



UNIVERSIDADE FEDERAL DO CEARÁ

CAMPUS SOBRAL

DEPARTAMENTO DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E DE

COMPUTAÇÃO

MESTRADO ACADÊMICO EM ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO

DANILO ALVES OLIVEIRA

**DETECÇÃO AUTOMÁTICA DE CISTOS ODONTOLÓGICOS EM IMAGENS DE
RADIOGRAFIA PANORÂMICA**

SOBRAL

2019

DANILO ALVES OLIVEIRA

DETECÇÃO AUTOMÁTICA DE CISTOS ODONTOLÓGICOS EM IMAGENS DE
RADIOGRAFIA PANORÂMICA

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

Orientador: Prof. Dr. Iális Cavalcante de Paula Júnior

SOBRAL

2019

DANILO ALVES OLIVEIRA

DETECÇÃO AUTOMÁTICA DE CISTOS ODONTOLÓGICOS EM IMAGENS DE
RADIOGRAFIA PANORÂMICA

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

Aprovada em:

BANCA EXAMINADORA

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

Prof. Dr. João Paulo do Vale Madeiro
Universidade Federal do Ceará (UFC)

Prof. Dr. Marcelo Bonifácio da Silva Sampieri
Universidade Federal do Ceará (UFC)

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

À meus queridos avós, Raimundo Alves Neto e Maria das Dores Fernandes, por toda a dedicação prestada em auxílio de meus estudos e pelo caráter ensinado para a vida.

AGRADECIMENTOS

Agradeço a Deus, pela capacidade concebida para concluir mais essa etapa em minha vida acadêmica e pela inspiração da ciência.

À minha querida mãe Antônia de Maria Alves, meus avós e familiares que me auxiliaram durante todos os momentos e me incentivaram a persistir firme nessa jornada.

À meu pai e avô, que apesar de não terem me acompanhado fisicamente, suas presenças foram contantes em meus pensamentos e ações.

À meu orientador Prof. Iális Cavalcante de Paula Júnior, por todo auxílio, oportunidade, confiança, paciência, pelas condições que me proporcionou e dedicação no decorrer deste trabalho.

Aos professores do curso de Odontologia da UFC, Prof. Dr. Filipe Nobre Chaves e Prof. Dr. Marcelo Bonifácio da Silva Sampieri, pela parceria, conhecimentos repassados e pelas valiosas contribuições no desenvolvimento dessa pesquisa.

Aos meus amigos, que sempre se fizeram presentes, buscando compartilhar alegrias, auxiliando nas dificuldades, motivando e ajudando em meu amadurecimento.

Aos membros e colegas da Sala de Estudos do Programa de Pós-Graduação em Engenharia Elétrica e de Computação (PPGEEC), por todo auxílio prestado, conhecimentos compartilhados, risos e amizade adquirida ao longo dessa jornada.

E por fim, a todos que ajudaram de forma direta ou indireta no desenvolvimento dessa pesquisa.

“A maravilhosa disposição e harmonia do universo só pode ter tido origem segundo o plano de um Ser que tudo sabe e tudo pode. Isso fica sendo a minha última e mais elevada descoberta.”

(Isaac Newton, *Philosophiae naturalis principia mathematica*, Livro III)

RESUMO

Os cistos odontogênicos são cavidades patológicas revestidas por epitélio e possuem conteúdo líquido, semi líquido ou gasoso em seu interior. Quando se expandem, causam desordem e pressionam outras estruturas, causando dor ao paciente. Devido a essa capacidade expansiva, é necessário identificá-lo precocemente para prevenir seu avanço ao longo dos seios faciais, causando diversos problemas a saúde do paciente. A melhor forma para identificação desses cistos é o diagnóstico por radiografias dentárias. No entanto, realizar esse tipo de diagnóstico não é tarefa simples, pois a interpretação das imagens de radiografia é um desafio até para especialistas mais experientes. Isso ocorre pela subjetividade da análise dos especialistas, influenciadas pelas limitações do sistema visual humano, fadiga e estresse. Nesse contexto, este trabalho propõe uma metodologia para identificação de regiões císticas em imagens de radiografia panorâmica. Para este fim, é investigada a aplicação de técnicas clássicas de processamento digital de imagens, tais como morfologia matemática, multinível de Otsu, limiarização entre outras. Também são realizadas análises com os descritores de textura *Local Binary Pattern* e Haralick, combinados aos métodos para aprendizado de máquina de Redes Neurais Artificiais (RNA), Máquina de Vetor de Suporte (*Support Vector Machine*) e Máquina de Aprendizado Extremo (*Extreme Learning Machine* (ELM)). Por fim, foi investigada a aplicação da Rede Neural Convolutiva (*Convolutional Neural Network* (CNN)) comparando aos métodos clássicos de visão computacional. Os resultados mostram que a melhor abordagem testada é da CNN com Dice médio de 0,70. Outros resultados concluem que as maiores métricas para a segmentação na região interna da mandíbula foi do multinível de Otsu e da análise de textura foi do descritor de Haralick com ELM e RNA. Por fim, os estudos de segmentação a partir da região patológica resultaram na técnica de contornos ativos com o filtro morfológico como melhor combinação para a segmentação final.

Palavras-chave: Cistos. Segmentação. Classificação. Diagnóstico. Textura. Odontologia. Visão Computacional.

ABSTRACT

Odontogenic cysts are pathological cavities lined by epithelium and have liquid, semi-liquid or gaseous contents inside. When they expand, they cause disorder and pressure other structures, causing pain to the patient. Due to this expansive capacity, it is necessary to identify it early to prevent its advance along the facial sinuses, causing several problems to the patient's health. The best way to identify these cysts is the diagnosis by dental radiographs. However, making this type of diagnosis is not a simple task, as the interpretation of radiographic images is a challenge even for more experienced specialists. This occurs due to the subjectivity of the experts' analysis, influenced by the limitations of the human visual system, fatigue and stress. Therefore, this work proposes a methodology for identifying cystic regions in panoramic radiography images. For this purpose, the application of classical techniques of digital image processing, such as mathematical morphology, Otsu multilevel, thresholding, among others, is investigated. Analyzes are also performed with the texture descriptors *Local Binary Pattern* and Haralick, combined with the methods for machine learning of Artificial Neural Networks (ANN), Support Vector Machine (*Support Vector Machine*) and Machine Extreme Learning Machine (*Extreme Learning Machine* (ELM)). Finally, it was investigated the application of the Convolutional Neural Network (*Convolutional Neural Network* (CNN)) comparing to the classical methods of computer vision. The results show that the best approach tested is from CNN with an average index of 0.70. Other results conclude that the greatest metrics for segmentation in the internal region of the mandible was of the multilayer of Otsu and the texture analysis was of the descriptor of Haralick with ELM and RNA. Finally, segmentation studies from the pathological region resulted in the technique of active contours with the morphological filter as the best combination for the final segmentation.

Keywords: Cysts. Segmentation. Classification. Diagnosis. Texture. Dentistry. Computer vision.

LISTA DE FIGURAS

Figura 1 – Radiografias intrabucais	24
Figura 2 – Radiografia panorâmica.	25
Figura 3 – Cistos em exames de radiografias.	26
Figura 4 – Etapas básicas de um sistema de VC.	29
Figura 5 – Métodos de limiarização.	33
Figura 6 – Processo de um CR.	34
Figura 7 – Exemplo do <i>K-means</i>	36
Figura 8 – Propagação da interface no <i>Fast Marching Method</i>	37
Figura 9 – Pontos vizinhos da grade.	39
Figura 10 – Perceptron simples.	44
Figura 11 – Hiperplanos formados pelo SVM.	46
Figura 12 – Fluxo das metodologias empregadas.	50
Figura 13 – Exemplos dos exames adquiridos.	51
Figura 14 – Imagens do sistema CystMARK.	51
Figura 15 – Exame com 3º molar.	52
Figura 16 – Imagem do GT	53
Figura 17 – Etapa de pré-processamento.	54
Figura 18 – Processo de obtenção da mandíbula.	55
Figura 19 – Algoritmos de segmentação inicial.	57
Figura 20 – Regiões selecionadas pelo multinível de Otsu.	58
Figura 21 – Exemplo de extração dos descritores.	59
Figura 22 – Exemplo de resultado da classificação nas regiões do multinível de Otsu.	61
Figura 23 – Fluxograma do processo utilizado na segmentação final.	62
Figura 24 – Algoritmos de segmentação inicial.	63
Figura 25 – Relação entre GT e segmentação.	65
Figura 26 – Resultados dos algoritmos de segmentação inicial para o grupo de 32 imagens.	68
Figura 27 – Resultados dos algoritmos de segmentação inicial para o grupo de 12 imagens.	69
Figura 28 – Resultados dos algoritmos de segmentação final para o grupo de 32 imagens.	75
Figura 29 – Resultados dos algoritmos de segmentação final para o grupo de 12 imagens.	76
Figura 30 – Imagens geradas pelas metodologias com o grupo de 32 imagens.	81
Figura 31 – Imagens geradas pelas metodologias com o grupo de 12 imagens.	81

LISTA DE TABELAS

Tabela 1 – Tamanhos de <i>kernel</i> da SVM.	60
Tabela 2 – Resultados de classificação com o Haralick para grupo de 32 imagens. . . .	70
Tabela 3 – Resultados de classificação com o LBP para grupo de 32 imagens.	71
Tabela 4 – Resultados de classificação com o Haralick para grupo de 12 imagens. . . .	72
Tabela 5 – Resultados de classificação com o LBP para grupo de 12 imagens.	72
Tabela 6 – Médias de Ac, Se e Es das metodologias.	78
Tabela 7 – Médias obtidas com o coeficiente Dice e outras avaliações.	79

LISTA DE ABREVIATURAS E SIGLAS

Ac	Acurácia
Adam	<i>Adaptive moment estimation</i>
ATM	Articulação Temporomandibular
CA	Contorno Ativo
CAD	<i>Computer-Aided Diagnosis</i>
CNN	<i>Convolutional Neural Network</i>
CR	Crescimento de Região
DCNN	<i>Deep Convolutional Neural Network</i>
Es	Especificidade
FMM	<i>Fast Method Marching</i>
FN	Falso Negativo
FO	<i>First Order</i>
FP	Falso Positivo
GLCM	<i>Grey-Level Co-occurrence Matrix</i>
GLRLM	<i>Gray-Level Run Length Matrix</i>
GT	<i>Ground Truth</i>
IA	Inteligência Artificial
IHME	<i>Institute for Health Metrics and Evaluation</i>
LBP	<i>Local Binary Pattern</i>
MDI	Momento da Diferença Inversa
OMS	Organização Mundial de Saúde
PDI	Processamento Digital de Imagens
RBF	<i>Radial Basis Function</i>
RNA	Redes Neurais Artificiais
ROC	<i>Receiver Operating Characteristic</i>
ROI	<i>Regions of Interest</i>
Se	Sensibilidade
SVM	<i>Support Vector Machine</i>
VC	Visão Computacional
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo

SUMÁRIO

1	INTRODUÇÃO	14
1.1	Trabalhos relacionados	15
1.2	Objetivos	19
<i>1.2.1</i>	<i>Objetivo Geral</i>	<i>19</i>
<i>1.2.2</i>	<i>Objetivos Específicos</i>	<i>19</i>
1.3	Organização do Trabalho	20
2	FUNDAMENTAÇÃO TEÓRICA	21
2.1	Cistos odontogênicos	21
2.2	Imagens Odontológicas	23
<i>2.2.1</i>	<i>Cistos nas imagens de radiografia</i>	<i>25</i>
2.3	Análise de imagens odontológicas	26
2.4	Sistema de Visão Computacional	29
2.5	Segmentação	30
<i>2.5.1</i>	<i>Métodos de limiarização</i>	<i>31</i>
<i>2.5.2</i>	<i>Crescimento de Região</i>	<i>33</i>
<i>2.5.3</i>	<i>Segmentação com K-means</i>	<i>35</i>
<i>2.5.4</i>	<i>Método de Contorno Ativo</i>	<i>35</i>
<i>2.5.5</i>	<i>Fast Marching Method</i>	<i>37</i>
2.6	Operações de Morfologia Matemática	39
2.7	Descritores de Textura	40
<i>2.7.1</i>	<i>Descritores de Haralick</i>	<i>41</i>
<i>2.7.2</i>	<i>Local Binary Pattern</i>	<i>42</i>
2.8	Algoritmos de aprendizado de máquina	43
<i>2.8.1</i>	<i>Redes Neurais Artificiais</i>	<i>44</i>
<i>2.8.2</i>	<i>Support Vector Machine</i>	<i>45</i>
<i>2.8.3</i>	<i>Extreme Learning Machine</i>	<i>46</i>
<i>2.8.4</i>	<i>Redes Neurais Convolucionais</i>	<i>46</i>
2.9	Métricas de avaliação	47
2.10	Considerações Finais	48
3	METODOLOGIA	49

3.1	Formação do banco de imagens	49
3.2	Pré-processamento	53
3.3	Obtenção de regiões da mandíbula	54
3.4	Segmentação das regiões iniciais	56
3.5	Extração de características e classificação	58
3.6	Segmentação final do cisto	61
3.7	Aplicação da CNN	63
3.8	Cálculo das métricas	65
3.9	Considerações Finais	66
4	RESULTADOS	67
4.1	Resultados da segmentação inicial	67
4.2	Resultados de classificação por textura	70
4.3	Resultados de segmentação final	74
4.4	Resultados finais da metodologia	78
4.5	Considerações Finais	82
5	CONCLUSÕES E TRABALHOS FUTUROS	83
	REFERÊNCIAS	85

1 INTRODUÇÃO

A boca é um importante órgão para manutenção da vida, funcionando desde uma simples entrada de alimento para seres primitivos a um órgão de estrutura robusta em organismos mais evoluídos. Nesse último grupo, ela é constituída por um complexo conjunto muscular, ósseo e nervoso. Para o homem, ela oferece diversas funções sociais e biológicas, atuando no corpo como entrada para o sistema digestivo além de auxiliar na respiração (ERIKSEN; DIMITROV, 2003; MCDONALD; MACFARLANE, 2018).

Na aplicação de sua função biológica, ocasionada pela entrada de alimentos, a cavidade oral também é utilizada como ambiente de entrada e morada para vírus e bactérias, sendo um dos habitats mais povoados do corpo humano com cerca de 6 milhões de bactérias (EDLUND *et al.*, 2015). A presença de restos de alimentos e a característica de umidade torna esse órgão um ambiente ideal para a proliferação desses micro-organismos que podem ocasionar diversos tipos de doenças bucais.

As patologias relacionadas à boca estão presentes em grande parte da população mundial e representam um importante desafio para as autoridades públicas. Segundo o estudo *Global Burden of Disease 2017*, feito pelo *Institute for Health Metrics and Evaluation* (IHME), a estimativa de pessoas com incidência de doenças orais foi de 3,59 bilhões em 2017 (IHME, 2018; JAMES *et al.*, 2018), representando elevado crescimento em relação a 2015 com 3,5 bilhões de casos (KASSEBAUM *et al.*, 2017). Esses dados são obtidos em 195 países desde 1990 e servem como referência para a Organização Mundial de Saúde (OMS).

No Brasil, a Pesquisa Nacional de Saúde Bucal, realizada em 2010 pelo Ministério da Saúde, aponta melhorias nas condições de saúde oral. Nessa estimativa, a presença de cárie na população com idade de 12 anos foi de 53% enquanto que, na mesma pesquisa, em 2003 foi de 69%. No entanto, em relação aos problemas periodontais, o percentual de indivíduos que apresentam algum tipo de problema foi elevado, apresentando 37% para a idade de 12 anos, 49,1% para a faixa de 15 a 19 anos, 82,2% para os adultos de 35 a 44 anos e 98,2% nos idosos de 65 a 74 anos (BRASIL, 2012).

Apesar de resultados animadores, como a redução de cáries para algumas faixas etárias, a pesquisa também mostrou que aspectos regionais impactam estas condições (RONCALLI, 2011). Outra condição apresentada por essa, é que houve alto índice de incidência de outros problemas orais, demonstrando que estes ainda são bem presentes na população brasileira.

As doenças bucais também impactam significativamente a economia, sejam com

custos diretos de tratamento ou indiretos, tal como a perda de produtividade no trabalho. Segundo Listl *et al.* (2015), a OMS estima que o impacto econômico total das doenças dentárias no mundo foi de U\$ 442 bilhões em 2010. No tratamento dessas doenças os gastos mundiais estimados são de U\$ 298 bilhões anuais, representando uma média de 4,6% dos gastos globais com saúde.

Apesar de todos os investimentos no tratamento de doenças bucais, a incidência continua alta, sendo a cárie dentária a com maior número de casos (JAMES *et al.*, 2018). Outro tipo de patologia que se destaca são as lesões maxilares, por poderem apresentar características degenerativas e de proliferação com potencial para colocar em risco a vida do paciente (NEVILLE *et al.*, 2009; IMRAN *et al.*, 2016).

Essas lesões que apresentam características clínicas e radiográficas específicas, são ocorrências relevantes na patologia oral e maxilofacial (LOUREDO *et al.*, 2017). Suas incidências provocam alterações nos exames de radiografia, sendo divididas em: tumores odontogênicos, cistos dos maxilares, pseudocistos, patologia óssea e neoplasias malignas, sendo as mais frequentes, os tumores e cistos. Os últimos são mais comuns na clínica odontológica, enquanto os tumores apresentam menor ocorrência. (PONTES *et al.*, 2012).

Na prática clínica, os cistos odontogênicos são patologias encontradas comumente por profissionais radiologistas. Estas podem ser detectadas pesquisando-se em locais que apresentem sintomas indicados pelos pacientes ou em algum exame de forma incidental, podendo representar um desafio ao diagnóstico (COSTA *et al.*, 2016).

Um estudo recente feito por Farias *et al.* (2019), com dados de 2006 a 2017 do estado da Bahia, mostra que a maioria das lesões orais intraósseas são císticas, com 47,5% de incidência. Outra pesquisa feita por Gordón-Núñez *et al.* (2010), no Rio Grande do Norte, obteve que a taxa incidente de cistos odontogênicos foi de 11%, para exames coletados em 38 anos. Esses dados demonstram que os cistos odontogênicos apresentam considerável incidência nas lesões maxilares, sendo a de maior ocorrência dentre as intraósseas.

1.1 Trabalhos relacionados

Tendo em vista a necessidade da detecção precoce de cistos e os desafios oferecidos pelas imagens odontológicas, diversos trabalhos de VC foram desenvolvidos na busca de algoritmos que detectem, segmentem e classifiquem estas patologias (MIKULKA *et al.*, 2012; BANUMATHI *et al.*, 2006). Essas pesquisas buscam técnicas e metodologias que obtenham melhores resultados no auxílio ao diagnóstico, seja para uma etapa do sistema de CAD ou o

processo completo.

Li *et al.* (2007) propuseram um método semi-automático para detectar lesões em imagens de raio-X periapicais. As imagens são segmentadas em três regiões, através de duas funções *level set* competitivas. A entrada dessas é um contorno inicial fornecido pelo classificador *Support Vector Machine* (SVM). Com as regiões segmentadas é criado um mapa de incertezas juntamente com técnica de médias para selecionar as áreas mais radiolúcidas, isolando os dentes e localizando as lesões. O método foi testado em 60 imagens e apresentou 9 Falsos Positivos (FPs) e 181 lesões detectadas.

Nurtanio *et al.* (2011) propuseram o uso de *Snakes* para detecção semi-automática de lesões dentárias causadas por cistos e tumores em radiografias panorâmicas. No pré-processamento é utilizado o filtro gaussiano, na etapa de segmentação é empregado um método baseado no modelo de deformação *snake*, através de modificações nos parâmetros relativos a pesos de energia e função objetivo. O método foi aplicado em 24 imagens através de posições iniciais fornecidas por especialistas e obtiveram acurácia média de 99,67%.

Vijayakumari *et al.* (2012) realizaram um estudo sobre a análise de cistos usando informação de textura. Em sua metodologia são investigadas técnicas de pré-processamento e padrões de textura para diferentes tipos de cistos. Na primeira etapa foram testadas a análise de blocos, extensão de contraste e abertura por reconstrução. A textura utilizada foi obtida pelas estatísticas da GLCM e foram avaliados parâmetros de contraste, correlação, energia, homogeneidade e média. Os testes foram realizados em 3 imagens, cada qual com um tipo diferente de radiografia. O melhor pré-processamento encontrado foi a extensão de contraste e o parâmetro que apresenta maior descrição entre os tipos de cistos é a energia.

Nurtanio *et al.* (2013) desenvolveram um sistema de classificação de cistos dentários em imagens de radiografia panorâmicas. Sua metodologia consiste em utilizar um filtro gaussiano para remover o ruído e realizar a extração de características dos cistos através de marcações feitas por especialistas. As texturas extraídas são baseadas em estatísticas de primeira ordem (FO, *First Order* em inglês), matriz de co-ocorrência de nível de cinza (GLCM, do inglês *Grey-Level Co-occurrence Matrix*) e matrizes de comprimento de corridas de cinza (GLRLM, do inglês *Gray-Level Run Length Matrix*). Estas foram classificadas em um SVM e obtiveram melhores resultados para a combinação de FO e GLRLM com acurácia de 87,18% e área abaixo da curva *Receiver Operating Characteristic* (ROC) de 0,9444.

Alsmadi (2016) desenvolveu um método automático para segmentação de lesões em

imagens de radiografia panorâmica. Um filtro de média é aplicado na imagem de entrada para atenuar o ruído. Após isso, o Fuzzy C-Mean com Neutrosófia é utilizado para agrupar e obter os graus de associação entre os pixels. A integração com a lógica da neutrosófia tem objetivo de reduzir o ruído sem desfoque nas bordas das imagens. A validação foi testada em 95 imagens e os resultados obtidos foram de 94,71% de similaridade, 95,92% de sensibilidade e 94,12% de especificidade, com razão de FPs de 6,1%.

Birdal *et al.* (2016) propuseram uma metodologia automatizada para detecção de lesões periapicais ocasionadas por cistos ou outros fatores, utilizando imagens de radiografia panorâmica. Nessa proposta, as mandíbulas são separadas utilizando a transformada discreta de wavelet com regressão polinomial. Após isso, é feita uma varredura radial angular para segmentar os dentes e localizar os ápices das raízes em que será aplicado o crescimento de região. O método foi testado em 9 imagens e obteve sensibilidade de 70% e especificidade de 89%.

Devi *et al.* (2018) propuseram uma técnica automática híbrida para detectar e extrair regiões císticas usando propriedades simétricas de curvas geodésicas *Isophote* e o *Fast Method Marching* (FMM). Cada curvatura *Isophote* possui um isocentro associado. A segmentação utilizada é iniciada a partir do máximo isocentro das curvas obtidas. Nesse ponto, é aplicado o FMM e obtido o cisto. O método foi testado com 3 imagens de radiografia e obteve acurácia média de 95% e coeficiente de similaridade Dice de 0,89 (SØRENSEN, 1948; DICE, 1945).

