre quais padrões visuais são decisivos para a classificação.

5.3 Trabalhos Futuros

Como continuidade natural desta pesquisa, algumas direções são particularmente promissoras:

- **Ampliação e diversificação do *dataset*:** aumentar o número de radiografias e incluir dados provenientes de diferentes fontes e equipamentos pode elevar a robustez e a capacidade de generalização dos modelos, reduzindo vieses de aquisição.
- **Sistemas completos de detecção e segmentação:** desenvolver um *pipeline* que opere diretamente em radiografias panorâmicas completas, incorporando detecção e/ou segmentação de regiões suspeitas, tornaria a solução mais aplicável em contexto real. Abordagens multiestágio (detecção → classificação) ou arquiteturas *end-to-end* podem ser investigadas.
- **Pré-processamento avançado e aprendizado de realce:** explorar técnicas mais sofisticadas, como *autoencoders*, modelos de realce aprendidos automaticamente e abordagens baseadas em difusão, pode permitir melhorias no contraste e na separação entre classes sem depender de parâmetros fixos.
- **Arquiteturas modernas e híbridas:** avaliar arquiteturas profundas mais recentes, incluindo modelos híbridos entre CNNs e Transformers, pode ampliar a capacidade de modelar dependências de longo alcance e padrões globais presentes nas radiografias.
- **Interpretabilidade e análise de erro:** incorporar métodos como Grad-CAM e variantes de mapas de atenção para justificar predições e investigar sistematicamente falsos positivos e falsos negativos, contribuindo para maior confiabilidade e aceitação em aplicações clínicas.
- **Validação externa e prototipação:** realizar validação em bases independentes e desenvolver um protótipo de ferramenta de apoio (por exemplo, interface de triagem) permitiria avaliar usabilidade e impacto do sistema em fluxo de trabalho profissional.

Em síntese, os resultados apresentados confirmam o potencial de redes neurais profundas, especialmente quando associadas à *transfer learning* e a estratégias de realce adequadas, para apoiar tarefas de classificação em radiografias PAN. Ao mesmo tempo, reforçam que avanços futuros devem priorizar generalização, explicabilidade e integração em *pipelines* completos de detecção, aspectos essenciais para transição do ambiente experimental para aplicações práticas.

REFERÊNCIAS

- ABDURRAZZAQ, A.; JUNOH, A. K.; MUHAMAD, W. Z. A. W.; YAHYA, Z.; MOHD, I. An overview of multi-filters for eliminating impulse noise for digital images. **TELKOMNIKA (Telecommunication Computing Electronics and Control)**, v. 18, n. 1, p. 385–393, 2020.
- ALVARES, L. C.; EURICO, E.; FREITAS, J. A. d. S. Manual de técnica radiográfica: radiografia intra-orais: periapical, interproximal e oclusal. 1965.
- BALLARD, D. H.; BROWN, C. M. **Computer vision**. [S. l.]: Prentice Hall Professional Technical Reference, 1982.
- BARELLI, F. **Introdução à visão computacional: Uma abordagem prática com Python e OpenCV**. [S. l.]: Editora Casa do Código, 2018.
- BIRDAL, R. G.; GUMUS, E.; SERTBAS, A.; BIRDAL, I. S. Automated lesion detection in panoramic dental radiographs. **Oral Radiology**, v. 32, n. 2, p. 111–118, May 2016. ISSN 1613-9674. Disponível em: <https://doi.org/10.1007/s11282-015-0222-8>.
- BRASIL. Lei nº 13.709, de 14 de agosto de 2018. **Diário Oficial da República Federativa do Brasil**, Brasília, DF, 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/113709.htm. Acesso em: 26 fev. 2024.
- COSTA, F. R.; ESTEVES, C.; BACELAR, M. T. Lesões benignas da mandíbula: uma revisão pictórica. **Acta Radiológica Portuguesa**, v. 28, n. 108, p. 25–36, 2016.
- DEEPAK, G. D.; BHAT, S. K. Optimization of deep neural networks for multiclassification of dental x-rays using transfer learning. **Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization**, Taylor & Francis, v. 12, n. 1, p. 2272976, 2024. Disponível em: <https://doi.org/10.1080/21681163.2023.2272976>.
- FEJERSKOV, O.; URIBE, S.; MARIÑO, R. J. Dentistry in a historical perspective and a likely future of the profession. **Career Paths in Oral Health**, Springer, p. 3–19, 2018.
- FENYO-PEREIRA, M.; CHILVARQUER, I.; HAYEK, J. E.; CHILVARQUER, L. W. **Fundamentos de Odontologia Radiologia Odontológica e Imaginologia, 2º Edição**. São Paulo, SP, Brasil: LIVRARIA SANTOS EDITORA LTDA, GEN - Grupo Editorial Nacional, 2013.
- FERLITSCH, A. **Deep Learning Patterns and Practices**. [S. l.]: Manning Publications, 2021. ISBN 9781638356677.
- FUKUSHIMA, K.; MIYAKE, S. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. In: **Competition and Cooperation in Neural Nets**. [S. l.]: Springer, 1982. p. 267–285.
- GÉRON, A. **Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems**. [S. l.]: O’Reilly Media, 2019.
- GONZALES, R. C.; WINTZ, P. **Digital image processing**. [S. l.]: Addison-Wesley Longman Publishing Co., Inc., 1987.
- GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing**. 3. ed. Upper Saddle River, NJ: Prentice Hall, 2008.