Poedjiastoeti e Suebnukarn (2018) desenvolveram uma rede neural convolucional (CNN, do inglês *Convolutional Neural Network*) para detecção automática de ameloblastomas e tumores odontogênicos queratocísticos. O banco de dados conta com 500 imagens de radiografia panorâmica. As imagens são pré-processadas com o logaritmo inverso e equalização de histograma. A arquitetura utilizada é a VGG-16 (ZHANG *et al.*, 2015), que consiste de 16 camadas. Essa rede foi pré-treinada com o banco ImageNet (RUSSAKOVSKY *et al.*, 2015) e refinada treinando-se com o conjunto de 400 exames. A CNN foi aplicada em 100 imagens e obteve acurácia de 83%, especificidade de 83,3% e sensibilidade de 81,8%.

Ariji *et al.* (2019) propõem avaliar o desempenho da detecção de lesões com aprendizado profundo em imagens de radiografia panorâmicas. O modelo foi criado usando-se a rede neural profunda DetectNet, da NVIDIA, com estimação de momento adaptativo (Adam) e treinamento com 500 épocas. São utilizadas 210 imagens no treinamento e os testes foram realizados com dois grupos, o primeiro com 50 imagens do conjunto de treinamento e o segundo com 25 imagens não vistas pela rede. Os resultados apresentaram sensibilidade de 88% para os

dois conjuntos, com nenhum FP no primeiro grupo e 0,04 FP/imagem no segundo grupo.

Yong *et al.* (2019) propuseram um método automático para detectar cinco tipos de cistos em imagens de radiografia panorâmica. Este utiliza uma rede neural convolucional profunda (DCNN, do inglês *Deep Convolutional Neural Network*) obtida através da modificação do YOLO-v3 (REDMON; FARHADI, 2018) para detectar cistos odontogênicos. Este último apresenta pré-treinamento feito utilizando-se os pesos convolucionais do modelo daknet53 (REDMON; FARHADI, 2018). O banco utilizado consta com 1182 exames e a partir desses foram realizadas transformações de rotação, flip horizontal e correção de gama aumentando-se o número de imagens. Os resultados apresentaram precisão média de detecção (mAP) de 94,14%, precisão de 0,99 e recall de 0,93.

Com base na literatura especializada, percebem-se os diversos esforços feitos pela comunidade acadêmica em busca de técnicas computacionais que auxiliem o diagnóstico de cistos odontogênicos. Essas utilizam desde técnicas clássicas de CV a procedimentos mais recentes com aprendizado profundo.

Nestas propostas, a técnica mais comum para redução de ruído é a filtragem das imagens pelo filtro gaussiano. Quanto a segmentação, as metodologias utilizam técnicas variadas, sejam com novas propostas ou com as baseadas em algoritmos clássicos, tais como contornos ativos, c-means, level set, FMM, crescimento de região, entre outros. Também é possível perceber o uso de ferramentas de aprendizado de máquina, tais como SVM, redes neurais artificiais, lógica fuzzy, CNN, dentre outras. Considerando-se os descritores, o mais frequente para classificar padrões de cistos é o de textura relacionado à matriz de GLCM.

Também é possível perceber que diversas propostas não apresentam a localização automática do cisto, sendo esta manualmente marcada para então ser realizada a segmentação. Além disso, a maioria dos trabalhos não apresenta análises comparativas de técnicas nas etapas intermediárias da metodologia, não ficando claro o motivo de se usar determinado algoritmo.

Na literatura pesquisada, não se encontra um método para análise da influência do tamanho do espaço amostral no cálculo da matriz de GLCM e na identificação de cistos. Também não há um método que investigue a textura *Local Binary Pattern* (LBP) e o uso do classificador *Extreme Learning Machine* nestas aplicações.

Considerando a importância clínica e acadêmica que a identificação de cistos engloba, os desafios oferecidos pelas imagens panorâmicas, bem como a necessidade de métodos que colaborem na interpretação de exames e auxiliem dentistas na decisão diagnóstica, a presente

pesquisa propõe uma metodologia para identificação automática de cistos em radiografias panorâmicas. Para isto, é desenvolvido um algoritmo de extração de parâmetros da imagem da mandíbula e análise de técnicas de multinível de Otsu, *K-means* e limiar adaptativo no processo de segmentação inicial. Para detecção de regiões pertencentes aos cistos, são investigadas as texturas de Haralick e LBP como descritores, considerando a influência do tamanho da janela das amostras nos resultados.

Os métodos de aprendizado de máquinas utilizados na classificação são as Redes Neurais Artificiais, *Support Vector Machine* e *Extreme Learning Machine*. Também é proposto um estudo das técnicas de segmentação a partir de regiões internas a cistos, analisando-se os modelos de contornos ativos, *Fast Method Marching* e crescimento de região, verificando-se a influência dos métodos de pré-processamentos e da variação do parâmetro de sensibilidade. Além disso, pretende-se implementar uma CNN para detectar os cistos e comparar os resultados com a metodologia proposta, confrontando uma abordagem clássica de VC com um modelo moderno de aprendizado profundo. Por fim, pretende-se avaliar as investigações realizadas na pesquisa por meio de métricas comumente utilizadas na literatura, como a acurácia, sensibilidade, especificidade e coeficiente Dice.

1.2 Objetivos

1.2.1 *Objetivo Geral*

Desenvolver uma metodologia, através de estudos sobre técnicas de Visão Computacional, para a identificação automática de cistos dentários em imagens de radiografia panorâmica, fornecendo-se a profissionais de Odontologia uma alternativa de auxílio na interpretação dessas imagens.

1.2.2 *Objetivos Específicos*

- Aplicar técnicas de segmentação na região interna da mandíbula, buscando-se separar as regiões com maior possibilidade de cisto;
- Investigar a capacidade de discriminação da região cística através das texturas de Haralick e LBP, com variação do tamanho de janela amostral e diferentes classificadores;
- Aplicar técnicas de segmentação a partir de regiões internas do cisto, avaliando-se

- a influência das técnicas de pré-processamento;
- Analisar o uso de um modelo de aprendizado profundo para identificação de cistos.
- Comparar as técnicas utilizadas na pesquisa aplicando-se métricas estatísticas;

1.3 Organização do Trabalho

A organização deste trabalho está estruturada em 5 capítulos. O capítulo 2 abrange os aspectos teóricos com relação às imagens de radiografia e as técnicas utilizadas na pesquisa, bem como as métricas de avaliação dos resultados. O capítulo 3 apresenta a metodologia proposta e sua implementação. No capítulo 4 são apresentados e discutidos os resultados das investigações. Por fim, o capítulo 5 apresenta as conclusões, considerações, contribuições e trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Neste Capítulo, são apresentadas as descrições das imagens de radiografias odontológicas, com suas características físicas e a forma como os cistos são nelas identificados. Em seguida, são descritas as etapas de um sistema típico de VC, apresentando-se as técnicas utilizadas. Também são abordados os métodos relacionados à segmentação e à morfologia matemática, com suas operações utilizadas. Então, são apresentados os descritores de textura, bem como os algoritmos de aprendizado de máquina. Por fim, são relatadas as métricas de avaliação aplicadas.

2.1 Cistos odontogênicos

Os cistos odontogênicos são patologias que geralmente apresentam características assintomáticas, com exceção dos cistos infectados, que são sintomáticos. Apresentam expansão cortical e erosão, com crescimento lento, tornando-os demasiadamente largos (SANTOS *et al.*, 2007). Esse efeito causa desordens na região mandibular, pois o cisto infla reabsorvendo a região óssea e pressionando outras estruturas anatômicas, provocando dor ao paciente (BANUMATHI *et al.*, 2006).

Os cistos odontogênicos são definidos como cavidades patológicas revestidas por epitélio odontogênico e presentes nos ossos gnáticos. A sua relação anatômica com a estrutura dentária é relevante para a detecção, pois realça no exame uma aparência radiográfica lucente, com limites definidos, uni ou multiloculares (COSTA *et al.*, 2016). São classificados em cistos odontogênicos de desenvolvimento ou inflamatórios, de acordo com sua origem (NEVILLE *et al.*, 2009; SU *et al.*, 2019; MARTIN; SPEIGHT, 2017).

O primeiro grupo apresenta fatores de origem desconhecidos, são associados como resultado de reação inflamatória providas de infecções e traumas maxilares. A resposta inflamatória possibilita a proliferação de remanescentes epiteliais, obtidos da inclusão de epitélio odontogênico de revestimento nas linhas de fusão dos processos embrionários, provenientes do ectoderma. Dessa forma, são associados a tecidos epiteliais que participam na odontogênese. São divididos em: cisto gengival, ceratocisto, dentígero, gengival do recém nascido, periodontal, glandular e calcificante. (SANTOS *et al.*, 2007; GHOM; GHOM, 2014; NEVILLE *et al.*, 2009).

O segundo grupo apresenta como fator precipitante inflamações que ocasionam a futura região cística. Sua origem são restos epiteliais de Malassez, epitélio crevicular e revestimento sinusal ou de trajetos fistulosos. As respostas inflamatórias nessas regiões aumentam

a produção de ceratinócitos, que resultam no aumento de proliferação do epitélio. São divididos em: cistos radiculares e colateral (GHOM; GHOM, 2014; NEVILLE *et al.*, 2009).

Cistos odontogênicos de desenvolvimento são compostos de cápsula e epitélio, apresentam preenchimento líquido ou mole, semelhantes a sacos e bolsas. A parede do cisto é semelhante a uma membrana semipermeável, permeável na direção do líquido e não permeável na direção externa sólida, característica essa que contribui para que o crescimento seja expansivo. O tipo mais comum desse grupo é o dentígero, que representa aproximadamente 20% dos cistos revestidos por epitélio nos ossos gnáticos (MIKULKA *et al.*, 2012; NEVILLE *et al.*, 2009).

Os cistos odontogênicos inflamatórios são patologias provenientes da proliferação de epitélio devido à inflamação. Bactérias ou um traumatismo que gera modificações pulpares podem infeccionar os canais reticulares e ocasionar essas lesões. Isso ocorre principalmente quando o tecido tenta erradicar a infecção dos agentes externos e causar reabsorção, levando a perda óssea. O tipo mais comum desse grupo é o cisto radicular, que possui a origem de sua inflamação na periodontite apical (MARTIN; SPEIGHT, 2017; GHOM; GHOM, 2014).

Esses tipos de cistos apresentam-se como uma cavidade revestida de forma parcial ou completa por epitélio proliferativo, contendo material semi-sólido ou líquido em seu interior. São geralmente assintomáticos, indicando sintomas expressivos quando atingem grandes proporções (BILODEAU; COLLINS, 2017; NEVILLE *et al.*, 2009).

Dentre os tipos de cistos, Grossmann *et al.* (2007) obtiveram que o de maior incidência na população brasileira é o radicular, com 61%, seguido pelo dentígero, com 25,3% e o queratocisto, com 7,2%. Esses dados foram obtidos com exames coletados no período de 51 anos no estado de Minas Gerais.

Segundo Varinauskas *et al.* (2006), os sintomas mais observados de cistos odontogênicos são a tumefacção, descarga purulenta através de fístula, parestesia, mobilidade dentária, dor e assimetria facial.

A detecção precoce dessas patologias é de fundamental importância, seja para evitar esses sintomas ou que os cistos se espalhem ao longo dos seios faciais. No caso de alguns cistos que apresentam caráter infeccioso, essa expansão pode propagar a infecção à corrente sanguínea causando risco à vida do paciente (SOUSA *et al.*, 2003). Por estes motivos, a detecção em estágios iniciais é essencial para o tratamento e estado pós-diagnóstico do paciente. Segundo Nurtanio *et al.* (2011), a identificação precoce da lesão cística reduz o gasto do tratamento e simplifica a prática clínica.

A melhor forma de identificação precoce desses problemas dentários é através de exames de radiografia, uma vez que, devido a localização dos cistos ser em tecidos que unem-se a ossos e dentes, estes ficam ocultos abaixo da superfície do prato cortical, sendo invisíveis durante as práticas de exames visuais (LI *et al.*, 2007).

Essa avaliação imagiológica é uma tarefa onerosa, pois demanda especialidades específicas, tais como percepção de padrões imagiológicos, sinais secundários e comportamento de prevalência das patologias buscadas. Esses exames podem apoiar na tomada de decisões e prevenir procedimentos desnecessários, provendo um auxílio significativo ao diagnóstico (COSTA *et al.*, 2016).

2.2 Imagens Odontológicas

A grande precursora das imagens médicas foi a radiografia, pois permitiu representar estruturas internas do corpo humano em duas dimensões. A descoberta dos raios-X ocorreu em 1885 pelo físico alemão Wilhelm Conrad Röntgen, através de estudos com raios catódicos e o tubo de Hittorf-Crookes. Ele aplicou essa nova forma de energia através de partes do corpo humano e capturou a saída em uma placa fotográfica, obtendo assim a primeira radiografia. (AMBIKA *et al.*, 2012; FEJERSKOV *et al.*, 2018; MARTINEZ, 2013).

Os raios-X são gerados utilizando-se um tubo a vácuo com um par cátodo e anodo. O anodo é carregado com carga positiva e o cátodo é aquecido, ocasionando a liberação dos elétrons, que se movimentam na direção do anodo. A colisão dos elétrons com o núcleo do anodo gera uma energia conhecida como raio-X. Essa é controlada pela tensão aplicada nos eletrodos e a corrente nos fios do cátodo. A imagem de radiografia é formada ao colocar-se um elemento entre a fonte de raio-X e um filme sensível a esta radiação (GONZALEZ; WOODS, 2009; WHAITES; DRAGE, 2013).

A descoberta de Röntgen possibilitou novas aplicações na área médica e acadêmica e suas contribuições com raios-X foram condecoradas com o Prêmio Nobel de Física em 1901. Os profissionais de odontologia reconheceram rapidamente as vantagens oferecidas pela radiografia e a incorporaram em sua área. A primeira imagem de radiografia dos dentes foi concebida 14 dias depois da descoberta dos raios-X, pelo Dr Frederic Otto Walkhoff, professor da Universidade de Braunschweig, que realizou uma intra-oral em sua própria boca (AMBIKA *et al.*, 2012; STELT, 2005).

A radiografia de filmes dominou o meio clínico por muitos anos e, mesmo com

diversos avanços tecnológicos, ela continua sendo comumente aplicada na radiografia oral e maxilofacial. A evolução natural dessas imagens foi a radiografia digital, que eliminou o processo químico de filme e favoreceu melhor visualização das estruturas. Uma grande vantagem dessa radiografia é a menor dose de radiação que o paciente é exposto. (LIMA, 2017; IANNUCCI; HOWERTON, 2016).

A imagem de radiografia digital utiliza sensores de raios-X, em substituição do filme fotográfico, sendo obtidas fazendo-se com que os raios-X que atravessam um paciente atinjam um dispositivo fluorescente, que converte a radiação em luz. Esse sinal de luz é captado por um sistema de digitalização fotossensível, que os converte em sinais elétricos digitais (GONZALEZ; WOODS, 2009; WHITE; PHAROAH, 2014).

As imagens de raio-X que representam a estrutura dentária em duas dimensões são conhecidas como radiografias planares, sendo divididas em intrabucal e extrabucal. A Figura 1 apresenta as intrabucais. Estas são formadas pelo posicionamento do equipamento na parte interna da boca. Seus tipos mais comuns são: periapical, interproximal e oclusal.

Figura 1 – Radiografias intrabucais



(a) Interproximal



(b) Oclusal



(c) Periapical

Fonte: Imagens 1a e 1b são adaptadas de White e Pharoah (2014). Imagem 1c foi obtida em parceria com o curso de Odontologia - UFC / Campus Sobral

As radiografias periapicais exibem todo o dente e osso alveolar circunjacente. As radiografias interproximais (*bite-wings*) mostram as coroas dos dentes e a crista alveolar adjacente. As radiografias oclusais exibem os dentes e osso alveolar com uma área maior que nas radiografias periapicais (WHITE; PHAROAH, 2014; WHAITES; DRAGE, 2013). Existem diversos tipos de radiografias extrabucais, entretanto a mais utilizada é a panorâmica. A Figura 2 apresenta um exemplo desse tipo.

Figura 2 – Radiografia panorâmica.



Fonte: Imagem obtida em parceria com o curso de Odontologia - UFC / Campus Sobral.

As radiografias panorâmicas são utilizadas clinicamente para o diagnóstico de problemas que precisam de ampla visualização dos maxilares, tais como avaliação de traumatismo, localização de terceiros molares, patologias extensas, lesões suspeitadas e anomalias de desenvolvimento (IANNUCCI; HOWERTON, 2016).

2.2.1 Cistos nas imagens de radiografia

A maioria dos cistos são originados por um conjunto grande de restos de epitélios odontogênicos remanescentes da formação dos dentes, ocorrendo frequentemente nos ossos maxilares. A formação dos cistos é dentro do osso e podem ocorrer em qualquer lugar do maxilar. O local mais comum de formação é na região dentada, acima do canal do nervo alveolar inferior, e são mais raros na região do côndilo e processo coronóide (GHOM; GHOM, 2014; WHITE; PHAROAH, 2014).

Devido a suas características de formação mandibular, o exame mais comum para sua detecção é a radiografia panorâmica, que apresenta de forma completa essa região. Além disso, através dessas imagens, é possível identificar a formação de cistos em estágios iniciais, antes que estes causem mais transtornos ao paciente.

Nas radiografias, os cistos apresentam-se como lesões radiolúcidas, cuja forma específica depende de seu tipo. Geralmente, são redondos ou ovais, e preenchidos por fluido. Quando um cisto apresenta longa duração, pode apresentar calcificação distrófica, o que proporciona à imagem interna um aspecto esparso (WHITE; PHAROAH, 2014).

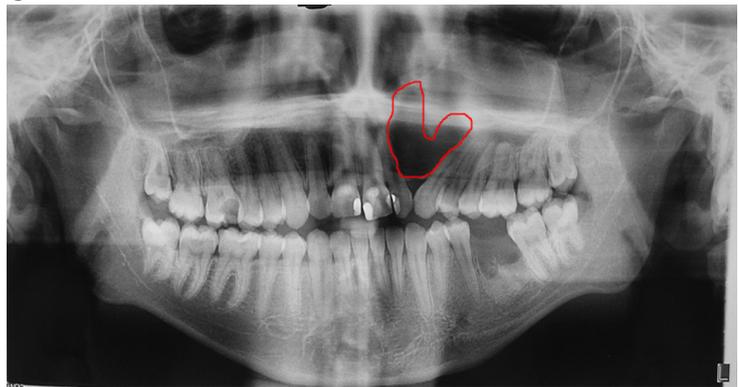
Uma das características dessas patologias é o crescimento lento, que pode causar o deslocamento das estruturas e reabsorção de dentes. Essa região de reabsorção apresenta forma curva e acentuada. Quando se expandem, os cistos podem provocar inchaço da mandíbula, exibindo no exame uma forma curva e suave, transformando as corticais vestibular ou lingual em um fino limite cortical (WHITE; PHAROAH, 2014).

A Figura 3 apresenta exames de radiografias com regiões de cistos demarcados em vermelho.

Figura 3 – Cistos em exames de radiografias.



(a) Cisto em exame periapical



(b) Cisto em exame panorâmico

Fonte: Imagens obtidas em parceria com o curso de Odontologia - UFC / Campus Sobral.

Nessas imagens são apresentados dois cistos periodontais apicais em exames de radiografia, nas quais são observadas as características radiográficas comentadas. Na Figura 3a é possível observar a tonalidade radiolúcida e localização na região próxima à raiz do dente, característico desse tipo de cisto. Na Figura 3b, é possível observar o aspecto expansivo da lesão caracterizado pelo deslocamento dos dentes na maxila.

2.3 Análise de imagens odontológicas

As imagens de radiografia dentária representam um grande avanço para o diagnóstico de patologias bucais. Além da identificação, são necessárias também ao tratamento, como por exemplo, no planejamento de intervenções cirúrgicas ou no acompanhamento do progresso

destas. Essas imagens apresentam detalhes de estruturas externas e internas da região bucal, como canais radiculares dos dentes e região da gengiva (WHITE; PHAROAH, 2014; IANNUCCI; HOWERTON, 2016).

Dessa forma, a detecção de cistos por meio destas é imprescindível e seu uso é frequente em odontologia, principalmente para avaliar lesões bucais que afetam o tecido ósseo da maxila e mandíbula. Apesar desses exames apresentarem informações importantes para diagnosticar anomalias císticas, eles devem ser complementados com outros dados, tais como anamnese e exame histopatológico. Em alguns casos, somente a radiografia é determinística na conclusão do diagnóstico, tais como na detecção de corpos estranhos, dentes retidos, fraturas radiculares, entre outros.

As evoluções tecnológicas das imagens de radiografia culminam na representação digital. Esse novo tipo de imagem elimina o processamento químico utilizado nas imagens com filme convencional e requer menor radiação, reduzindo a dose absorvida pelo paciente (WHITE; PHAROAH, 2014). As vantagens das imagens digitais são significativas, pois apresentam disponibilidade imediata, possibilidade de aprimoramento, reconstrução e análise assistida por computador (LI *et al.*, 2007).

Apesar do progresso que as imagens de radiografia ofereceram para a área odontológica, sejam analógicas ou digitais, sua análise não constitui uma tarefa simples, pois essas precisam ser inspecionadas por um profissional com treinamento especializado. (BIRDAL *et al.*, 2016; LI *et al.*, 2007). Mesmo que a participação dessas imagens em meio clínico seja comum, sua interpretação é um desafio até para especialistas experientes. Um dos principais desafios está no fator dessas representar a anatomia tridimensional das estruturas em duas dimensões (WHITE; PHAROAH, 2014; BIRDAL *et al.*, 2016). Além disso, a inspeção feita pelo profissional apresenta caráter subjetivo, já que não utiliza medidas quantitativas, estando sujeita à limitação do sistema visual humano. Como por exemplo, a capacidade de percepção da variação de tons muito próximos na escala de cinza.

Outros fatores que podem interferir na interpretação são a fadiga, o estresse e distração. Os resultados podem ser mal interpretados resultando em uma avaliação errônea e um diagnóstico incorreto do especialista, que pode afetar diretamente a qualidade de vida do paciente (BIRDAL *et al.*, 2016; LI *et al.*, 2007). Esses aspectos podem ser mitigados através de ferramentas computacionais avançadas que contribuam com a examinação e identificação de patologias.

Programas que ajudam profissionais de saúde no diagnóstico são conhecidos como *Computer-Aided Diagnosis* (CAD) e são aplicados em meio clínico através de imagens médicas. No campo odontológico, esses sistemas vêm apoiando dentistas em diversas aplicações (OHASHI *et al.*, 2016; KATSUMATA; FUJITA, 2014). Alguns destes não se limitam apenas a sistemas para melhorar a qualidade da imagem, mas também sistemas inteligentes capazes de fornecer informações extras. Como por exemplo, identificação das estruturas anatômicas, desordens dentárias e possíveis padrões patológicos.

Esses sistemas CAD inteligentes evitam a arbitrariedade na análise, já que permitem a identificação automática de patologias e estruturas dentárias. Dessa forma, evitam a subjetividade superando as limitações visuais humanas e fornecendo métricas quantitativas sobre um determinado resultado de exame (MAHMOUD *et al.*, 2016). Para este fim, esses sistemas comumente utilizam técnicas de Processamento Digital de Imagens (PDI) e Inteligência Artificial (IA). A união dessas áreas do conhecimento é conhecida como Visão Computacional (VC).

Os sistemas que utilizam VC apresentam etapas de reconhecimento que, além de processar a imagem, fornecem resultados interpretativos. É possível rotular objetos e agrupá-los de acordo com suas características. Esses sistemas utilizam técnicas estatísticas, algoritmos de IA e reconhecimento de padrões (GONZALEZ; WOODS, 2009).

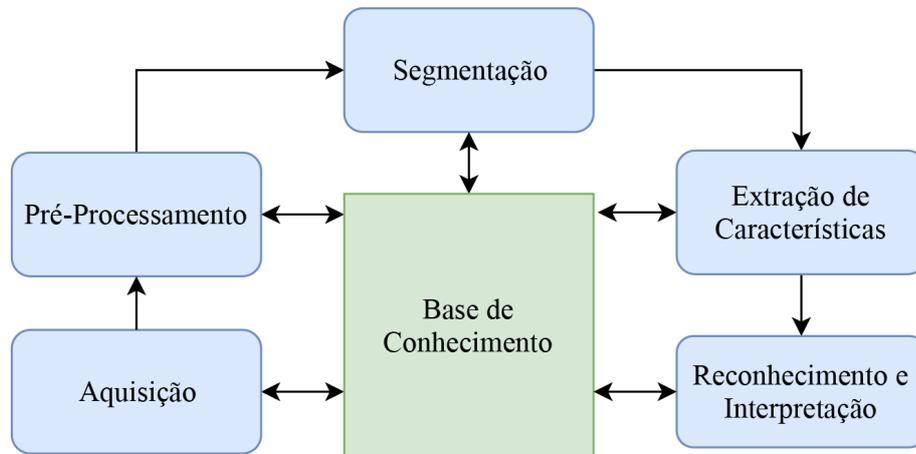
A aplicação de VC em imagens de radiografia dentária apresenta diversas vantagens, tais como a manipulação digital do exame, alterações de contraste, realce de estruturas anatômicas e identificação automática de padrões patológicos. Esses sistemas tornam possíveis novas perspectivas na área odontológica, fornecendo ferramentas que melhoram a compreensão dos exames (LIMA, 2017; STELT, 2005).

Porém, utilizar sistemas de VC em imagens de radiografia não constitui uma tarefa simples. Segundo Li *et al.* (2007), a análise das radiografias dentárias, em comparação com outros tipos de imagens, é um desafio para o processamento clássico de imagens. Principalmente devido a fatores como: má qualidade de aquisição, presença de ruído, artefatos de amostragem e baixo contraste; topologia complexa; orientação arbitrária dos dentes; e ausência clara de demarcação entre as regiões de interesse e demais estruturas. Esse último fator é significativo na detecção de cistos, dado que as estruturas problemáticas tendem a apresentar desordens complicadas e são normalmente acopladas com estruturas saudáveis.

2.4 Sistema de Visão Computacional

A área da VC corresponde ao campo de estudo que objetiva simular nas máquinas a capacidade da visão. Para isso, fazem-se necessários captar cenas como imagens, melhorá-las, separar objetos de interesse e interpretá-las. Normalmente apresentam como entrada do processo uma imagem e fornecem na saída atributos ou interpretações (GONZALEZ; WOODS, 2009; BACKES; JUNIOR, 2016). A Figura 4 apresenta o diagrama de um típico sistema de VC.

Figura 4 – Etapas básicas de um sistema de VC.



Fonte: Adaptado Gonzalez e Woods (2009).

De forma geral, esses sistemas possuem cinco etapas básicas, todas elas aliadas a uma base de conhecimento do problema abordado, conforme descrito a seguir:

- *Aquisição de imagens*: Este é o primeiro processo do fluxo apresentado na Figura 4. Nesta etapa, a cena de entrada é captada através de algum dispositivo específico. Na aplicação odontológica, essas imagens são obtidas por máquinas de raio-X e apresentam diversos fatores que influenciam sua qualidade, tais como posicionamentos, intensidade e o ruído natural ocasionado pela radiação.
- *Pré-processamento*: Etapa responsável por manipular a imagem de forma que seja mais apropriada para as fases subsequentes. Envolve operações como: remover ruídos, realizar restaurações, realce, suavizar ou salientar bordas. Nas radiografias odontogênicas, essa etapa é fundamental para redução do ruído aderido na fase anterior e para obter melhor discriminação das estruturas anatômicas.
- *Segmentação*: Busca dividir os elementos da imagem em regiões, agrupando-as ou separando-as de acordo com propriedades pré-definidas. Essa é uma das etapas mais importantes para o sistema, pois uma falha nesse processo pode

comprometer o resultado final. Nas imagens odontológicas, essa etapa apresenta alto grau de dificuldade, pois os elementos não possuem limites bem definidos entre si, além de sobreposições das estruturas mandibulares.

- *Extração de características:* Utiliza os objetos selecionados na etapa anterior e extrai características que melhor os representem para o problema abordado. São métodos matemáticos que simplificam um conjunto grande de dados para representa-lo com maior precisão. Quanto melhor o atributo extraído descrever o objeto de interesse, maiores serão as chances de sucesso na classificação. Na literatura pesquisada, a característica mais utilizada na representação das regiões cística são as texturas.
- *Reconhecimento e interpretação:* Fase que emprega os atributos dos objetos, extraídos no processo anterior, para classificar e agrupa os elementos de acordo com suas características, fornecendo um rótulo a cada objeto. Os algoritmos que participam dessa etapa fazem a seleção de classes de acordo com uma base de conhecimento anterior. Utiliza técnicas estatísticas, reconhecimento de padrões e IA.
- *Base de conhecimento:* Embora não seja uma etapa de um sistema de VC, esta representa todo o conhecimento prévio sobre o domínio de odontologia, com relação ao comportamento das patologias, características, estruturas anatômicas, entre outras. Essa representação se relaciona com as demais, demonstrando a interação dos conhecimentos do problema com as técnicas de VC.

2.5 Segmentação

A segmentação subdivide a imagem em seus objetos ou regiões com características semelhantes, tais como: níveis de cinza, textura, contraste ou cor. O processo de segmentação está completo quando os objetos de interesse de uma aplicação forem detectados. Dessa forma, o problema a ser tratado é quem define se essa divisão pode apresentar maior ou menor detalhe. Essa é uma etapa essencial dos sistemas de VC, pois sua precisão influencia diretamente no resultado final do sistema (GONZALEZ; WOODS, 2009; KANG *et al.*, 2009).

O principal objetivo desse processo é categorizar as regiões similares da imagem e atribuir um rótulo a cada pixel, de forma a indicar a qual grupo este pertence. As técnicas utilizadas e a quantidade de regiões obtidas dependem do problema, e o conhecimento a respeito

da cena trabalhada é fundamental (ZAITOUN; AQEL, 2015).

Existem diversas técnicas de segmentação, mas de forma geral, elas podem ser classificadas de acordo com as propriedades nas quais se baseiam, sendo agrupadas em duas categorias: descontinuidade e similaridade (GONZALEZ; WOODS, 2009).

Os métodos de segmentação que utilizam propriedades baseadas em descontinuidade detectam as variações abruptas de intensidade de níveis entre um pixel e seus vizinhos, sendo possível encontrar em tais regiões as bordas dos objetos. Essas técnicas utilizam comumente uma convolução bidimensional entre uma matriz específica e a imagem. Exemplos desses métodos são os operadores de Sobel, Prewitt e Roberts (GONZALEZ; WOODS, 2009).

Os métodos de segmentação por similaridade separam os objetos pela proximidade destes em relação a uma característica ou propriedade. Neste processo, as regiões podem ser conectadas ou isoladas. Métodos desse tipo são a limiarização e crescimento de região (CR) (GONZALEZ; WOODS, 2009).

2.5.1 Métodos de limiarização

A limiarização é a técnica de segmentação mais simples e comum, cuja eficiência computacional permite sua extensa aplicação a diversos sistemas de VC. Seu objetivo é agrupar conjuntos de pixels que ocupem a mesma faixa de intensidade de níveis de cinza, partindo do princípio de que os pixels de um objeto concentram-se em uma dada faixa de níveis de cinza a qual difere de regiões de fundo da imagem (GONZALEZ; WOODS, 2009).

Dessa forma, é possível obter um limiar de nível de cinza L que separe as regiões de interesse do objeto com as do fundo da imagem. A técnica de limiarização fornece uma intensidade específica para cada faixa de níveis de cinza, separada a partir de L . Quando separa em duas regiões distintas, essa função é chamada de binarização e pode ser definida como

$$g(x,y) = \begin{cases} 0, & \text{se } f(x,y) < L \\ 1, & \text{se } f(x,y) \geq L \end{cases}, \quad (2.1)$$

em que $f(x,y)$ é o nível de cinza presente no ponto de localização (x,y) (GONZALEZ; WOODS, 2009).

A limiarização pode ser considerada global ou local. A primeira apresenta o mesmo valor de limiar para toda a imagem, enquanto a segunda possui valores de limiar específicos para cada pixel na imagem. Essas definições dada para L são representadas como

$$L = I[x,y,p(x,y),f(x,y)], \quad (2.2)$$

em que $p(x,y)$ corresponde às propriedades consideradas da vizinhança de determinado ponto (x,y) . Caso l seja definido somente em termos de $f(x,y)$, então é um limiar global. Se l depender de $p(x,y)$ além de $f(x,y)$, então é considerado um limiar local.

Essa técnica pode apresentar também mais de um limiar global na separação de suas regiões, técnica conhecida como multilimiarização. O valor do limiar é quem determina o sucesso dessa segmentação, e sua má escolha pode gerar resultados insatisfatórios. Um dos métodos que auxiliam na busca por valores ótimos de limiar é o método de Otsu (OTSU, 1979).

Esse método é ótimo no sentido de que maximiza a variância entre os grupos. Utiliza a ideia de que as classes com limiares bem estabelecidos devem ser diferentes em relação aos valores de intensidade de seus pixels. Esse limiar ótimo buscado pelo método de Otsu é um valor que oferece a melhor separação entre as classes.

Considerando-se C_1 e C_2 como as duas classes de uma imagem e G como a maior intensidade de pixel, pode-se definir que o método de Otsu busca um valor ideal l no intervalo $[1, 2, \dots, G - 1]$ que maximiza a equação 2.3 da variância entre as duas classes, definida como (OTSU, 1979):

$$\sigma^2 = P_1(\mu_1 - \mu_g)^2 + P_2(\mu_2 - \mu_g)^2. \quad (2.3)$$

Em que, P_1 e P_2 são as probabilidades de ocorrência das classes C_1 e C_2 , μ_1 e μ_2 são as intensidades médias dos pixels atribuídos às mesmas classes respectivamente, e μ_g é a intensidade média de toda a imagem. Esses valores são calculadas como

$$\mu_1 = \sum_{i=0}^l iP(i)/P_1, \quad (2.4)$$

$$\mu_2 = \sum_{i=l+1}^G iP(i)/P_2, \quad (2.5)$$

$$\mu_g = P_1\mu_1 + P_2\mu_2. \quad (2.6)$$

Em que $P(i)$ é uma componente do histograma normalizado da imagem, que é calculado com a quantidade de pixels de intensidade ' i ' dividida pelo total de pixels. Para cada valor de l , é calculado um σ^2 correspondente, e o limiar escolhido é o que obtiver maior valor de σ^2 (BACKES; JUNIOR, 2016).

Essa técnica pode ser estendida para obter mais de um limiar, separando as imagens em mais que duas regiões semelhantes. Esse método de limiarização é conhecido como multinível de Otsu, e busca determinar valores de limiares ideais que separem os pixels da imagem em grupos. Assim, a variância da equação 2.3 é modificada para:

$$\sigma^2(T_1, T_2, \dots, T_k) = P_1(m_1 - m_g)^2 + P_2(m_2 - m_g)^2 + \dots + P_{k+1}(m_{k+1} - m_g)^2. \quad (2.7)$$

Em que P_k é a probabilidade ocorrência da classe k , e T_1, T_2, \dots, T_k representam os limiares, m_k é a taxa média dos pixels atribuídos à classe k e m_g é a intensidade média global (LIAO *et al.*, 2001).

A limiarização pode ser feita de formas diversas para cada região de vizinhança da imagem. Essa técnica é conhecida como limiarização adaptativa. Nesse processo, o valor do limiar é alterado para diferentes partes da imagem. Bradley e Roth (2007) propuseram uma técnica dessa categoria, que determina o limiar com base na intensidade média local, seguindo estatísticas de primeira ordem com uma janela centrada em torno dos pixels da imagem.

A Figura 5 apresenta os resultados desses métodos aplicados em imagens.

Figura 5 – Métodos de limiarização.



(a) Imagem original

(b) Global

(c) Adaptativo

(d) Multinível de Otsu

Fonte: Autor.

A Figura 5a apresenta uma imagem de entrada e as Figuras 5b, 5c e 5d representam, respectivamente, os resultados das técnicas de limiarização global, adaptativo e com o multinível de Otsu. A limiarização global utiliza limiar de 128 e o multinível de Otsu é aplicado para 4 limiares.

2.5.2 Crescimento de Região

O CR é um algoritmo utilizado para reunir pixels conectados com propriedades semelhantes para que formem uma região. A propriedade comumente utilizada é a intensidade de

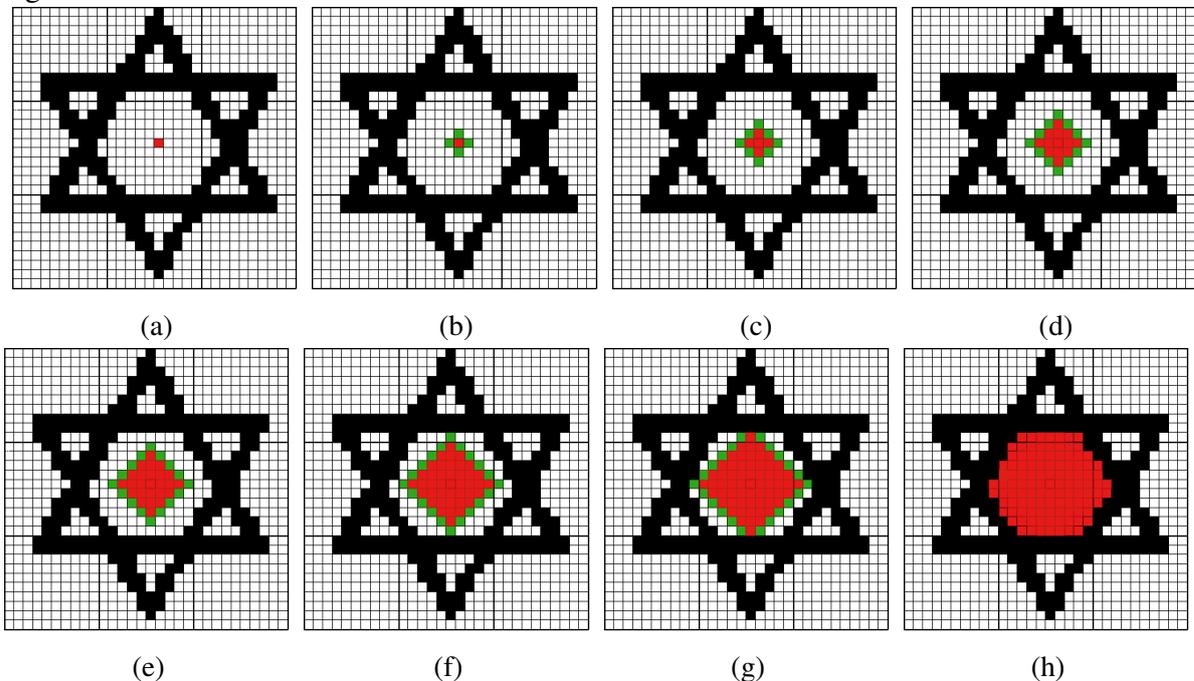
níveis de cinza. Caso o pixel verificado esteja dentro de uma tolerância adotada, ele é incorporado à região.

O objetivo desse método é selecionar um grupo de pixels conhecidos como sementes, e usar uma propriedade, como por exemplo a cor, para verificar a semelhança dos adjacentes. Os contíguos que apresentarem propriedade igual ou semelhante às sementes iniciais são selecionados para uma nova iteração. Esses novos pixels são utilizados como meio de propagação para continuar o processo até que não sejam incluídos mais pixels que satisfaçam à condição de similaridade (TANG, 2010).

Angelina *et al.* (2012) resumem esse processo em 3 etapas básicas: (1) Inicializar uma região R_{e0} com um ou mais pixels; (2) Agrupar em R_{e0} todos os vizinhos que satisfazem uma suposição ou condição específica; (3) A iteração continua até que a convergência seja alcançada.

A Figura 6 apresenta uma descrição mais detalhada desse algoritmo, aplicado em uma imagem, na qual é possível observar a região de interesse crescendo até os limites de vizinhança possíveis.

Figura 6 – Processo de um CR.



Fonte: Autor.

A Figura 6a representa a imagem de entrada do algoritmo com a semente em vermelho. Na Figura 6b, a vizinhança é representada por pixels verdes. Suas propriedades são verificadas, e caso sejam semelhantes às da semente, serão adicionados para serem novas

sementes. Na Figura 6c, a vizinhança dos pixels adicionados será verificada para ser incorporada à região de interesse. As Figuras 6d, 6e, 6f, 6g e 6h representam as demais iterações até que a vizinhança resultante não tenha as propriedades buscadas.

2.5.3 Segmentação com *K-means*

O algoritmo *K-means* é uma técnica não hierárquica que produz grupos com sobreposição. Ela é empregada em dados de forma a agrupá-los de acordo com características próprias. De forma geral, é um algoritmo que busca separar *clusters* para que minimize o erro quadrático médio entre a média de um grupo e os pontos que o compõem (MACQUEEN *et al.*, 1967).

Considere $X = x_1, x_2, \dots, x_n$ como a entrada n -dimensional de pontos a serem separados em K grupos, formando os *clusters* $C = c_1, c_2, \dots, c_k$. O erro quadrático entre a média μ_k e os pontos em um *cluster* c_k é definida na equação a seguir (JAIN, 2010):

$$J(c_k) = \sum_{x_i \in c_k} \|x_i - \mu_k\|^2. \quad (2.8)$$

Essa função modificada para a soma dos erros de todos os *clusters* é definida como:

$$J = \sum_{k=1}^K \sum_{x_i \in c_k} \|x_i - \mu_k\|^2. \quad (2.9)$$

Assim, o algoritmo busca minimizar a função objetivo denotada pela equação 2.9 convergindo para um mínimo local. Sua aplicação pode ser feita em quatro etapas principais: (1) Escolher K sementes nas amostras que serão agrupadas; (2) Atribuir a cada amostra o rótulo da semente mais próxima; (3) Reposicionar a semente para a posição média das amostras com seu rótulo; (4) Repetir os passos até a nova posição da semente convergir com a posição atual ou atingir um limite de interações (BACKES; JUNIOR, 2016).

A Figura 7 apresenta o resultado desse algoritmo aplicado em imagens.

Na Figura 7b, apresenta-se o resultado do algoritmo *K-means* aplicado à imagem da Figura 7a com $k = 3$. Observa-se que os pixels da imagem foram agrupados por semelhança de cor, os tons mais claros são representados por branco, os mais escuros por preto e as cores intermediárias, como azul e marrom, como cinza.

2.5.4 Método de Contorno Ativo

O método de Contorno Ativo (CA), introduzido por Kass *et al.* (1988), seleciona regiões da imagem pela detecção de seu contorno. Inicia-se com uma região curva que se ajusta,

Figura 7 – Exemplo do *K-means*.



(a) Imagem de entrada



(b) Resultado do *K-means*

Fonte: Imagem 7a adaptada do quadro *A Noite Estrelada* de Vincent van Gogh e imagem 7b obtida como resultado do *K-means*.

expandindo-se ou contraindo-se na estrutura de interesse. Essas curvas apresentam propriedades que definem seu comportamento ao longo das iterações, tais como energia interna e externa.

Esse modelo busca encontrar bordas e determinar fronteiras de regiões, através de minimização da função que representa a energia. Essa característica faz as curvas reagirem a forças e proporciona comportamento dinâmico nas iterações, de forma que a energia é reduzida, fazendo a curva encontrar a região de borda (KASS *et al.*, 1988).

O contorno ativo busca parametrizar uma curva geométrica expressa por $v(s) = (x(s), y(s))$, em que s representa o domínio de parâmetros de forma que $s \in [0, 1] \rightarrow \mathbb{R}^2$. A energia dessa curva é descrita como

$$E = \int_0^1 E_{int}(v(s)) + E_{ext}(v(s)) ds, \quad (2.10)$$

em que E_{int} é a energia interna e E_{ext} é a energia externa. A primeira representa forças relacionadas ao formato do contorno, modificando propriedades de elasticidade e suavização da curva. A segunda energia é o conjunto de forças que dependem da imagem e não do contorno. Nessa componente, são comumente utilizadas técnicas de segmentação por descontinuidades (CAVALCANTE, 2010; KASS *et al.*, 1988).

Outra técnica de contorno ativo baseada em segmentação por região foi proposto por Chan e Vese (2001). O modelo define uma curva de evolução C , dentro do conjunto aberto Ω , pertencente a \mathbb{R}^2 . O subconjunto aberto, formado pela curva C , é ω , de forma que $\omega \subset \Omega$ e C é a fronteira da região de ω . Essa curva é definida como $C = \partial\omega$, em que $\partial\omega$ indica as bordas do conjunto ω . As regiões que estão dentro da região ω são chamadas de *inside*(C) e as regiões externas a região ω são chamadas de *outside*(C).

Dessa forma, considera-se que uma imagem I é formada por duas regiões de intensidades aproximadamente constantes e com valores distintos, esse modelo busca minimizar as energias definidas como

$$F_1(C) + F_2(C) = \int_{inside(C)} |I(x, y) - c_1|^2 dx dy + \int_{outside(C)} |I(x, y) - c_2|^2 dx dy, \quad (2.11)$$

em que, $F_1(C)$ e $F_2(C)$ são as energias para regiões internas e externas, c_1 e c_2 são as médias das intensidades para mesmas regiões, respectivamente (CHAN; VESE, 2001).