- GONZALEZ, R. C.; WOODS, R. E. **Processamento digital de imagem**. [S. l.]: Pearson, 2010.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. [S. l.]: MIT Press, 2016.
- HA, E.-G.; JEON, K. J.; CHOI, H.; LEE, C.; CHOI, Y. J.; HAN, S.-S. Automatic diagnosis of retention pseudocyst in the maxillary sinus on panoramic radiographs using a convolutional neural network algorithm. **Scientific Reports**, Nature Publishing Group UK London, v. 13, n. 1, p. 2734, 2023.
- JUNIOR, F. V. L. C. O uso de aprendizado profundo na classificação de ressonâncias magnéticas para detecção de tumor cerebral. 2023.
- KAUR, D.; KAUR, Y. Various image segmentation techniques: a review. **International Journal of Computer Science and Mobile Computing**, v. 3, n. 5, p. 809–814, 2014.
- KIM, S.; CASPER, R. Applications of convolution in image processing with matlab. **University of Washington**, p. 1–20, 2013.
- KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. **arXiv preprint arXiv:1412.6980**, 2014. Disponível em: <https://arxiv.org/abs/1412.6980>.
- KUMAR, V. S.; KUMAR, P. R.; YADALAM, P. K.; ANEGUNDI, R. V.; SHRIVASTAVA, D.; ALFURHUD, A. A.; ALMAKTOOM, I. T.; ALFTAUKHAH, S. A. A.; ALSHARARI, A. H. L.; SRIVASTAVA, K. C. Machine learning in the detection of dental cyst, tumor, and abscess lesions. **BMC Oral Health**, Springer, v. 23, n. 1, p. 833, 2023.
- LI, S.; FEVENS, T.; KRZYŻAK, A.; JIN, C.; LI, S. Semi-automatic computer aided lesion detection in dental x-rays using variational level set. **Pattern Recognition**, Elsevier, v. 40, n. 10, p. 2861–2873, 2007.
- LOPES, H. P.; JUNIOR, J. F. S. Endodontia: biologia e técnica. In: **Endodontia: biologia e técnica**. [S. l.: s. n.], 1999. p. 650–650.
- MALLYA, S.; LAM, E. **White and Pharoah's Oral Radiology: Principles and Interpretation**. [S. l.]: Elsevier Health Sciences, 2018.
- NEVES, L. A. P.; NETO, H. V.; GONZAGA, A. Avanços em visão computacional. **Omnipax, Curitiba, PR**, v. 1, 2012.
- NEVILLE, B. W.; DAMM, D. D.; ALLEN, C. M.; CHI, A. C. **Patologia Oral e Maxilofacial**. 3. ed. Rio de Janeiro: Elsevier, 2009. ISBN 9788535230895.
- NURTANIO, I.; PURNAMA, I. K. E.; HARIADI, M.; PURNOMO, M. H. Cyst and tumor lesion segmentation on dental panoramic images using active contour models. **IPTEK: The Journal for Technology and Science**, v. 22, n. 3, p. 152–158, 2011.
- OLIVEIRA, D.; BARRETO, J. B.; MESQUITA, I. M.; JR, I. C. P.; CHAVES, F. N.; SAMPIERI, M. B. S.; MADEIRO, J. P. Analysis of the influence of pre-processing techniques with convolutional neural networks for automatic detection of cysts in wisdom teeth. **Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization**, Taylor & Francis, v. 11, n. 3, p. 299–310, 2023. Substitui a entrada incompleta 'oliveira2022analysis' (sem volume/número/doi). Disponível em: <https://doi.org/10.1080/21681163.2022.2035258>.