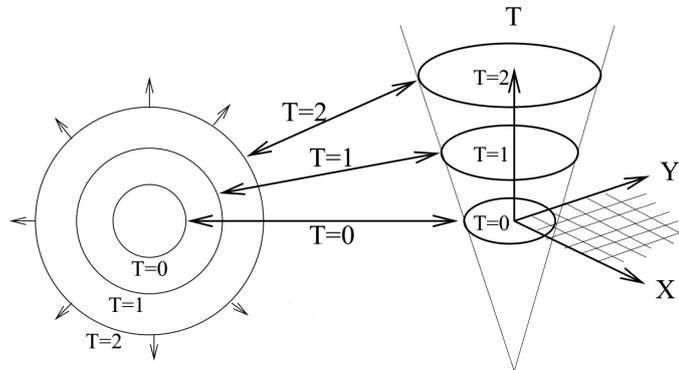
Pela equação 2.11 percebe-se que, caso a curva C esteja fora do objeto, então $F_1(C) > 0$ e $F_2(C) \approx 0$. Caso esteja dentro do objeto, então $F_1(C) \approx 0$ e $F_2(C) > 0$. E se estiver em regiões dentro e fora do objeto as energias são $F_1(C) > 0$ e $F_2(C) > 0$.

Portanto, o objetivo desse modelo é obter $F_1(C) \approx 0$ e $F_2(C) \approx 0$. A ideia é que as iterações façam evoluir a curva para encontrar a região em que a energia interna e externa sejam nulas, e que corresponde à borda do objeto.

2.5.5 Fast Marching Method

O método da marcha rápida, do inglês *Fast Marching Method* (FMM), proposto por Sethian (1999) segue o princípio de causalidade, sendo relacionado ao algoritmo de Dijkstra, que calcula o caminho de menor custo em uma rede. Esse método representa a propagação de uma interface verificando o valor de borda, em que, para cada ponto (x, y) no espaço, existe um tempo T associado. Uma interface é geometricamente uma curva ou superfície que separa dois meios que estão interagindo (PEIXOTO; VELHO, 2000).

Figura 8 – Propagação da interface no *Fast Marching Method*



Fonte: adaptado de Sethian (1999).

Esse método foi desenvolvido definindo-se uma aproximação numérica da distância

pela solução da equação Eikonal, definida por

$$|\nabla T|F = 1, \quad (2.12)$$

em que T é o tempo em que a interface se expande a cada iteração, ocupando uma nova posição no espaço, e F é a velocidade de avanço. A função T é construída a partir da variação de uma interface inicial, e apresenta modificação constante. A figura 8 apresenta a propagação da interface sobre uma matriz discreta pelo FMM (SETHIAN, 1999).

Na Figura 8, a curva em $T = 0$ é a interface inicial, e, de acordo com a evolução do algoritmo, essa região é expandida a cada nova iteração. Naturalmente, as disposições da superfície em T são conhecidos somente nos pontos internos a sua interface. No caso, $T = 1$ e $T = 2$ não são conhecidos na curva inicial. O algoritmo FMM permite calcular os valores de T de forma eficiente, à medida que ocorrem as propagações.

O FMM foi desenvolvido para resolver problemas de borda da equação de Eikonal. Esta é uma especificação da equação de Hamilton-Jacobi que representa a propagação positiva de interfaces. A solução de viscosidade sobre essa equação é representada por

$$|\nabla u(x)| = F(x), \quad (2.13)$$

em que $F(x)$ é a velocidade na direção normal da curva em um ponto x . Essa solução pode ser representada para multi-dimensões como

$$|\nabla u(x, y, z)| = f(x, y, z). \quad (2.14)$$

Uma aproximação desse gradiente para multi-dimensões pode ser escrita como

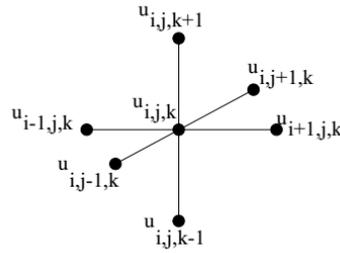
$$|\nabla u| \approx \left[\begin{array}{c} \max(D_{ijk}^{-x}u, -D_{ijk}^{+x}u, 0)^2 + \\ \max(D_{ijk}^{-y}u, -D_{ijk}^{+y}u, 0)^2 + \\ \max(D_{ijk}^{-z}u, -D_{ijk}^{+z}u, 0)^2 \end{array} \right]^{1/2} = f_{ijk}, \quad (2.15)$$

em que, D^x , D^y e D^z são operadores de diferenças finitas nos eixos de coordenadas de frente e trás, representados por sinal positivo e negativo respectivamente.

O FMM testa os valores vizinhos da interface como entrada u na equação 2.15 para resolvê-la. A cada iteração, novos pontos são calculados e podem ser aderidos a interface. (SETHIAN, 1999; DEVI *et al.*, 2018).

A equação 2.15 é uma expressão quadrática para u_{ijk} , que é resolvida pelo FMM utilizando-se os valores dos vizinhos na grade. A Figura 9 demonstra a grade de pontos vizinhos de u_{ijk} utilizados para resolver essa equação.

Figura 9 – Pontos vizinhos da grade.



Fonte: adaptado de Sethian (1999).

2.6 Operações de Morfologia Matemática

Em processamento de imagens, operações de morfologia matemática são operações baseadas na teoria dos conjuntos que possibilitam representar e descrever formas. O princípio consiste em transladar um conjunto sobre a imagem e fazer relações entre ele e os elementos presentes na imagem.

As operações básicas dessas técnicas são conhecidas como erosão e dilatação. Elas são fundamentais para o processamento morfológico, pois derivam-se em outras transformações, como por exemplo abertura e fechamento. Nas expressões seguintes, B é o conjunto de translação, conhecido como elemento estruturante, e A é o conjunto no qual ocorrerá as operações.

A erosão entre os dois conjuntos A e B no espaço Z^2 , é definida como

$$A \ominus B = \{z | (B_z) \subseteq A\}, \quad (2.16)$$

em que z é o conjunto de todos os pontos em que a translação de B por z esteja contida em A . Aplicada em imagens binárias resulta na redução ou afinamento de objetos (GONZALEZ; WOODS, 2009).

A dilatação entre os mesmos conjuntos é expressa como

$$A \oplus B = \{z | (\hat{B})_z \cap A \neq \emptyset\}, \quad (2.17)$$

em que $(\hat{B})_z$ é a translação da reflexão de \hat{B} , em torno da origem por z , e z é o conjunto de todos os deslocamentos de forma que $(\hat{B})_z$ e A se sobreponham em pelo menos um elemento. A aplicação dessa operação aumenta a área dos objetos (GONZALEZ; WOODS, 2009).

A combinação dessas duas primitivas formam as operações de abertura e fechamento, que apresentam característica de suavizar contornos. A abertura rompe pequenas ligações entre objetos e elimina saliências, e é expressa conforme a equação abaixo:

$$A \circ B = (A \ominus B) \oplus B. \quad (2.18)$$

O fechamento, realiza o preenchimento de lacunas ou buracos presentes nos objetos e é expresso conforme a equação abaixo (GONZALEZ; WOODS, 2009):

$$A \bullet B = (A \oplus B) \ominus B. \quad (2.19)$$

As operações morfológicas também podem ser aplicadas em imagens com tons de cinza. Neste caso, os conjuntos são definidos com um elemento estruturante, $b(x, y)$, transladando sobre uma imagem $f(x, y)$. Algoritmos representantes dessa aplicação são as transformadas *top-hat* e *bottom-hat*, que combinam as operações de abertura e fechamento.

A transformada *top-hat* é definida como

$$T_{hat}(f) = f - (f \circ b). \quad (2.20)$$

De forma semelhante, a transformada *bottom-hat* é definida como

$$B_{hat}(f) = (f \bullet b) - f. \quad (2.21)$$

Esses operadores são utilizados para remover objetos de uma imagem através do elemento estruturante. A transformada *bottom-hat* é utilizada para objetos escuros em um fundo claro e a *top-hat* para objetos claros em um fundo escuro. A remoção desses elementos em tons de cinza decorre na alteração do brilho da imagem, podendo ser combinado ambos operadores para modificação do contraste (GONZALEZ; WOODS, 2009).

2.7 Descritores de Textura

Descritores de textura são empregados para descrever regiões da imagem de forma a discriminá-las. São medidas matemáticas utilizadas para representar a textura percebida visualmente pelo homem, e quantificá-la.

A descrição de elementos por textura considera características como suavidade, rugosidade e regularidade. As principais formas de descrever essas características em uma imagem são através das abordagens estatísticas, estrutural e espectral (GONZALEZ; WOODS, 2009). Dentre esses descritores, os de Haralick destacam-se como uma textura clássica de VC. Além destes, outros descritores têm sido desenvolvidos e aplicados com imagens médicas, como exemplo a textura LBP.

2.7.1 Descritores de Haralick

Os descritores de Haralick foram propostos por Haralick *et al.* (1973) que descrevem as texturas com base em estatísticas de segunda ordem, através da relação espacial entre níveis de cinza da imagem. São calculados por meio de matrizes estatísticas de coocorrência de níveis de cinza.

A matriz de coocorrência é quadrada e possui tamanho definido pela quantidade de níveis de cinza da imagem ou região analisada. Ela descreve a frequência de transições entre níveis de cinza, de forma que, para uma matriz, G , um dado valor $G(i, j)$ da matriz é a quantidade de vezes que o nível i variou para o nível j em uma determinada direção. Os ângulos utilizados são de 0, 45, 90, 135 e simetrias destes (HARALICK *et al.*, 1973).

Com a matriz de co-ocorrência, são calculadas as probabilidades de ocorrência das combinações que a constituem, formando-se outra matriz. Através desta nova matriz, são extraídos os 14 atributos estatísticos de textura definidos por Haralick *et al.* (1973). Alguns destes que serão descritos a seguir. Nessas definições, p é a matriz de correlação normalizada, N_g é o número de níveis de cinza distintos na imagem e μ_x , μ_y , σ_x , σ_y são as médias e desvios padrões de linhas e colunas.

A energia (definida originalmente como segundo momento angular), representa uma medida de homogeneidade da imagem, denotada por:

$$Energia = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p(i, j)^2. \quad (2.22)$$

O contraste representa o momento de diferença da matriz p ou a quantidade de variações locais presentes, definido por:

$$Contraste = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \right\}_{|i-j|=n}. \quad (2.23)$$

A correlação é uma medida que representa o quanto um pixel está correlacionado com os outros pixels da imagem, e é definida por:

$$Correlação = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (ij)p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}. \quad (2.24)$$

A variância é uma medida de heterogeneidade em forma de desvio dos valores p da matriz, representado por:

$$Variância = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu)^2 p_{i,j}. \quad (2.25)$$

Em que μ representa as média das distribuições. O momento da diferença inversa (MDI) é a homogeneidade da matriz de co-ocorrência e representa o quão próximo os elementos estão da diagonal. É calculado como

$$MDI = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \frac{1}{1 + (i-j)^2} p(i, j). \quad (2.26)$$

A entropia representa a dispersão dos níveis de cinza da imagem e mede a informação contida em p , podendo ser definida como

$$Entropia = - \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p(i, j) \log(p(i, j)). \quad (2.27)$$

Os demais atributos são: soma da média, soma da variância, soma da entropia, diferença da variância, diferença da entropia, duas medidas de informação da correlação e coeficiente de correlação máximo. A formulação desses demais descritores pode ser encontrada em Haralick *et al.* (1973).

2.7.2 Local Binary Pattern

Este algoritmo de representação de textura, proposto por Ojala *et al.* (2002), é robusto às variações da escala de cinza e rotação, sendo assim invariante a essas transformações. Sua implementação é simples e eficiente do ponto de vista computacional.

O LBP consiste em atribuir rótulos binários para cada pixel da imagem. Esses rótulos são obtidos através de uma operação de vizinhança do pixel analisado. A textura da imagem é o histograma dos rótulos obtidos pelo algoritmo (OJALA *et al.*, 2002).

O operador LBP é baseado em um conjunto circular e simétrico de vizinhos P em uma distância de raio R , sendo denotado por $LBP_{P,R}^{u^2}$. Nessa representação, o parâmetro R define o espaço do operador. P está relacionado à quantização do espaço angular, e u^2 é a definição de quantas variações entre bits são necessárias para o padrão uniforme.

Os pixels vizinhos de um dado pixel central, dentro do espaço controlado por (P, R) , são verificados. A diferença entre os pixels da vizinhança e o central é a entrada de uma função s , definida como (OJALA *et al.*, 2002)

$$s(x) = \begin{cases} 1, & \text{caso } x \geq 0 \\ 0, & \text{caso } x < 0 \end{cases} \quad (2.28)$$

Dessa forma, analisando-se os vizinhos no espaço (P, R) , considerando que g_p é o nível de cinza do pixel vizinho corrente e g_c é o nível de cinza do pixel central, o código LBP,

que caracteriza a textura para o ponto g_c é calculado como (OJALA *et al.*, 2002)

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (2.29)$$

em que 2^p é um peso atribuído a cada sinal $s(g_p - g_c)$, com o objetivo de transformar as diferenças em um código único LBP. Na equação 2.29, P define a quantidade de bits do código e 2^P são os possíveis valores do LBP.

Uma modificação desse operador consiste em utilizar para a textura somente os chamados padrões uniformes, o que implica calcular a uniformidade do padrão binário obtido com o LBP. Um código binário é considerado uniforme se apresentar no máximo duas variações de bits (OJALA *et al.*, 2002).

Dessa forma, os pixels que recebem os rótulos LBP são aqueles que apresentam padrão uniforme. Os que não o apresentam são marcados com somente um rótulo, de forma a agrupá-los. O número de rótulos para mapeamento desses padrões é definido como

$$nP = P(P - 1) + 3, \quad (2.30)$$

em que, o número de rótulos é denotado por nP . Considerando-se $P = 8$, o mapeamento apresenta 59 rótulos. Se não utilizasse a ideia de padrão uniforme, o mapeamento apresentaria 256 rótulos.

2.8 Algoritmos de aprendizado de máquina

Aprendizado de máquina é uma área da IA que consiste em técnicas computacionais que buscam detectar padrões em conjuntos de dados. Esse reconhecimento pode ser utilizado para o algoritmo aprender a executar uma tarefa sozinho, realizar previsões ou desenvolver sistemas inteligentes. Essas técnicas utilizam-se de sistemas de aprendizado, que definem a forma como determinado algoritmo aprende com os dados de entrada, utilizando conhecimento prévio ou não.

Esses sistemas apresentam características e abordagens particulares, mas de forma geral, utilizam-se de mesmos mecanismos de aprendizagem, tais como: aprendizado supervisionado, não-supervisionado ou por reforço (BOUSQUET *et al.*, 2011; WITTEN *et al.*, 2016).

Os sistemas computacionais que utilizam aprendizado de máquina são aplicados, por exemplo, na identificação de objetos em imagens, transformação de fala em texto, combinar itens de notícias, entre outras. Essas aplicações inicialmente exigiam conhecimento prévio do problema e uma análise cuidadosa para a escolha dos atributos de descrição. Modelos modernos

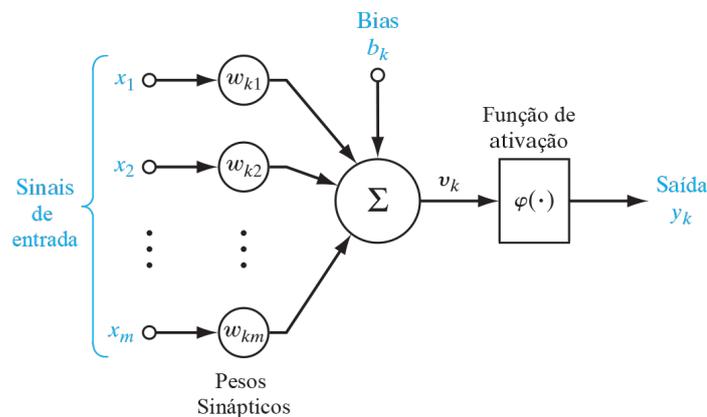
dessas técnicas, conhecidos como aprendizado profundo, superam essas limitações (LECUN *et al.*, 2015). Nas sub-seções seguintes, são apresentados os algoritmos de aprendizado de máquina utilizados nessa pesquisa.

2.8.1 Redes Neurais Artificiais

As Redes Neurais Artificiais (RNA) modelam o neurônio biológico e simulam seu comportamento, sendo utilizadas na resolução de problemas computacionais complexos, como por exemplo, a classificação de amostras em problemas não-lineares.

Essa abordagem foi inicialmente proposta por McCulloch e Pitts em 1960, que desenvolveram um modelo matemático para representar um neurônio biológico. Esta formulação é conhecida como neurônio artificial ou perceptron, e simula a decisão se um sinal deve seguir ou não (HAYKIN *et al.*, 2009). A Figura 10, apresenta o modelo perceptron simples do neurônio artificial.

Figura 10 – Perceptron simples.



Fonte: Adaptado de Haykin *et al.* (2009).

Nessa representação, o conjunto x_1, x_2, \dots, x_m são os sinais de entrada, $w_{k1}, w_{k2}, \dots, w_{km}$ são os pesos associados às entradas, b_k é o bias, que tem efeito de aumentar ou reduzir a entrada da rede, φ é a função de transição do neurônio e y representa a saída. O parâmetro φ é conhecido como função de ativação. Exemplos das possíveis funções de ativação são: a sigmóide, degrau, tangencial e gaussiana.

As redes neurais precisam passar por uma etapa de aprendizado. Nessa etapa, a rede é treinada para que os pesos w sejam ajustados através de algoritmos de treinamento. O objetivo é que os pesos sejam modificados de forma que as saídas sejam satisfatórias para padrões do treinamento conhecido, reduzindo-se a diferença entre a saída obtida e a desejada (BISHOP,

2006).

As redes neurais mais simples que empregam esses modelos são conhecidas como redes perceptron simples. São utilizadas para resolução de problemas de natureza lineares. Uma abordagem dessas redes para resolução de problemas não lineares é a rede conhecida como *Multi-Layer Perceptron* (MLP). Essa rede é caracterizada por apresentar três ou mais camadas: uma camada de entrada e uma de saída, com uma ou mais camadas intermediária. Os neurônios apresentam ligações entre as diferentes camadas e pesos sinápticos para cada ligação.

2.8.2 *Support Vector Machine*

Máquina de Vetores de Suporte, do inglês *Support Vector Machine* (SVM), é um algoritmo de inteligência computacional baseado em aprendizado estatístico, que utiliza maximização da margem entre os dados para separação em classes. Esse algoritmo é essencialmente binário, mas pode ser estendido a trabalhar com multiclases, além da possibilidade de aplicação em problemas de classificação ou regressão (HSU *et al.*, 2016; CORTES; VAPNIK, 2017).

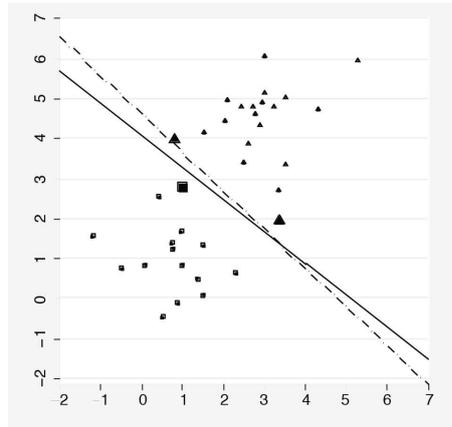
Esse classificador é caracterizado por apresentar boa capacidade de generalização, ser robusto no trabalho de dados com grande dimensão e apresentar convexidade do problema de otimização em seu treinamento.

As SVMs originalmente são utilizadas na classificação de conjuntos de dados linearmente separáveis. Esses algoritmos buscam uma equação de hiperplano ótimo seguindo a teoria de aprendizado estatístico, de forma que a separação apresente margem maximizada. Os pontos de dados mais próximos do hiperplano são chamados de vetores de suporte (GUENTHER; SCHONLAU, 2016).

A Figura 11 apresenta um exemplo de escolha do hiperplano para duas classes. Nessa figura, observa-se a possibilidade de dois hiperplanos para separar a classe de triângulos e quadrados. O SVM busca aquele que apresentar maior margem entre os vetores de suporte, representados pelos dados destacados. Esse hiperplano é denotado com a linha contínua.

Uma generalização do algoritmo SVM foi proposta com o objetivo de dividir dados não-linearmente separáveis. Essa tarefa é feita por meio de funções que mapeiam as amostras do domínio de entrada para um novo espaço, de forma que nesse novo domínio esses dados sejam linearmente separáveis. Estas são conhecidas como funções de *kernel*. As mais comuns são: Gaussianas, *Radial Basis Function* (RBF) e polinomiais (GUENTHER; SCHONLAU, 2016).

Figura 11 – Hiperplanos formados pelo SVM.



Fonte: Adaptado de Guenther e Schonlau (2016).

2.8.3 *Extreme Learning Machine*

Esse algoritmo foi proposto por Huang *et al.* (2004) como uma técnica de aprendizado para redes neurais *feedforward* com uma camada oculta. A abordagem consiste em atribuir pesos aleatórios na entrada da camada oculta, que não serão ajustados e independem dos dados de treinamento. Dessa forma, é eliminada a utilização do gradiente descendente no treinamento, que apresentava convergência lenta.

Com essas considerações, o modelo resulta em um sistema linear e os pesos sinápticos da camada de saída são determinados analiticamente, através de uma operação com a matriz inversa generalizada da matriz de saída da camada oculta. Portanto, o treinamento da ELM não apresenta iteratividade como o *backpropagation*, e sua velocidade de aprendizado pode chegar a milhares de vezes mais rápido, enquanto tende a apresentar melhor generalização (HUANG *et al.*, 2004).

As limitações obtidas pela ELM consistem justamente na escolha aleatória dos pesos, que pode gerar um conjunto de dados que desencadeie em uma matriz de pesos não ótimos da camada de saída, fornecendo resultados ruins. Além disso, a matriz de saída da camada oculta pode ser singular e impossível de ser invertida.

2.8.4 *Redes Neurais Convolucionais*

Redes neurais convolucionais são um tipo de rede neural que substitui a multiplicação geral da matriz pela operação de convolução em pelo menos uma camada (GOODFELLOW *et al.*, 2016). Essa técnica ficou consagrada na literatura por LeCun *et al.* (1998), que a aplicaram em diversas tarefas de reconhecimento de imagens.

A estrutura de uma CNN apresenta camadas de convolução, agrupamento e classificação, que podem ser agrupadas de formas diferentes, conforme a arquitetura. A camada de convolução é formada por um conjunto de mapas (semelhantes a neurônios) com diferentes vetores de pesos, para que recursos diversos possam ser extraídos. A camada de agrupamento realiza uma média local, fazendo uma subamostragem e reduzindo a resolução do mapa de recursos. A última camada é uma rede neural clássica *feedforward* para a classificação das características extraídas.

Os mapas que fazem as convoluções na imagem de entrada são conhecidos como filtros de *kernel*, e seus valores são obtidos dos pesos dos neurônios, através da etapa de treinamento com o *backpropagation*. Para cada filtro dessa camada, é gerada uma nova imagem pela sua convolução com a entrada.

O agrupamento é conhecido como *pooling*, sendo realizada a redução de sua dimensionalidade na representação dos dados. Essa redução é feita substituindo a saída da camada anterior, em um determinado local, por um valor estatístico. A abordagem mais comum utilizada é o *max pooling* que computa o máximo valor de um determinado local do mapa de características para ser a saída, eliminando os demais dessa região (GOODFELLOW *et al.*, 2016).

As CNNs representam um grande avanço para a área de VC, pois permitem a detecção de padrões em imagens com a extração de características realizada automaticamente. Essa técnica é considerada uma abordagem de aprendizado profundo, pois utiliza grande número de camadas, e tem sido aplicada em diferentes campos.

2.9 Métricas de avaliação

As métricas utilizadas para avaliar a metodologia são estatísticas clássicas utilizadas na área da saúde. Tais métricas objetivam comparar as predições obtidas pelos algoritmos com os rótulos fornecidos pelos profissionais de odontologia. As marcações dos especialistas são conhecidas como *Ground Truth* (GT) e a comparação com os resultados da metodologia gera primitivas utilizadas em categorias de diagnóstico, tais como Verdadeiro Positivo (VP), Verdadeiro Negativo (VN), Falso Positivo (FP) e Falso Negativo (FN).

Uma amostra é considerada VP se for rotulada como cisto pela metodologia e constar como cisto pelos especialistas. Caso essa seja marcada como não-cisto pela metodologia e esteja rotulada como cisto pelos especialistas, então é um FN. VN é quando essa amostra é marcada como não-cisto pela metodologia e consta como não-cisto pelos especialistas. Por fim, FP é

quando a amostra é marcada como cisto pela metodologia e é rotulado como não-cisto pelos especialistas.

Cada amostra analisada pela metodologia é rotulada em uma das categorias apresentadas acima. Essas classes são utilizadas para o cálculo das métricas de Acurácia (Ac), Sensibilidade (Se), Especificidade (Es) e Coeficiente Dice.

A Ac representa a taxa de acertos total, ou seja, o quanto o resultado obtido pela metodologia se aproximou do GT, e é calculada como

$$Ac = \frac{VP + VN}{VP + VN + FP + FN}. \quad (2.31)$$

A Se representa a taxa de amostras rotuladas como patógenas que realmente correspondem com o GT e mede o quanto a metodologia é hábil em rotular VPs. É calculada como

$$Se = \frac{VP}{VP + FN}. \quad (2.32)$$

A Es é a taxa de não-cistos corretamente rotulados pela metodologia dentre todas as amostras rotuladas como não-cistos, medindo o quanto a metodologia é hábil em rotular VNs, sendo expressa como

$$Es = \frac{VN}{VN + FP}. \quad (2.33)$$

O coeficiente Dice, proposto por Dice (1945) e Sørensen (1948), representa o grau de semelhança entre duas amostras. Em VC, pode ser utilizado para avaliar o quanto uma região marcada pelo algoritmo se aproxima do GT. Essa informação é útil para avaliar a qualidade da segmentação de imagens. Esse índice é calculado como

$$Dice = \frac{2VP}{2VP + FP + FN}. \quad (2.34)$$

2.10 Considerações Finais

Este capítulo apresentou toda a base teórica necessária para o entendimento da metodologia empregada, bem como as características das imagens utilizadas e o aspecto visual dos cistos nos exames. Foram detalhadas as etapas de um sistema de VC, os métodos de segmentação utilizados, os descritores de textura, os algoritmos de aprendizado de máquina e as métricas estatísticas de avaliação.

O próximo capítulo abordará a metodologia utilizada nas investigações propostas, bem como detalhes dos experimentos realizados e algoritmos desenvolvidos.

3 METODOLOGIA

As etapas necessárias para a detecção automática de cistos em imagens de radiografia panorâmica, que são descritas neste capítulo, utilizam as técnicas apresentadas no capítulo 2. A Figura 12 ilustra o fluxograma da metodologia adotada e cada uma das etapas envolvidas.

A metodologia inicia na etapa 1 da figura, com o processo de formação do banco de imagens e marcação do GT, descrito na seção 3.1. De posse das imagens, a etapa 2 descrita na seção 3.2 apresenta os pré-processamentos utilizados, a fim de preparar as imagens para processos posteriores. A etapa 3, descrita na seção 3.3, apresenta o processo realizado para obter a região de interesse da mandíbula. O estágio 4, descrito na seção 3.4, apresenta a aplicação de algoritmos para selecionar as áreas mais semelhantes a cistos. O processo de extração de texturas e treinamento de classificadores é representado no estágio 5 e descrito na seção 3.5. O estágio 6 apresenta a segmentação final do cisto, descrito na seção 3.6. A etapa 7 representa a aplicação da CNN, descrita na seção 3.7, que é treinada e testada a partir das imagens de entrada e das imagens com a segmentação inicial. Por fim, a seção 3.8 representada no estágio 8 descreve os cálculos utilizados para obter as métricas, comparando o GT com as marcações resultantes dos algoritmos.

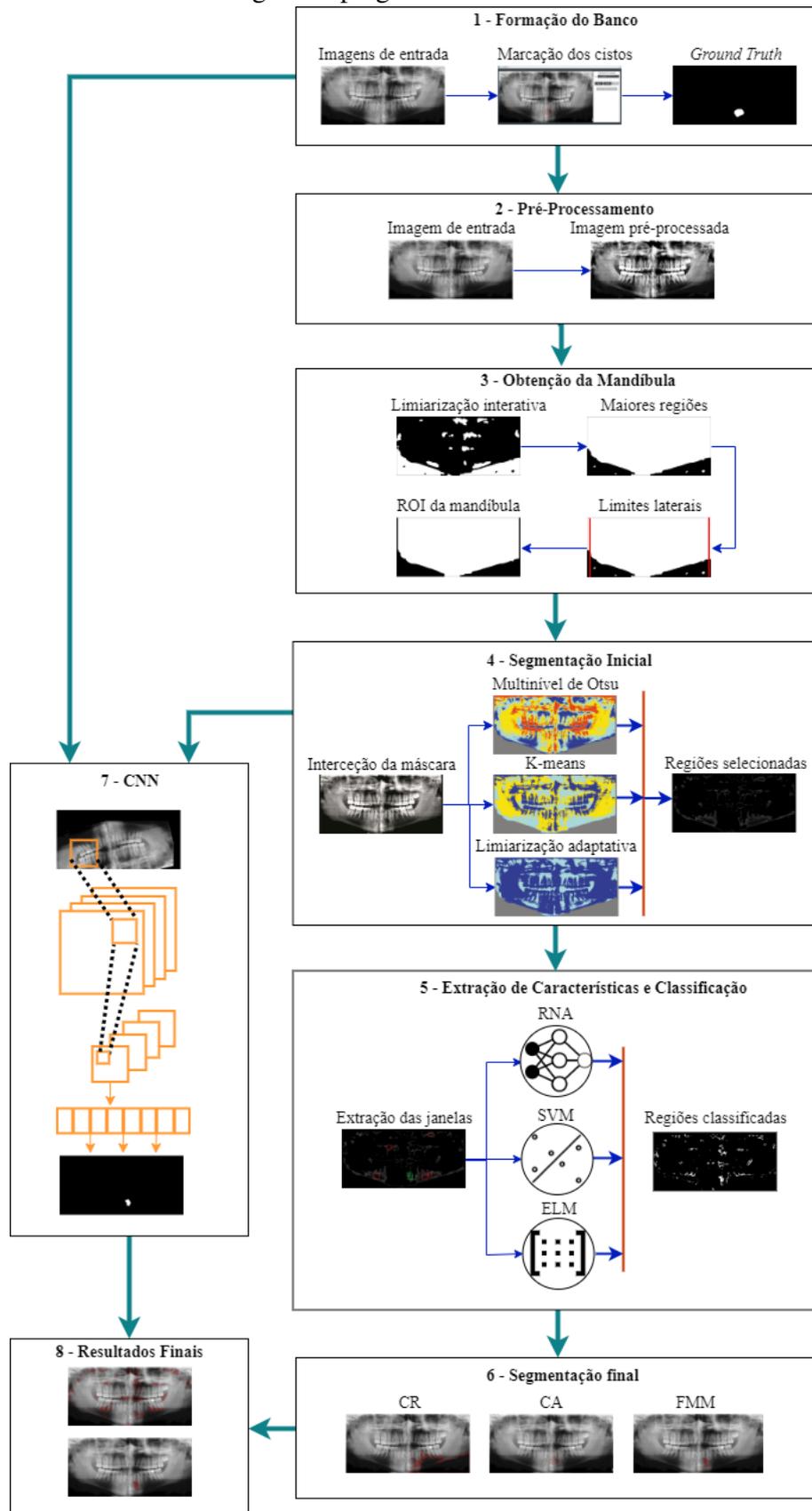
São realizados, de acordo com o fluxo seguido, três experimentos: O primeiro, com classificação por textura (etapas 1, 2, 3, 4, 5, 6 e 8), aplicação direta da CNN nas imagens (etapas 1, 7 e 8) e uma proposta híbrida da CNN com a metodologia desenvolvida, utilizando como entrada da CNN as imagens segmentadas (etapas 1, 2, 3, 4, 7 e 8). Os algoritmos da metodologia empregada nessa pesquisa foram desenvolvidos, majoritariamente, em MATLAB, com exceção do sistema para marcação, que utiliza Java e PHP, e da implementação da CNN, que utiliza Python.

3.1 Formação do banco de imagens

Os exames utilizados nessa pesquisa são radiografias panorâmicas, obtidas em parceria com o curso de Odontologia da UFC - *Campus* Sobral. Essas imagens foram processadas em níveis de cinza, com codificação de 8 bits, e redimensionadas para tamanho de 1050 x 500 *pixels*. As imagens apresentam duas origens: digitalizadas de exame de filme e de radiografia digital. A Figura 13 apresenta exemplos de exames utilizados nessa pesquisa.

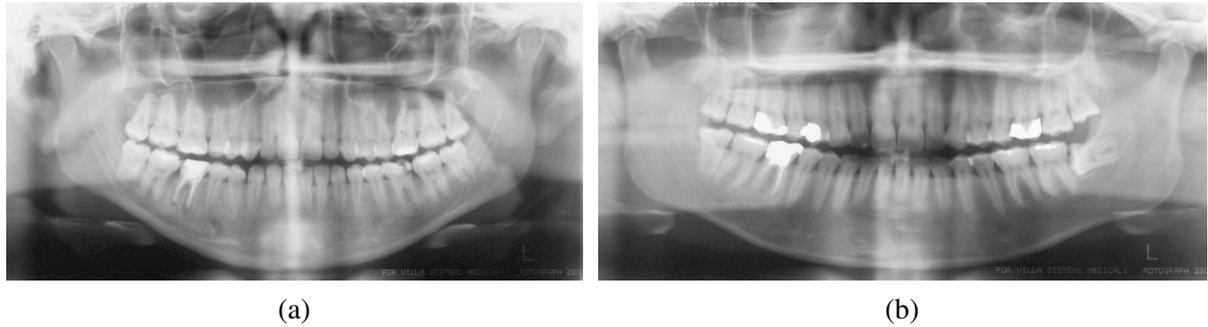
Iniciou-se o processo de formação do banco através da análise dos exames e indicação

Figura 12 – Fluxo das metodologias empregadas.



Fonte: Autor.

Figura 13 – Exemplos dos exames adquiridos.

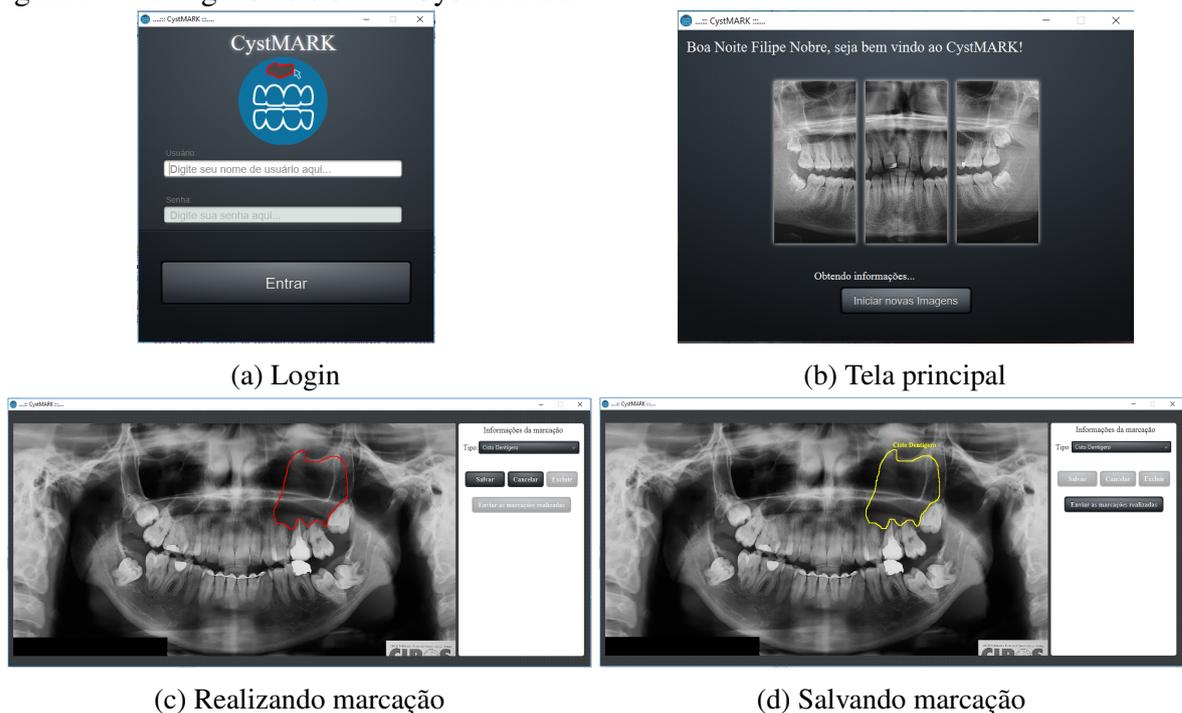


Fonte: Imagens obtidas em parceria com o curso de Odontologia - UFC / *Campus* Sobral

pelos especialistas dos locais que apresentam cistos. Essas marcações foram feitas por dois cirurgiões-dentistas especializados em radiografia do curso de Odontologia da UFC - *Campus* Sobral. Nessa etapa, desenvolveu-se um programa denominado *CystMARK*, para que auxilie na seleção das patologias e formação do GT.

Essa aplicação foi desenvolvida como parte desse trabalho, em JavaFX, uma plataforma multimídia que proporciona ao usuário uma interface rica e iterativa. Através dela, o especialista não só pôde avaliar os exames e marcar as regiões císticas, como também indicar o tipo de cisto que foi marcado. Esse sistema obtém as imagens de uma API desenvolvida em PHP, a qual também recebe as coordenadas das marcações feitas e as salva em arquivos de texto. A Figura 14 apresenta as imagens desse sistema.

Figura 14 – Imagens do sistema *CystMARK*.

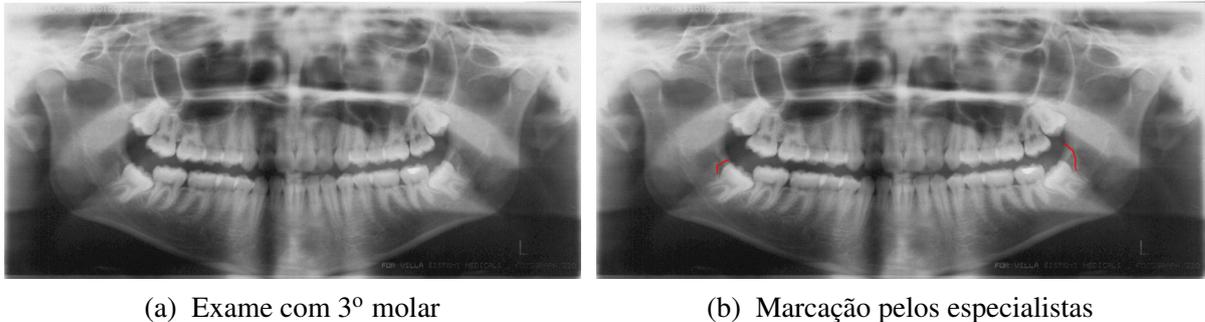


Fonte: Autor

A Figura 14a apresenta a tela de login e, a Figura 14b mostra a tela inicial, que informa ao usuário se há novos exames a serem analisados. A Figura 14c ilustra uma marcação sendo feita no programa, e a Figura 14d apresenta o momento de confirmação da marcação para salvar as coordenadas e o tipo de cisto indicado.

Ao final, foram analisados 600 exames, sendo que 242 apresentam algum tipo de cisto indicado por pelo menos um especialista. A grande maioria desses 242 exames possuem cistos dentígeros do 3º molar, que apresentam padrão diferenciado dos demais tipos e possível posição conhecida. Esses cistos apresentaram diversas divergências de marcações entre especialistas, e optou-se em não utilizá-los na pesquisa. A Figura 15 apresenta um exame com essa categoria da patologia.

Figura 15 – Exame com 3º molar.



Fonte: Imagem 15a obtida em parceria com o curso de Odontologia - UFC / Campus Sobral e Imagem 15b formada pela marcação dos especialistas.

A Figura 15a apresenta um exame com esse tipo de cisto, e a Figura 15b a marcação indicada pelos especialistas em vermelho. Dentre todas as imagens selecionadas com cistos, houve concordância em 86 e discordância em 156. Dessa forma, foram realizadas reuniões a fim de obter indicações com o consenso de ambos. Por fim, foi estabelecido como padrão de referência da pesquisa, o GT com 32 exames de cistos não dentígeros do 3º molar. Este tipo conta com uma imagem binária para cada exame, com a região de interesse (do inglês *Regions of Interest (ROI)*), demarcada pelos especialistas. A Figura 16 mostra um exame com marcação e sua imagem binária com a ROI. É possível observar que as marcações indicadas pelos especialistas não correspondem às fronteiras do cisto, que é o padrão buscado. Dessa forma, foi preciso marcar as fronteiras císticas com base nas indicações dos especialistas.

Observou-se nesta etapa que as imagens adquiridas pela digitalização de exames de filmes apresentavam problemas de aquisição que influenciaram nas etapas subsequentes, além de algumas amostras digitais que também possuíam problemas de qualidade. Assim, o banco foi

Figura 16 – Imagem do GT



(a) Marcação dos especialistas

(b) Máscara binária

Fonte: Imagens obtidas pelo processo de marcação.

dividido em dois grupos de exames: um com as 12 imagens de melhor qualidade e outro com todas as obtidas do GT, ambos sendo utilizados em todas as etapas da metodologia. Com isso, foi possível avaliar a influência da qualidade de aquisição no sistema desenvolvido.

3.2 Pré-processamento

Nessa etapa, são realizados procedimentos para melhoria da qualidade das imagens e ajustes das intensidades de tons, para melhor aplicação das etapas posteriores. O processo inicia-se com a execução do filtro gaussiano nas imagens, a fim de atenuar o ruído de alta frequência produzidos pelo processo de aquisição. Esse filtro provê à imagem um aspecto de suavização das bordas, sendo o mesmo aplicado pela convolução da função gaussiana às imagens. Essa função é expressa como

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \quad (3.1)$$

em que σ controla a largura de banda do filtro e x e y representam os valores do eixo das abscissas e ordenadas, respectivamente. A convolução da imagem com essa função gera uma nova imagem através da discretização e aproximação dos valores resultantes. Na aplicação desse filtro, nos exames, utilizou-se uma máscara 3x3 e $\sigma = 1$, valores esses obtidos empiricamente.

Após a filtragem, é realizada uma operação de aumento de contraste nas imagens, buscando-se acentuar as diferenças de intensidades de cinza nas estruturas, para diferenciá-las com melhor eficiência. As operações de melhoria de contraste realizam modificações no histograma das imagens, para realçar determinadas estruturas de interesse. A técnica aplicada para esse fim utiliza as operações morfológicas descritas na seção 2.6, sendo baseada em Birdal

et al. (2016). A equação 3.2 apresenta a operação de contraste morfológico utilizada.

$$I_c = (I + T_{hat}(I)) - B_{hat}(I), \quad (3.2)$$

em que I é a imagem de entrada, e I_c é a saída com maior contraste. O termo T_{hat} representa o resultado da técnica *top-hat*, e o termo B_{hat} representa o resultado para a técnica *bottom-hat*. Essa operação fornece I_c com maior contraste entre as estruturas de que a imagem I . A Figura 17 apresenta o aspecto adquirido da imagem após a etapa de pré-processamento.

Figura 17 – Etapa de pré-processamento.



(a) Exame original

(b) Exame pré-processado

Fonte: Imagem 17a obtida em parceria com o curso de Odontologia - UFC / Campus Sobral e Imagem 17b obtida após pré-processamento.

A Figura 17a apresenta um exame original, e a Figura 17b, o resultado do pré-processamento. Observa-se que a técnica fornece à imagem maior diferenciação entre as estruturas anatômicas.

3.3 Obtenção de regiões da mandíbula

Essa etapa busca separar a região da mandíbula do restante da imagem. Esse processo é necessário para restringir a área de busca dos cistos à região maxilar interna, local da ocorrência dessas patologias. Ao final, espera-se obter uma imagem binária com uma ROI indicando a área de busca para próximas etapas.

Para esse fim, é utilizado um processo de limiarização iterativa que busca um valor de limiar para separar as regiões externas abaixo da curvatura do maxilar do restante da imagem. Essas áreas externas são mais escuras.

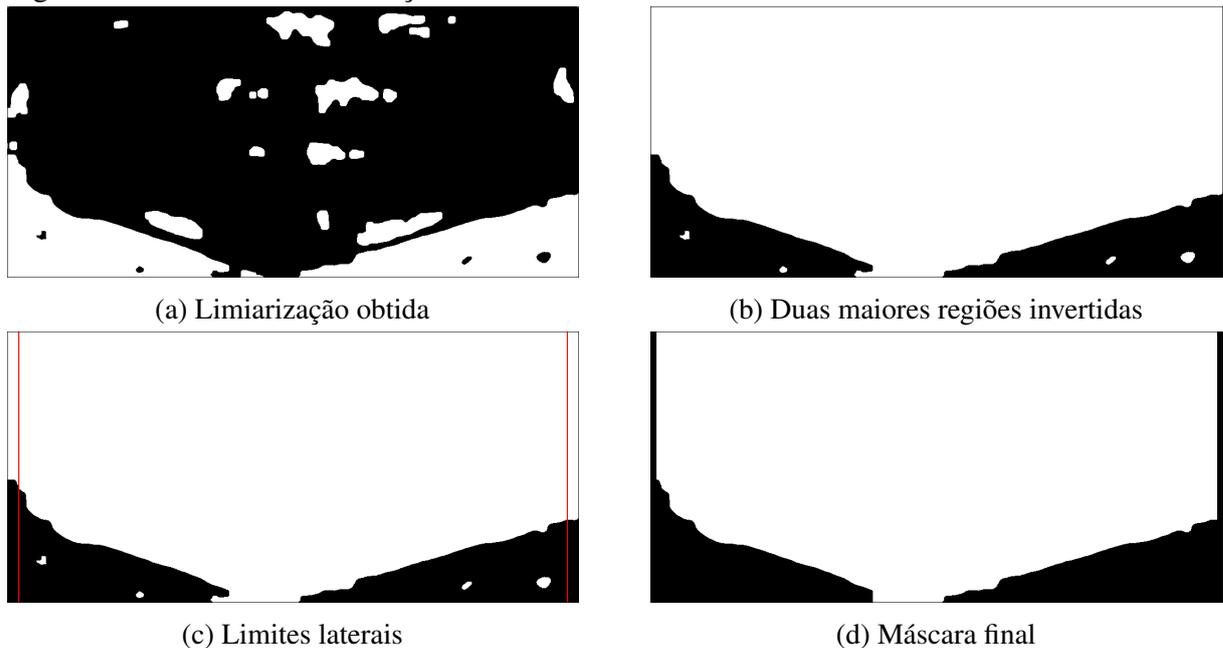
Esse processo é executado enquanto as duas maiores áreas externas (lado direito e esquerdo) forem menores que 20000, valor este obtido empiricamente. A cada iteração, a imagem é binarizada globalmente (vide subseção 2.5.1) com o limiar corrente e é realizada a

operação de fechamento (vide seção 2.6). Além disso, nesse processo, também é calculado o limite inferior da mandíbula. Esse limite é utilizado para cortar uma parte inferior da imagem, evitando que as regiões externas do lado direito e esquerdo tenham conexão. O resultado da limiarização é apresentado na Figura 18a.