PIZER, S. M.; AMBURN, E. P.; AUSTIN, J. D.; CROMARTIE, R.; GESELOWITZ, A.; GREER, T.; ROMENY, B. ter H.; ZIMMERMAN, J. B.; ZUIDERVELD, K. Adaptive histogram equalization and its variations. **Computer Vision, Graphics, and Image Processing**, v. 39, n. 3, p. 355–368, 1987. ISSN 0734-189X. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0734189X8780186X>.

QUEIROZ, J. E. R. de; GOMES, H. M. Introdução ao processamento digital de imagens. **Rita**, v. 13, n. 2, p. 11–42, 2006.

RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: **Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015**. [S. l.]: Springer, 2015. (Lecture Notes in Computer Science, v. 9351), p. 234–241.

RUDER, S. An overview of gradient descent optimization algorithms. **arXiv preprint arXiv:1609.04747**, 2016. Disponível em: <https://arxiv.org/abs/1609.04747>.

SALVADOR, J. F.; MEDEIROS, P. L.; JUNIOR, O. F.; CAPELOZZA, A. L. A. Anatomia radiográfica do canal mandibular e suas variações em radiografias panorâmicas. **Innovations Implant Journal**, v. 5, n. 2, 2010. ISSN 1984-5960. Disponível em: http://revodonto.bvsalud.org/scielo.php?script=sci_arttext&pid=S1984-59602010000200006. Acesso em: 23 fev. 2024.

SANDLER, M.; HOWARD, A.; ZHU, M.; ZHMOGINOV, A.; CHEN, L.-C. Mobilenetv2: Inverted residuals and linear bottlenecks. In: **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. [S. l.: s. n.], 2018. p. 4510–4520.

SANTOS, T. d. S.; ANTUNES, A. A.; AVELAR, R. L.; ANTUNES, A. P. Cistos odontogênicos: estudo epidemiológico de 72 casos. **Rev. Bras. Cir. Cabeça Pescoço**, v. 36, n. 1, p. 30–32, 2007.

SILVA, A. M. M. da; PATROCÍNIO, A. C.; SCHIABEL, H. Processamento e análise de imagens médicas. **REVISTA BRASILEIRA DE FÍSICA MÉDICA (ONLINE)**, 2019.

SIVASUNDARAM, S.; PANDIAN, C. Performance analysis of classification and segmentation of cysts in panoramic dental images using convolutional neural network architecture. **International Journal of Imaging Systems and Technology**, Wiley Online Library, v. 31, n. 4, p. 2214–2225, 2021.

SOKOLOVA, M.; LAPALME, G. A systematic analysis of performance measures for classification tasks. **Information Processing & Management**, Elsevier, v. 45, n. 4, p. 427–437, 2009.

SOUSA, E. L. R. d.; FERRAZ, C. C. R.; GOMES, B. P. F. d. A.; PINHEIRO, E. T.; TEIXEIRA, F. B.; SOUZA-FILHO, F. J. d. Bacteriological study of root canals associated with periapical abscesses. **Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology**, Elsevier, v. 96, n. 3, p. 332–339, 2003.

STEVENS, E.; ANTIGA, L.; VIEHMANN, T. **Deep Learning with PyTorch**. [S. l.]: Manning Publications, 2020.

SUN, M.; SONG, Z.; JIANG, X.; PAN, J.; PANG, Y. Learning pooling for convolutional neural network. **Neurocomputing**, Elsevier, v. 224, p. 96–104, 2017.

- TAN, M.; LE, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In: PMLR. **Proceedings of the 36th International Conference on Machine Learning (ICML)**. [S. l.], 2019. p. 6105–6114.
- VARGAS, A. C. G.; PAES, A.; VASCONCELOS, C. N. Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. In: **Proceedings of the XXIX Conference on Graphics, Patterns and Images (SIBGRAPI)**. [S. l.: s. n.], 2016. p. 1–4.
- VARINAUSKAS, V.; GERVICKAS, A.; KAVOLIUNIENE, O. Analysis of odontogenic cysts of the jaws. **Medicina (Kaunas)**, v. 42, n. 3, p. 201–207, 2006.
- VAZ, G. C. **Aplicação de modelos de deep learning para qualificação da área da engenharia biomédica: um estudo de caso de visão computacional em imagens de raio-x da região torácica**. Tese (Doutorado) – [sn], 2023.
- WHITE, S. C.; PHAROAH, M. J. **Oral Radiology: Principles and Interpretation**. [S. l.]: Elsevier Health Sciences, 2013.
- YOSINSKI, J.; CLUNE, J.; BENGIO, Y.; LIPSON, H. How transferable are features in deep neural networks? **arXiv preprint arXiv:1411.1792**, 2014.