A imagem binária é invertida, para que a região branca seja a região externa e são analisados os objetos brancos conectados presentes na imagem. Através dessa análise, são selecionadas as duas maiores regiões, caso as áreas dessas regiões não atinjam o tamanho especificado, o processo é reiniciado com o novo limiar.

As áreas encontradas são adicionadas em uma nova imagem. Essa imagem possui apenas valores 0, e após isso, apresenta 1 para as áreas adicionadas. Como a ROI da mandíbula deve ter valor 1, e as regiões externas valor 0, a imagem é invertida. O resultado desse processo é apresentado na Figura 18b.

Figura 18 – Processo de obtenção da mandíbula.



Fonte: Imagens obtidas durante o processo de obtenção da mandíbula.

O próximo passo é obter limites à direita e à esquerda, para que a ROI seja a área branca entre esses limites. O limite da esquerda é obtido através de uma busca que percorre da segunda coluna até a metade da largura da imagem. Em cada iteração é feita a soma dos valores de todas as linhas da coluna corrente, buscando qual apresenta a menor quantidade de valores 1. A que tiver essa condição será o limite à esquerda.

De forma semelhante, o limite da direita é obtido. A principal diferença é que as

colunas percorridas iniciam no fim da imagem e são decrementadas até a metade de sua largura. Nas iterações, é buscado, a coluna com menor quantidade de valores 1 para ser o valor desse limite.

Caso os limites obtidos sejam os próprios valores iniciais, coluna 2 para o limite da esquerda e a coluna 1030 para o limite da direita, então é feito um ajuste de $+10$ e -10 para esquerda e direita, respectivamente. A Figura 18c apresenta esses limites obtidos.

Com esses limites, são atribuídos zeros a todos os valores das colunas antes de limite à esquerda e depois do limite à direita. Dessa forma, a imagem de saída apresenta como região com valores 1 apenas a área interna da mandíbula. No entanto, observou-se que, em alguns casos, o processo de limiarização deixa áreas com valores 1 dentro da região externa, como é possível observar na Figura 18c. A fim de remover essas regiões, é feito um processo que busca o último pixel branco da região interna de cada coluna e zera todos os demais abaixo dele, retornando para a a imagem de saída somente a ROI da mandíbula. A Figura 18d apresenta as áreas excedentes aos limites removidas, bem como as regiões abaixo da mandíbula.

3.4 Segmentação das regiões iniciais

Após obtidas as ROI's das imagens, a próxima etapa consiste em selecionar, dentro da máscara, apenas regiões com maior semelhança à patologia buscada. Essa etapa é chamada de segmentação inicial. Seu objetivo é reduzir o custo computacional da localização do cisto, fornecendo para o algoritmo de IA apenas regiões com maior possibilidade de validação.

Para isso, faz-se necessário o uso de técnicas de segmentação, que buscam separar os pixels da imagem em grupos de semelhança. Essa seção apresenta a aplicação dos algoritmos de segmentação na área da mandíbula, buscando aquele com melhor capacidade de discriminação de cistos dentre as demais estruturas anatômicas. A descrição destes algoritmos é detalhada na seção 2.5.

Primeiramente, é feito uma operação *AND* entre a imagem pré-processada e a máscara obtida na etapa anterior. O resultado é utilizado para a aplicação dos algoritmos. As técnicas de segmentação utilizadas são: limiarização adaptativa, multinível de Otsu e *K-means*.

No limiar adaptativo, o tamanho da janela de vizinhança foi definido como

$$s = 2 \left\lfloor \frac{S}{16} \right\rfloor + 1, \quad (3.3)$$

em que, S correspondem as dimensões da imagem de entrada e $\lfloor \cdot \rfloor$ retorna o inteiro mais próximo

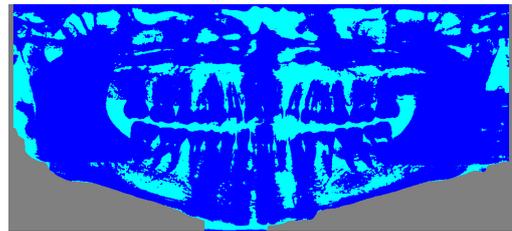
menor ou igual ao valor informado. O resultado dessa equação para as imagens do banco é uma vizinhança de 63 x 129 pixels.

O *K-means* utilizou distância quadrática euclidiana e inicialização dos centros de *clusters* com o algoritmo *k-means++*. Essa técnica de inicialização utiliza uma heurística para fornecer o local inicial dos *clusters* com convergência mais rápida. A Figura 19 apresenta imagens obtidas com esses métodos.

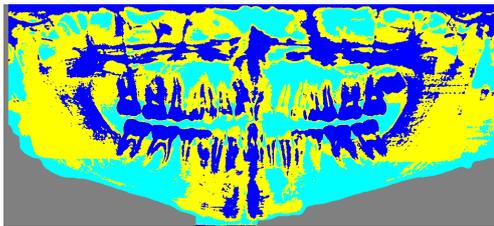
Figura 19 – Algoritmos de segmentação inicial.



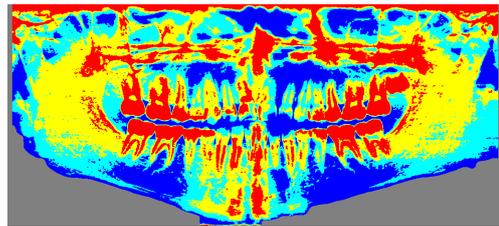
(a) Interseção entre a máscara e a imagem pré-processada



(b) Limiar adaptativo



(c) *K-means*



(d) Multínível de Otsu

Fonte: Imagens obtidas durante os experimentos.

Nessas imagens, cada aspecto de cor representa uma determinada região indicada pelo algoritmo. A Figura 19a apresenta a imagem obtida com a operação *AND* entre a imagem pré-processada e a máscara. A Figura 19b apresenta a imagem resultante da limiarização adaptativa, a Figura 19c mostra a saída obtida com o *K-means*, e a Figura 19d apresenta as regiões separadas com o multínível de Otsu. Nas imagens obtidas pelos algoritmos, as regiões em cinza representam as áreas externas removidas.

O parâmetro de sensibilidade do limiar adaptativo, o número de classes do *K-means* e a quantidade de limiares do multínível de Otsu foram estudados. A fim de ajustar a melhor configuração de parâmetros para os algoritmos, e dessa forma compará-los para selecionar o mais adequado para a presente metodologia. Em cada técnica, foi selecionada somente a região que apresentou a maior intercessão com os cistos pelo GT, calculada pela quantidade de pixels que intercedem. Na qual, são calculadas as porcentagens de área da mandíbula e cisto. O resultado

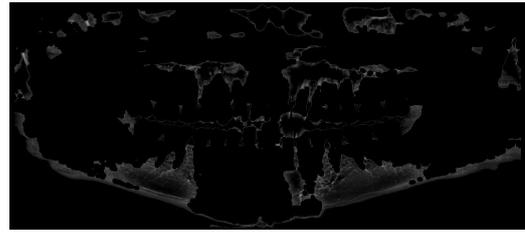
dessa análise será apresentado na seção 4.1.

O melhor resultado foi obtido com a segmentação do multinível de Otsu. Como essa técnica fornece diversos grupos de regiões nas imagens, é selecionado apenas o grupo que apresenta maior intercessão com o GT para todas as imagens. Dessa forma, somente um grupo do multinível de Otsu é empregado para formar a máscara com as regiões semelhantes a cistos dentro da ROI da mandíbula. A Figura 20 apresenta as regiões obtidas.

Figura 20 – Regiões selecionadas pelo multinível de Otsu.



(a) Máscara obtida



(b) Intercessão da máscara com a imagem pré-processada

Fonte: Imagens obtidas nos experimentos.

Na Figura 20a, apresenta-se a máscara obtida selecionando-se a classe escolhida do multinível de Otsu, e a Figura 20b apresenta a interseção dessa máscara com a imagem pré-processada. Esse processo é feito para todas as imagens, e as interseções são enviadas para a próxima etapa.

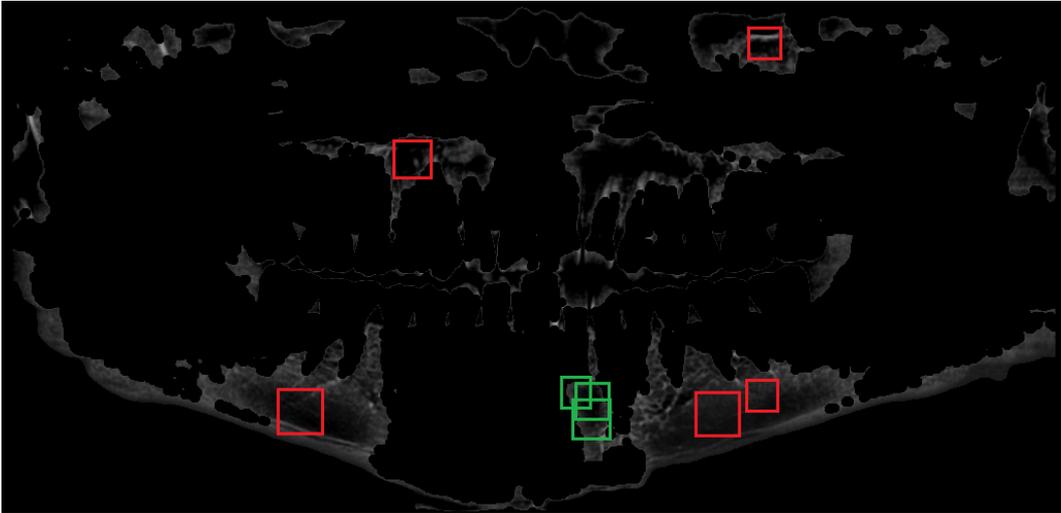
3.5 Extração de características e classificação

Nessa seção, são descritos os procedimentos para extração das texturas de Haralick e LBP, bem como a classificação destas por algoritmos de aprendizagem de máquina. Esses descritores precisam de um espaço de vizinhança para obtenção da textura. Dessa forma, é necessário um estudo preliminar sobre a influência do tamanho da janela amostral nos resultados para as diferentes texturas e classificadores.

A partir das regiões selecionadas pelo multinível de Otsu, é iniciada a extração de características. Esse processo consiste em fazer convoluções de janelas sobre a imagem. A métrica obtida pela janela representa a textura do pixel central desta. Assim, na medida que a janela percorre a imagem, de cada interseção do seu pixel central com uma região obtida pelo multinível de Otsu é extraída a textura. Então, é verificada se essa amostra pertence a uma

região de cisto. Em caso positivo, é rotulada como classe 1, do contrário como classe 0. Essas informações irão compor um arquivo no banco de características, em que cada linha representa uma amostra de textura obtida da imagem.

Figura 21 – Exemplo de extração dos descritores.



Fonte: Autor.

Esse processo é ilustrado na Figura 21. Nessa imagem, as janelas representadas em verde são as amostras utilizadas para descrever o cisto, e as em vermelho são para as demais regiões não císticas.

Os tamanhos de janela investigados são: 11x11, 15x15, 19x19, 23x23, 27x27, 31x31, 35x35, 39x39, 43x43, 47x47 e 51x51. Em cada imagem, são gerados 11 arquivos correspondentes as variações de janelas, para cada algoritmo de textura. Isso totaliza 22 bancos de características, sendo que cada banco representa as texturas de um determinado tipo de janela para todas imagens. Assim, são 704 arquivos de dados a serem aplicados na classificação.

Na extração da textura LBP, a janela é dividida em células, cada qual com seu histograma. Utiliza-se raio de seleção dos vizinhos com valor 1 e vizinhança-8. São extraídos 10 recursos LBP para cada amostra.

Na extração das texturas de Haralick, a janela é utilizada para cálculo da matriz de co-ocorrência de níveis de cinza, para que as estatísticas sejam calculadas. Sua formação utiliza pares de vizinhos com ângulo de 0° , sendo considerada a propriedade de simetria dos valores. Dentre as métricas calculadas, a correlação e a diferença de entropia apresentavam valor nulo para maior parte das amostras, sendo não utilizadas para formação dos bancos. Dessa forma, são obtidas 12 descritores de textura Haralick para cada amostra.

Esses bancos são utilizados para treino e classificação dos algoritmos de RNA,

ELM e SVM. Essas técnicas desenvolvem modelos com as amostras de treino e validam sua aprendizagem com amostras de teste.

A RNA utilizada é a *MLP* com 120 neurônios na camada oculta, função de ativação sigmóide, com 100 épocas e algoritmo de otimização *Levenberg-Marquardt*. A ELM é configurada com 500 neurônios na camada oculta e função de ativação sigmoide. Esses parâmetros foram obtidos através de testes exaustivos.

O SVM tem como *kernel* a função de base radial (RBF) e tamanho definido através de otimização bayesiana. Selecionou-se um banco aleatório de cada grupo e textura para otimizar o SVM, de modo que o parâmetro obtido é aplicado na classificação dos demais bancos do mesmo grupo e textura. Os tamanhos de *kernel* obtidos são apresentados na Tabela 1.

Tabela 1 – Tamanhos de *kernel* da SVM.

Grupo	Textura	kernel
12 imagens	Haralick	0,02336
	LBP	0,11595
32 imagens	Haralick	0,02574
	LBP	0,12976

Fonte: o autor.

Durante os testes de classificação, observou-se que o desbalanceamento dos bancos com relação às amostras prejudicava os resultados, pois a quantidade de exemplos não cíclicos é muito superior. Para eliminar o desbalanceamento, em cada imagem foi escolhido aleatoriamente amostras não cíclicas com quantidade equivalente à de cíclicas.

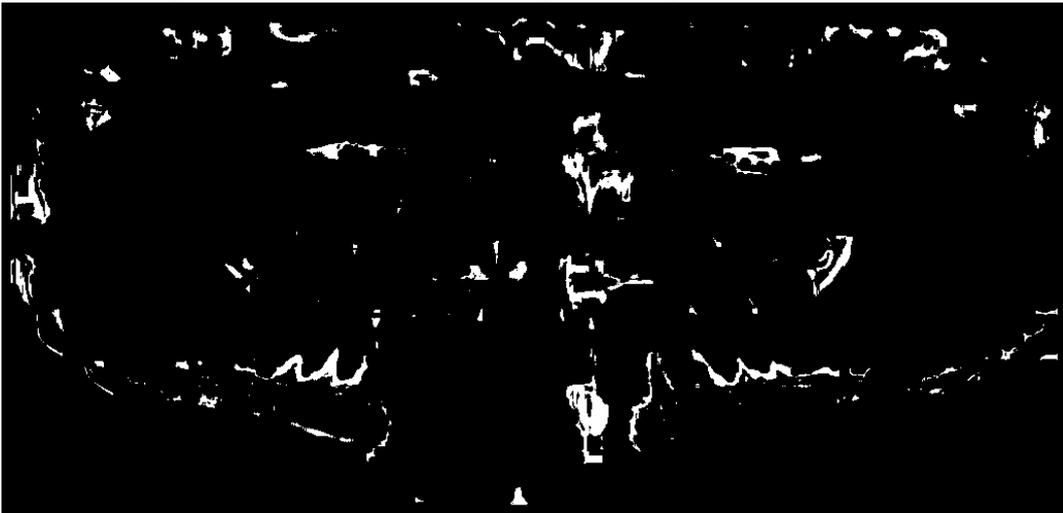
Os bancos, correspondentes a uma única variação de janela e o mesmo tipo de textura, são aplicados nos algoritmos de classificação. Esse processo é feito separando um arquivo do banco, correspondente às texturas de uma imagem, para ser utilizado como teste e os demais arquivos serem aplicados no treinamento. Essa operação é repetida alternando o arquivo de teste por outro, e continua até que todos os arquivos do banco sejam testados. Essa forma de validação cruzada é conhecida como *Leave-one-out*. Os resultados da classificação total do banco é a média das métricas obtidas com cada arquivo.

O processo descrito é realizado para cada variação de janela nos grupos de 12 e 32 imagens. São calculadas as métricas nos dois grupos, a cada combinação de janela, textura e classificador, de forma a avaliar seus resultados. A configuração com melhores métricas em cada grupo é utilizada para selecionar as regiões de cisto.

O algoritmo com melhor resultado nessa investigação é utilizado para classificar as regiões obtidas com o multinível de Otsu. Esse processo é feito através do *Leave-one-out*, treinando novamente o algoritmo com os arquivos através da melhor configuração obtida e testando diretamente na imagem correspondente ao arquivo separado para teste.

A textura de cada pixel da imagem que estiver na região selecionada pelo multinível de Otsu é enviado para o algoritmo já treinado. A decisão dele sobre a amostra é utilizada para preencher uma imagem de saída. Assim, caso seja positiva, o valor do pixel de mesma posição na imagem de saída é inserido para 1, indicando que é uma amostra de cisto, caso negativa, é inserido para zero, indicando uma amostra não cística. Esse processo continua modificando os arquivos de treino e testando na imagem corresponde ao arquivo separado para teste, até completar todas as imagens.

Figura 22 – Exemplo de resultado da classificação nas regiões do multinível de Otsu.



Fonte: Imagem obtida através dos experimentos.

Os resultados desse processo são imagens binárias, nas quais as regiões com valor 1 são indicações de cistos dadas pelo algoritmo de classificação. A Figura 22 apresenta uma saída obtida nessa etapa. Nas saídas obtidas pela classificação, é aplicado um filtro de mediana de 15x15 como pós-processamento, a fim de remover erros aleatórios ocasionados pelo algoritmo. Essas imagens são utilizadas na etapa de segmentação final do cisto.

3.6 Segmentação final do cisto

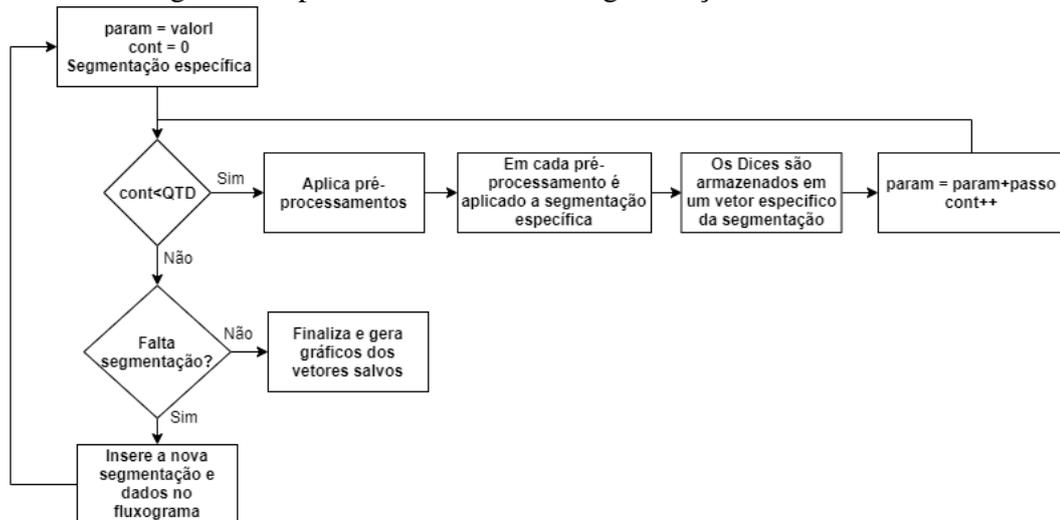
Essa etapa consiste em receber as marcações pós-processadas e realizar uma expansão das áreas obtidas, para que as regiões semelhantes e adjacentes com as indicações de cistos

sejam incorporadas como marcação. Existem diversos algoritmos para esse fim. Nessa pesquisa foram utilizados o CR, CA e FMM 2.5.

Para obter os parâmetros de cada algoritmo, é realizada uma investigação da variação desses em relação ao desempenho alcançado. Além disso, também são realizados testes da influência do pré-processamento nessas técnicas. As segmentações obtidas são comparadas com o GT para calcular o coeficiente de similaridade Dice, a fim de comparar e obter o melhor algoritmo de segmentação, parâmetro e pré-processamento a ser utilizado nessa etapa.

Os experimentos consistem em iniciar as segmentações a partir de um ponto dentro do cisto para verificar qual algoritmo consegue obter melhor definição da região de interesse. Esse processo é feito para cada técnica, variando seus parâmetros de sensibilidade, três vezes em cada imagem, sendo uma para cada tipo de pré-processamento. A Figura 23 apresenta o fluxograma desse procedimento.

Figura 23 – Fluxograma do processo utilizado na segmentação final.



Fonte: Autor.

Nesse fluxograma, *Segmentação específica* representa o algoritmo de segmentação que será utilizado nessa iteração, *param* é o valor do parâmetro de sensibilidade, *valorI* representa um valor inicial utilizado como primeira variação, *QTD* é a quantidade de variações de parâmetros e *passo* é o valor que incrementará a sensibilidade da segmentação.

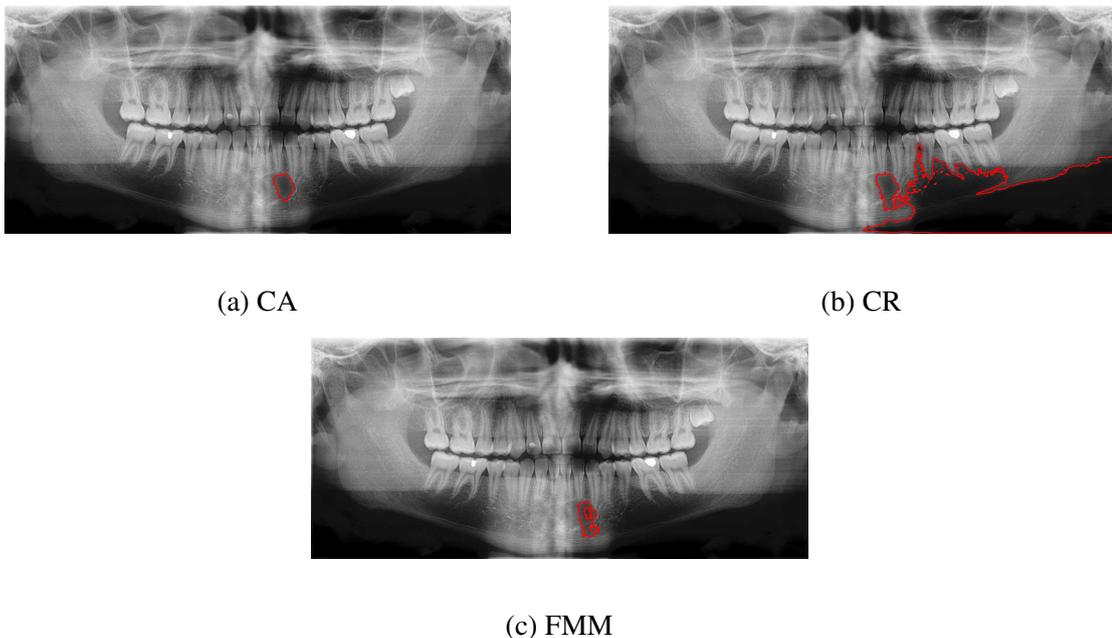
Os pré-processamentos utilizados são: filtro adaptativo, equalização de histograma e o filtro morfológico. Eles visam variar o contraste da imagem, e são aplicados a fim de que seja selecionado o que obtiver melhor resultado com as técnicas de segmentação. A Figura 24 apresenta segmentações obtidas por esses algoritmos com o mesmo ponto inicial.

Em cada variação de sensibilidade, são aplicados os algoritmos de pré-processamento

nas imagens. Os resultados destes são utilizados como entradas do algoritmo de segmentação específico, que inicia a partir de um ponto selecionado dentro do cisto. As médias de coeficientes Dice das imagens obtidas em cada pré-processamento são armazenadas em um vetor da segmentação, juntamente com o parâmetro utilizado.

Após isso, a sensibilidade da segmentação é incrementada pelo *passo* e o *cont* é incrementado. Caso ainda tenham variações a serem testadas ($cont < QTD$), o processo é refeito com a nova variação. Caso contrário, a interação com a segmentação específica é encerrada. Então é feita uma nova condição, que verifica se falta algum algoritmo de segmentação a ser testado. Se for positivo, o processo é reiniciado com a segmentação específica sendo o novo algoritmo a ser testado, e refazendo todas as iterações para ele. Caso todos já tenham sido testados, os dados salvos são utilizados para formar os gráficos das análises.

Figura 24 – Algoritmos de segmentação inicial.



Fonte: Imagens obtidas durante os experimentos.

Essas imagens ilustram o resultado das segmentações com relação ao exame original, mas foram aplicadas na imagem pré-processada com o filtro morfológico. A Figura 24a apresenta as regiões obtidas com o CA, a Figura 24b com o CR e Figura 24c com o FMM.

3.7 Aplicação da CNN

A CNN foi testada nessa pesquisa como uma alternativa ao processo clássico de VC. Essa técnica é aplicada a partir da imagem de entrada, sem a necessidade de segmentação

e extração de características específicas. Porém, faz-se necessário uma quantidade grande de amostras de radiografia, pois a CNN precisa de muitas amostras de treino. Nesse caso, cada imagem de um cisto é uma amostra para a CNN. Ademais, realizou-se uma técnica conhecida como aumento de dados (YONG *et al.*, 2019). Essa consiste em gerar novas amostras a partir de transformações das imagens originais. São utilizadas alterações de rotação, cisalhamento, *flip*, translações verticais e horizontais, e combinações destas para gerar um total de 10 modificações. Para o grupo de 12 imagens, são geradas 1284, e para o grupo de 32, são formadas 1280.

A arquitetura da CNN utilizada é a *U-net*, proposta por Ronneberger *et al.* (2015) para segmentação de imagens biomédicas de estruturas neuronais. Essa rede utiliza função de ativação ReLU, que é uma unidade linear retificadora que produz zero em metade de seu domínio. Sua aplicação zera todos os valores negativos de saída da camada anterior. Também é utilizada a técnica de *MaxPool*, um tipo de *pooling* que seleciona dentro da região somente o pixel de maior valor para seguir a próxima etapa (GOODFELLOW *et al.*, 2016).

A *U-net* consiste de duas etapas: uma de redução da amostragem e outra de expansão. A primeira apresenta repetição de 2 camadas de convolução 3x3 seguidas de uma ReLU e uma operação de *MaxPool* de 2x2 com passo 2. Isso é feito para cada bloco, dobrando o número de características dos canais.

Na etapa de aumento da amostragem, é feito uma convolução 2x2, uma concatenação com os dados retirados do mapa de características da etapa de redução da amostragem, duas convoluções 3x3, seguidas de uma ReLU, e uma *UpSampling* 2x2. Essa é uma técnica utilizada para aumentar o tamanho de amostragem (RONNEBERGER *et al.*, 2015). Isso é feito para cada bloco, que resulta na redução pela metade dos canais de características. A última camada utiliza uma convolução 1x1.

Após cada camada de convolução dos blocos, é feito uma operação de normalização em lotes, com tamanho 3. Apresentando essas características descritas, a arquitetura possui um total de 23 camadas.

Como são utilizados cinco camadas *MaxPool* com redução da resolução espacial por um fator de 2 em cada, então a imagem de entrada deve ser divisível por 32. Dessa forma, as imagens são reduzidas para um tamanho de 256x256, reduzindo portanto o custo computacional.

No modelo dessa arquitetura, é utilizada a função de perda que combina entropia binária cruzada e a perda de dados medida pelo coeficiente Dice. Essa função é utilizado pela rede para ajuste dos parâmetros no treinamento. Para acompanhar e medir o desempenho, é

utilizada a métrica Dice juntamente com sua perda.

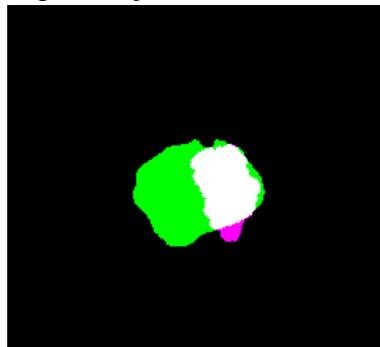
O algoritmo de otimização utilizado é a estimativa de momento adaptativo, conhecido como *Adam*, aplicado com 200 épocas e parada antecipada com paciência definida para 10, ou seja, caso não ocorra melhora de perdas nessa quantidade de épocas, o treinamento é interrompido.

A abordagem utilizada para validação cruzada da CNN é metodologia *Holdout*. O banco das imagens geradas é dividido em 3 grupos aleatórios, com 70% para treino, 10% para validação e 20% para teste. A média das métricas obtidas na aplicação do grupo de teste é o resultado da CNN. Esse procedimento também é testado em um modelo híbrido que propõe uma união da metodologia e a CNN, utilizando as etapas 1, 2, 3, 4, 7 e 8 (Figura 12). Nesse modelo, a rede é aplicada a partir das regiões selecionadas na etapa de segmentação inicial. Esses dois procedimentos são realizados para os bancos de 12 e 32 imagens.

3.8 Cálculo das métricas

As métricas utilizadas para avaliar a metodologia são calculadas comparando-se os resultados obtidos com os experimentos e o GT. As primitivas de VP, FP, VN e FN são encontradas através dos pixels da imagem resultante das abordagens. Dessa forma, cada pixel pertence a um desses grupos, e a quantidade de pixels de cada grupo presente em uma dado exame é utilizada para cálculo das métricas descritas na seção 2.9, representando a avaliação dessa imagem.

Figura 25 – Relação entre GT e segmentação.



Fonte: Autor.

A Figura 25 ilustra a relação entre uma determinada segmentação e o GT. Nessa figura, cada cor representa um grupo das primitivas. A cor branca são os VP's, verde são os FN's, ciano são os FP's e preto são os VN's. Observa-se que os VP's são os pixels pertencente a área

de intercessão entre as duas marcações, pois foram rotulados como cisto pela metodologia e pelo GT. FN's são os pixels que pertencem ao GT mas não foram rotulados pela metodologia. FP's são os pixels que foram rotulados como cistos, mas não constam no GT. E por último, VN's são os que não são cistos em nenhuma das duas imagens.

3.9 Considerações Finais

Nesse capítulo apresentou-se o processo utilizado para a detecção automática de cistos, explicando as etapas intermediárias da aplicação, bem como as abordagens e parâmetros utilizados nas técnicas. Percebe-se que a metodologia apresenta diversas investigações intermediárias e os resultados finais de detecção. Além disso, vale ressaltar que foi apresentado o sistema desenvolvido para realizar as marcações e algoritmo proposto para obtenção da região da mandíbula, através de limiarização iterativa.

No próximo capítulo, serão apresentados os resultados obtidos através dessa metodologia. Estes são separados em dois grupos: resultados intermediários e resultados finais da metodologia. O primeiro grupo busca as melhores técnicas e parâmetros para os métodos utilizados, e o segundo são os resultados da metodologia completa, utilizando os melhores métodos encontrados pelo primeiro grupo.

4 RESULTADOS

Este capítulo apresenta os resultados obtidos nos experimentos realizados a partir da metodologia especificada no Capítulo 3. Os métodos são aplicados na base coletada para os grupos de 12 e 32 imagens, com exceção da CNN, foram executadas em um notebook com processador Intel core i7 e 8 Gb de memória RAM. A CNN foi executada no Google Colab, um serviço de nuvem gratuito que conta com GPU's Tesla K80 e 13GB de RAM.

Em etapas intermediárias da metodologia, são feitos experimentos de técnicas computacionais, buscando selecionar a que apresente melhor métrica para ser incorporada à metodologia final. As discussões dessas investigações serão realizada nesse capítulo. Inicialmente, os resultados das etapas intermediárias serão discutidos, tais como: segmentação inicial, classificação de texturas e segmentação final. Em seguida, são apresentados os resultados obtidos com a metodologia proposta completa: utilizando textura, através da CNN diretamente nas imagens e, por último, o modelo híbrido proposto.

4.1 Resultados da segmentação inicial

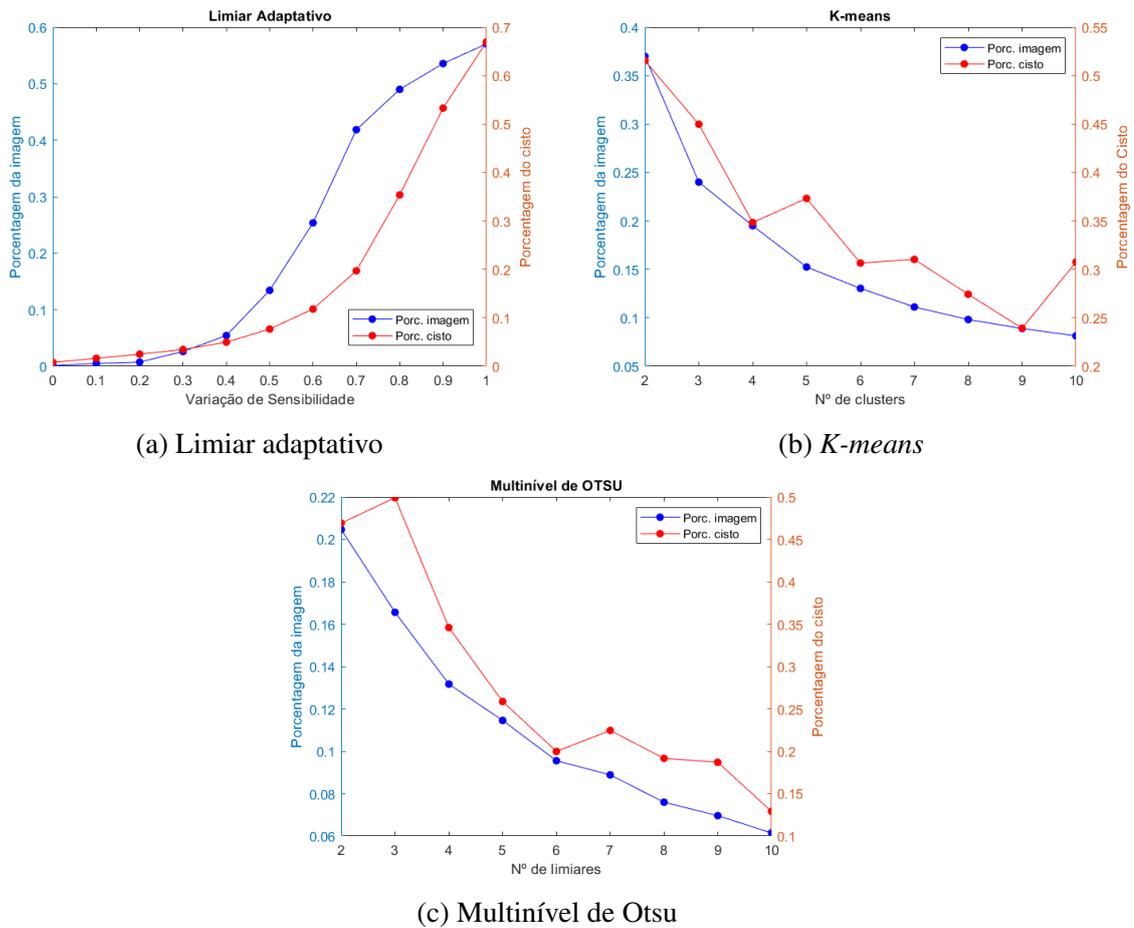
A segmentação inicial consiste no teste de técnicas de segmentação dentro da região da mandíbula. Esse estudo é feito através da variação de parâmetros, buscando o melhor algoritmo e configuração de parâmetro para essa etapa.

As métricas utilizadas para isso são as porcentagens de área do cisto e da maxila, avaliando para cada variação de parâmetro as porcentagens obtidas pelo algoritmo dentro da ROI. No limiar adaptativo é feito uma variação de sensibilidade de 0,1 entre 0 e 1, no *K-means* é feito um variação de 1 *cluster* entre 2 e 10 e no multinível de Otsu é feito uma variação de 1 limiar entre 2 e 10.

Nos gráficos obtidos (Figura 26), o eixo da esquerda com o gráfico em azul, representa a porcentagem de área da mandíbula que o algoritmo seleciona e o eixo da direita, com o gráfico em vermelho, a porcentagem de área do cisto. O objetivo desse estudo é selecionar o algoritmo e seus parâmetros que reduza ao máximo a área da mandíbula sem recair a do cisto. Os resultados obtidos nessa etapa, com os três algoritmos, foram para o grupo de 32 imagens.

Na Figura 26a, temos os resultados para o limiar adaptativo. Observa-se que o aumento da sensibilidade cresce as áreas da mandíbula e do cisto selecionadas pelo algoritmo, sendo o ótimo alcançado para sensibilidade 1.

Figura 26 – Resultados dos algoritmos de segmentação inicial para o grupo de 32 imagens.



Fonte: Autor

A Figura 26b apresenta os resultados obtidos para o *K-means*. Nesse gráfico, as variações de área do cisto e da mandíbula apresentam comportamento diferente. A área do cisto apresenta alternância de incremento e decremento, com máxima em 2 *clusters*. Analisando o comportamento da área da mandíbula, verifica-se que esta decai com o aumento dos *clusters*.

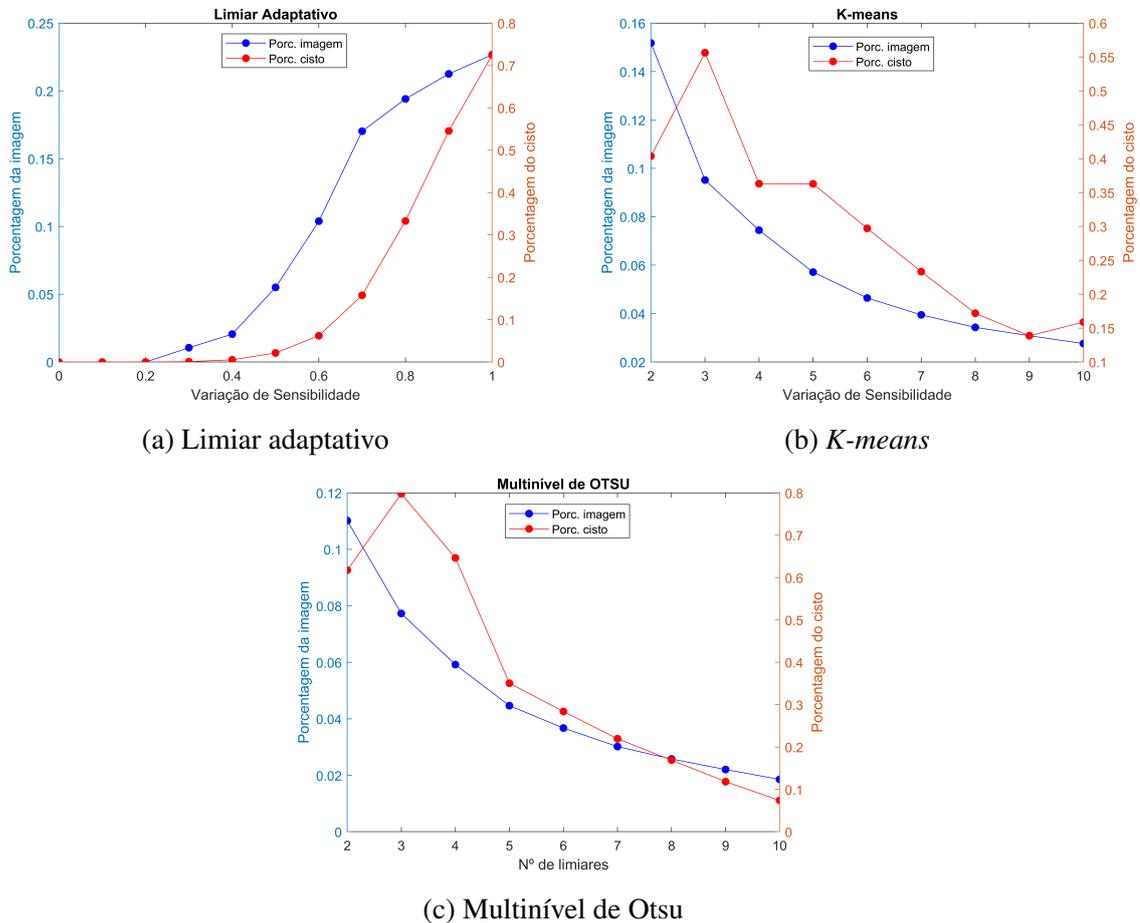
Os resultados com o multinível de Otsu (Figura 26c) apresenta valor máximo com 3 limiares e após isso regride, enquanto a área da mandíbula decresce sucessivamente com o aumento dos limiares.

Analisando esses resultados percebe-se que o limiar adaptativo apresenta máxima área de cisto próxima dos 70%. No entanto, para essa mesma sensibilidade, também apresenta alta taxa de regiões da mandíbula, com valor de 55%. O *K-means* possui maior taxa de área do cisto em torno de 52% e para a mesma quantidade de *clusters* apresenta taxa da mandíbula de 37%. O multinível de Otsu possui taxa de área do cisto máxima de 50% e, para a mesma quantidade de limiares, possui área da mandíbula de 16,5%.

Dessa forma, o algoritmo que melhor seleciona as regiões internas da mandíbula no

grupo de 32 imagens é o multinível de Otsu. Os resultados do mesmo experimento aplicado no grupo de 12 imagens é apresentado na Figura 27.

Figura 27 – Resultados dos algoritmos de segmentação inicial para o grupo de 12 imagens.



Fonte: Autor

Os resultados do limiar adaptativo (Figura 27a), tem características semelhantes ao comportamento observado para o grupo de 32 imagens, sendo a melhor sensibilidade obtida igual a 1.

A Figura 27b apresenta os resultados do *K-means*. Observa-se que a taxa de área do cisto atinge valor máximo para 3 *clusters*. Já a taxa da área da mandíbula decai na medida que o número de *clusters* aumenta.

A Figura 27c apresenta os resultados para o multinível de Otsu, que possui comportamento semelhante ao observado no grupo de 32 imagens. O melhor resultado é para o número de limiares igual a 3.

O limiar adaptativo foi o método com a maior taxa de cisto e da mandíbula (71%). Já o *K-means* apresenta taxa máxima de área cística de 55% e, para a mesma quantidade de *clusters*,

área da mandíbula de 9,5%. O multinível de Otsu possui maior taxa de área cística (80%) e, para mesmo limiar, a área da mandíbula de 8,5%.

Assim, o algoritmo que melhor contempla os requisitos buscados para o grupo de 12 imagens também é o multinível de Otsu. Portanto, para ambos os grupos é selecionada essa técnica com 3 limiares para a etapa de segmentação das regiões internas a mandíbula.

Comparando os dois grupos, percebe-se que os experimentos no grupo de 12 imagens apresentam maior porcentagem do cisto na maioria dos algoritmos, com exceção do *K-means* que obteve porcentagem inferior ao grupo de 32 imagens. Isso significa que a maioria das técnicas de segmentação avaliadas não conseguem distinguir bem as regiões císticas em imagens com baixa qualidade, incorporando parte significativa da área patológica a outro grupo de região.

4.2 Resultados de classificação por textura

Os experimentos de classificação por textura visam selecionar a combinação de tipo de textura, tamanho da janela amostral e algoritmo de classificação que apresentem melhor desempenho. Foram realizados os experimentos com três técnicas de aprendizado de máquinas, para 11 variações de janelas, com dois descritores de textura.

O critério de desempenho principal foi a acurácia, que representa a proporção total de acertos, e a técnica de validação utilizada foi a *Leave-one-out*. Nas tabelas apresentadas no decorrer dessa seção, os valores destacados em negrito representam as maiores taxas obtidas na métrica para cada classificador. A Tabela 2 apresenta os resultados obtidos, com o grupo de 32 imagens, para a textura Haralick.

Tabela 2 – Resultados de classificação com o Haralick para grupo de 32 imagens.

Janelas	RNA			SVM			ELM		
	Ac(%)	Se(%)	Es(%)	Ac(%)	Se(%)	Es(%)	Ac(%)	Se(%)	Es(%)
11x11	91,48	90,81	92,15	90,82	90,35	91,28	91,21	92,19	90,22
15x15	92,31	92,11	92,50	92,42	93,49	91,36	92,58	94,05	91,11
19x19	93,48	92,52	94,45	92,41	95,02	89,81	93,97	95,68	92,25
23x23	92,54	93,04	92,04	92,85	96,15	89,54	93,93	95,50	92,35
27x27	92,06	92,85	91,28	92,68	96,23	89,13	92,64	93,50	91,79
31x31	91,39	93,12	89,65	91,77	96,62	86,92	92,91	95,00	90,83
35x35	91,34	93,21	89,47	90,61	95,30	85,93	92,52	95,00	90,05
39x39	91,61	94,63	88,60	90,34	94,06	86,62	93,77	96,12	91,42
43x43	89,12	94,41	83,82	89,62	94,35	84,90	93,14	95,63	90,65
47x47	90,75	94,43	87,08	89,02	94,25	83,80	92,85	95,37	90,33
51x51	87,04	92,46	81,62	89,05	93,87	84,24	93,01	95,88	90,14

Fonte: o autor.

Nessa tabela são apresentados os dados resultantes da comparação entre os diferentes classificadores ao longo da variação de janelas amostrais. Observando o comportamento dos classificadores, é possível observar que o aumento do tamanho de janela melhora os resultados até determinado valor e tamanhos maiores que esse limite tendem a prejudicar as métricas.

Analisando as melhores acurácias, nota-se que a RNA obteve 93,48% para janela de 19x19, o SVM tem máxima de 92,85% para janela de 23x23 e a ELM obteve métrica superior de 93,97% com janela de 19x19. Logo, o classificador com a maior proporção de acertos das amostras nesse experimento é o ELM. Os resultados dos experimentos realizados com a textura LBP no banco de 32 imagens é apresentado na Tabela 3.

Tabela 3 – Resultados de classificação com o LBP para grupo de 32 imagens.

Janelas	RNA			SVM			ELM		
	Ac(%)	Se(%)	Es(%)	Ac(%)	Se(%)	Es(%)	Ac(%)	Se(%)	Es(%)
11x11	70,58	68,99	72,18	68,43	63,79	73,07	70,52	73,84	67,20
15x15	71,45	77,32	65,58	71,37	62,93	79,80	71,69	69,09	74,29
19x19	70,72	81,14	60,29	72,83	63,27	82,39	74,17	70,47	77,87
23x23	75,13	84,95	65,31	75,43	65,45	85,41	76,22	71,21	81,23
27x27	72,99	86,91	59,07	78,32	70,27	86,36	76,69	70,40	82,99
31x31	74,70	90,31	59,09	79,84	71,77	87,92	76,17	68,09	84,25
35x35	73,38	90,93	55,84	79,89	71,39	88,40	77,04	68,29	85,80
39x39	73,66	92,04	55,29	79,11	70,05	88,18	76,70	67,12	86,29
43x43	74,95	93,21	56,68	80,12	71,18	89,05	78,18	69,17	87,19
47x47	77,48	93,56	61,40	81,78	74,02	89,54	78,51	69,67	87,35
51x51	75,03	92,83	57,24	81,11	73,14	89,08	77,30	66,41	88,19

Fonte: o autor.

Para esse descritor, verifica-se que o aumento da área amostral tem influencia positiva, ou seja, o máximo desempenho de classificação é alcançado para as maiores dimensões de janela.

Logo, as maiores acurácias obtidas pelo algoritmo são de, 77,48% para a RNA, 81,78% para a SVM e 78,51% para a ELM, todos com a janela de 47x47, sendo o SVM o algoritmo que apresenta melhor resultado de classificação para essa textura.

Examinando os dados das tabelas 2 e 3, observa-se que em termos de sensibilidade, o melhor resultado é do descritor Haralick (96,62%) com a SVM e janela de 31x31, demonstrando ser essa combinação a que possui melhor capacidade em rotular corretamente uma amostra cística. A maior especificidade também é dessa textura (94,45%) com a RNA e janela de 19x19, demonstrando essa combinação como a que indica bem amostras não císticas.

Em geral, a textura Haralick possui maiores taxas de máximos em todas as métricas calculadas. Entre os dois descritores, percebe-se que o de Haralick atinge valor máximo de acurácia com janela menor, já o de LBP perdura para as ultimas variações de janela, deixando a

possibilidade de taxas superiores com dimensões maiores que as experimentadas.

A Tabela 4 apresenta os resultados obtidos pelos classificadores com o grupo de 12 imagens para a textura de Haralick

Tabela 4 – Resultados de classificação com o Haralick para grupo de 12 imagens.

Janelas	RNA			SVM			ELM		
	Ac(%)	Se(%)	Es(%)	Ac(%)	Se(%)	Es(%)	Ac(%)	Se(%)	Es(%)
11x11	89,15	86,88	91,41	85,88	88,33	83,43	88,94	89,16	88,72
15x15	91,90	91,32	92,47	88,46	90,39	86,52	90,65	89,96	91,35
19x19	92,85	93,50	92,20	91,89	94,26	89,52	93,34	92,85	93,82
23x23	96,17	95,06	97,28	93,92	96,14	91,71	94,59	93,77	95,41
27x27	95,11	95,21	95,00	93,35	95,80	90,89	94,99	94,90	95,09
31x31	95,25	95,72	94,78	93,23	96,18	90,28	94,79	94,14	95,45
35x35	94,69	95,73	93,64	93,25	96,10	90,39	95,56	96,50	94,62
39x39	92,43	96,59	88,27	92,78	93,78	91,77	96,01	96,68	95,35
43x43	92,89	96,07	89,72	91,65	91,62	91,68	95,40	95,84	94,96
47x47	91,95	95,14	88,76	90,65	89,94	91,35	94,12	93,15	95,09
51x51	90,39	94,22	86,56	89,83	88,40	91,25	89,28	84,10	94,46

Fonte: o autor.

Através dessa tabela, é possível perceber um comportamento semelhante ao observado na Tabela 2, em que os classificadores atingem valores máximos em uma determinada dimensão amostral e tendem a reduzir para tamanhos maiores.

Comparando os classificadores, observa-se a maior taxa para a RNA (96,17%) e para a SVM (93,92%), ambos com janela de 23x23, e para a ELM 96,01%, com janela de 39x39. Logo, a melhor técnica de classificação é a RNA.

Tabela 5 – Resultados de classificação com o LBP para grupo de 12 imagens.

Janelas	RNA			SVM			ELM		
	Ac(%)	Se(%)	Es(%)	Ac(%)	Se(%)	Es(%)	Ac(%)	Se(%)	Es(%)
11x11	66,38	57,51	75,24	62,26	45,89	78,62	63,52	55,44	71,59
15x15	68,97	67,01	70,94	64,59	45,51	83,67	64,36	52,47	76,25
19x19	71,11	72,13	70,10	64,02	41,55	86,49	64,16	47,52	80,80
23x23	74,39	77,48	71,31	62,98	36,43	89,53	64,37	45,98	82,75
27x27	71,97	77,63	66,31	64,54	37,78	91,30	64,79	44,40	85,18
31x31	74,91	80,20	69,61	67,05	42,23	91,87	65,64	45,48	85,80
35x35	75,89	81,17	70,61	68,67	45,30	92,03	64,55	41,97	87,14
39x39	75,63	83,00	68,25	71,86	51,00	92,73	65,24	42,32	88,17
43x43	78,25	81,97	74,54	71,73	50,22	93,23	65,83	44,58	87,08
47x47	77,79	87,09	68,49	73,35	52,01	94,69	67,02	47,07	86,98
51x51	76,84	86,71	66,97	76,53	58,65	94,42	67,90	46,65	89,14

Fonte: o autor.

A Tabela 5 apresenta os resultados dos testes aplicados com a textura LBP. Esses resultados são coerentes aos obtidos para o grupo de 32 imagens. A maioria das taxas apresentam

crescimento ao longo das variações de janela, com exceção da especificidade da RNA e da sensibilidade da ELM, que manifestam decaimento. Cabe observar que, as maiores taxas são encontradas para altas dimensões de janelas, assim como observado na Tabela 3.

Comparando os classificadores, verifica-se que os maiores índices de acurácia são, de 78,25% para a RNA (janela de 43x43), 76,53% para a ELM e 67,9% para o SVM (ambos com janela de 51x51). Dessa forma, o classificador com maior proporção de acertos nesse experimento é a RNA.

Analisando os dados das Tabelas 4 e 5, percebe-se que os três classificadores com o descritor de Haralick apresentaram sensibilidade elevada, mas a maior obtida é de 96,68% com a ELM (janela de 39x39). Em relação a especificidade, o mesmo descritor apresenta a maior taxa, de 97,28% para a RNA (janela de 23x23).

No geral, a textura Haralick demonstra resultados superiores ao LBP em todas as métricas analisadas. Percebe-se que o primeiro apresenta convergência de taxa máxima das métricas dentro do intervalo de janelas investigado, enquanto o último tende a apresentar taxas superiores para janelas maiores.

Percebe-se também que ambos os grupos de 12 e 32 imagens possuem comportamento semelhante para uma mesma textura ao longo das variações de janela. O grupo de 12 imagens com textura Haralick apresenta a maioria das taxas superiores às do grupo de 32, com exceção da sensibilidade da SVM. Com a textura LBP no grupo de 12 imagens, a maioria das métricas tem valores inferiores aos do grupo de 32, com exceção da especificidade nos três classificadores.

A diferença entre a maior acurácia do grupo de 12 e de 32 imagens, com o descritor de Haralick, é de 2,2%, indicando que o grupo com 12 imagens possui melhor resultado. Essa diferença com o LBP é de -3,58%, indicando que o grupo de 12 imagens apresentou resultados menores que o de 32. Isso demonstra que remover os exames de má qualidade melhora o desempenho da classificação com Haralick, mas piora com LBP, significando que algumas imagens removidas são bem representadas pelo último descritor.

Os resultados demonstram que as duas texturas apresentam diferenças significativas na capacidade de descrever as estruturas de acordo com a variação do tamanho de janela. O descritor de Haralick descreve melhor as estruturas classificadas para janelas de dimensões menores, enquanto o LBP desempenha melhor em espaços amostrais maiores.

Vale ressaltar que resultados melhores poderiam ser alcançados para outras variações

de janelas, no entanto, tamanhos maiores tornam o custo computacional inviável. As estatísticas de Haralick possuem mais sensibilidade em discriminar as regiões císticas com matrizes menores, e os histogramas LBP precisam de mais dados para representar as estruturas classificadas adequadamente. Vale ressaltar ainda que o método de Haralick apresenta mudanças abruptas das taxas com a variação de janelas, enquanto o LBP tem alterações mais amenas.

Em suma, a textura de Haralick apresenta resultados superiores à LBP na aplicação testada, com as melhores taxas de classificação nos dois grupos. No grupo de 32 imagens, os melhores resultados são com a ELM e janela de 19x19, e no grupo de 12 imagens, com a RNA e janela de 23x23.

4.3 Resultados de segmentação final

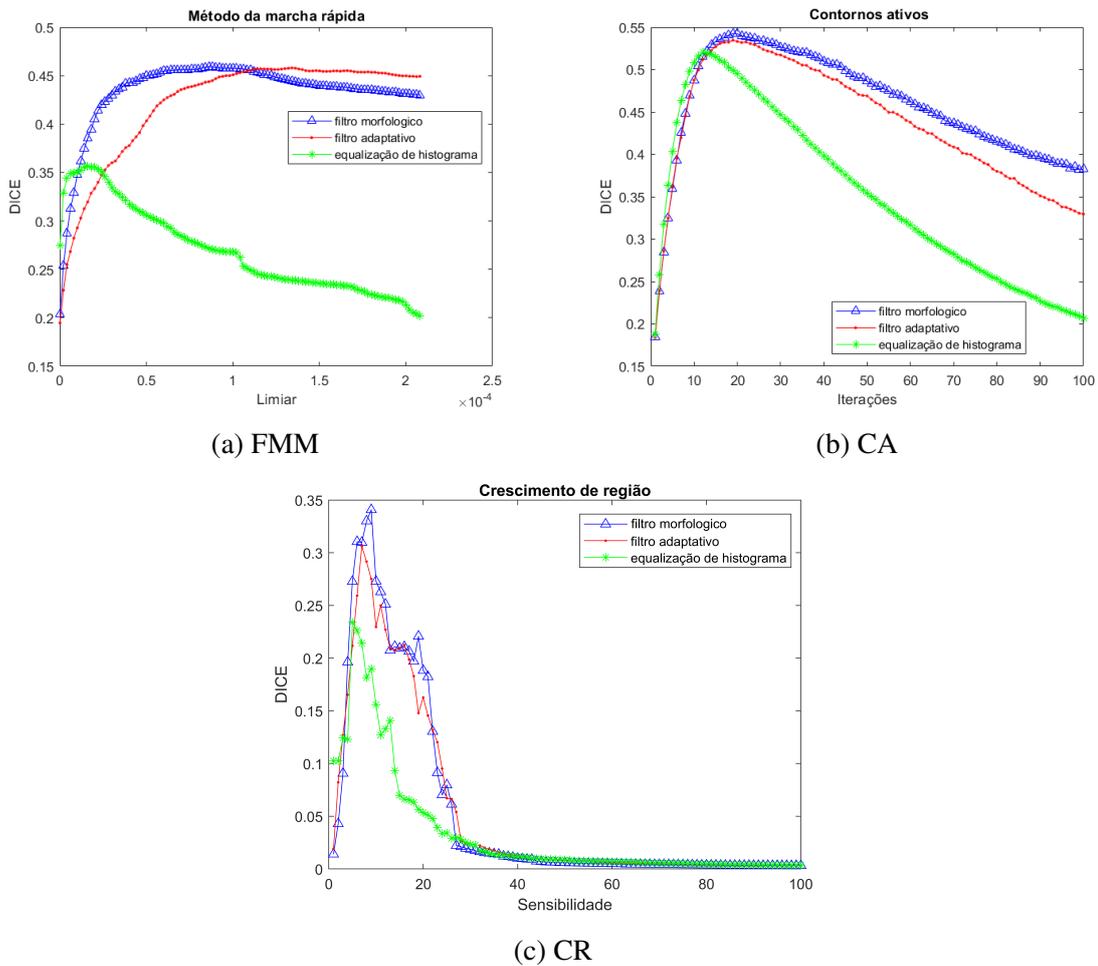
A segmentação final é a etapa que busca extrair a região cística. Para isso é preciso selecionar a combinação de algoritmos, parâmetros e pré-processamento que obtenha melhor desempenho na segmentação do cisto. As técnicas avaliadas nos experimentos foram: FMM, CR e CA. Os três tipos de pré-processamento analisados foram: filtro morfológico, equalização de histograma e filtro adaptativo.

Os algoritmos são aplicados nos exames a partir de um ponto dentro de cada cisto, sendo avaliados através do coeficiente Dice. Esse processo é feito variando-se os parâmetros e o método de pré-processamento. Os resultados para o grupo de 32 imagens são apresentados na Figura 28.

A Figura 28a apresenta os resultados obtidos com o FMM. É possível perceber que as técnicas de pré-processamento exercem influências diferentes para uma mesma variação de limiar. Com o aumento de limiar, os resultados obtidos com os filtros morfológico e adaptativo apresentam crescimento até atingirem um valor máximo. Após esse ponto, o primeiro reduz levemente e o segundo mantém valores aproximadamente constantes. A equalização de histograma também demonstra crescimento até um determinado limiar e reduz significativamente para valores maiores.

A Figura 28b demonstra os resultados obtidos com o CA. Nesse gráfico, os valores resultantes para as três técnicas possuem comportamento semelhante, atingem um máximo e reduzem consideravelmente. Os maiores coeficientes Dice obtidos são próximos, em taxa e número de iterações, e o decaimento varia de acordo com a técnica, sendo o mais rápido para a equalização de histograma.

Figura 28 – Resultados dos algoritmos de segmentação final para o grupo de 32 imagens.



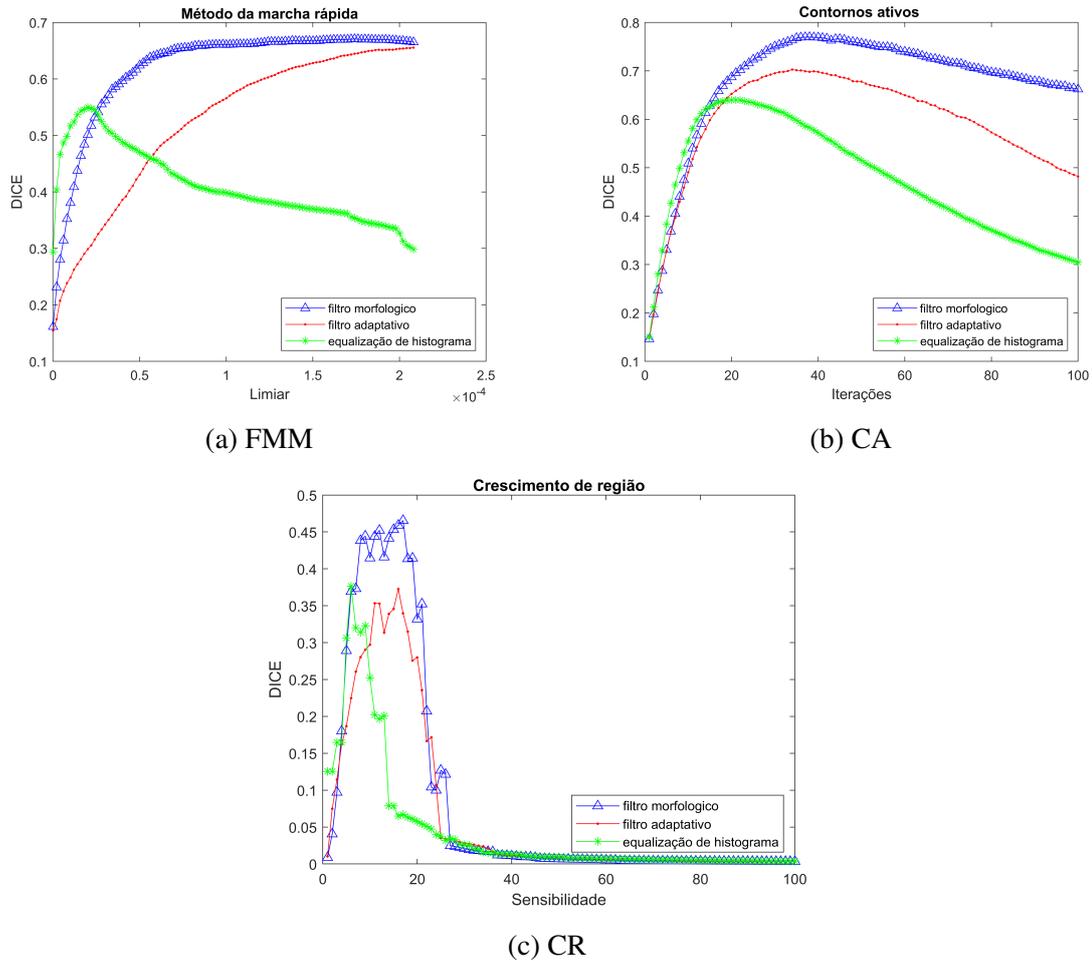
Fonte: Autor

A Figura 28c expressa os resultados obtidos com o CR. Observa-se que os dados resultantes apresentam variações abruptas, diferente do comportamento dos gráficos apresentados anteriormente. Para os valores de sensibilidade iniciais, os resultados variam alternadamente até atingir um máximo, quando então reduzem rapidamente chegando em 0. Os três pré-processamentos demonstram esse comportamento, diferenciando nas máximas taxas que atingem e na velocidade que reduzem.

O FMM alcança a maior taxa Dice, de aproximadamente 0,46, com os filtros adaptativo e morfológico, para os limiares de aproximadamente $0,58 \cdot 10^{-4}$ e $1,4 \cdot 10^{-4}$. O CA obteve maior métrica de 0,54 com o filtro morfológico e 21 iterações. O CR alcançou maior taxa com valor de aproximadamente 0,34 para sensibilidade próxima de 10.

Logo, a técnica que consegue melhor segmentação da região cística, no grupo de 32 imagens, é o CA, com filtro morfológico e 21 iterações. Os resultados obtidos para o grupo de 12 imagens é apresentado na Figura 29.

Figura 29 – Resultados dos algoritmos de segmentação final para o grupo de 12 imagens.



Fonte: Autor

A Figura 29a apresenta os resultados obtidos para o FMM. É possível observar que o resultado é similar ao FMM com o grupo de 32 (Figura 28a). No entanto, as variações alcançaram métricas maiores e o filtro adaptativo exibe comportamento modificado, demorando mais variações de limiares para atingir valor máximo. Vale ressaltar também que os filtros morfológico e adaptativo não demonstram tendências de redução para valores maiores que os testados.

A Figura 29b apresenta os resultados obtidos com o CA. Estes resultados possuem variações semelhantes às obtidas no grupo anterior. As diferenças observadas são: o aumento do Dice máximo obtido e a maior diferença entre os maiores valores de taxa para os pré-processamentos.

A Figura 29c demonstra que os resultados encontrados com o CR são semelhantes aos do grupo anterior, alterando os valores máximos de taxa e diferenciando melhor os resultados dos pré-processamentos. Observa-se também que os dados do filtro adaptativo demonstram mesmo

comportamento de valores ao do morfológico, no entanto com taxas menores.

Para o grupo de 12 imagens, o filtro morfológico consegue maior métrica para os três segmentadores. O FMM pontua maior Dice com valor de 0,68 e limiar de aproximadamente $1,7 \cdot 10^{-4}$. O CA possui maior taxa próxima de 0,78 com 37 iterações e o CR com máxima métrica de aproximadamente 0,47 e sensibilidade de 17.

Dessa forma, o segmentador que registra maior coeficiente Dice para o grupo de 12 imagens é o CA, com filtro morfológico e quantidade de iterações de 37. Portanto, para os dois grupos é selecionado o CA com o filtro morfológico na etapa de segmentação final, alterando apenas a quantidade de iterações.

Analisando a influência dos pré-processamentos nos dois grupos, nota-se que as melhores métricas são obtidas pela técnica de filtro morfológico, seguindo do adaptativo e da equalização de histograma. Este último tem a característica de modificar a distribuição dos valores do histograma reduzindo as diferenças de níveis de cinza entre cisto e mandíbula. Tal circunstância leva as técnicas de segmentação a extrapolar a região buscada.

Os filtros morfológico e adaptativo obtiveram melhores taxas devido a suas características intrínsecas de mudança de contraste, que deixaram o cisto mais escuro que a região da mandíbula, auxiliando na segmentação.

A combinação CA e filtro morfológico é interessante para a aplicação, pois o filtro modifica o contraste e torna as regiões do cisto mais escuras do que a sua vizinhança, deixando as bordas de transição entre as regiões bem definidas. Esse cenário favorece o CA alcançar as fronteiras dos cistos.

Um problema dessa combinação é no processamento de exames com duas regiões com tons levemente claros e escuros dentro da mesma área cística. O filtro resulta em maior ênfase à área escura, ficando a parte clara semelhante à mandíbula. Logo, o CA segmentaria somente a parte escura do cisto.

Os dois grupos apresentados nas Figuras 28 e 29 mostram que as imagens com menor qualidade influenciam nos resultados de segmentação final. Também é possível observar que, para o grupo de 12 imagens, as métricas obtidas são superiores em todos os experimentos realizados.

4.4 Resultados finais da metodologia

Nessa subseção são apresentados os resultados obtidos com as metodologias completas, aplicadas nos exames originais para adquirir as máscaras binárias de saída. As imagens obtidas com essa aplicação são comparadas ao *GT* para cálculo das métricas de avaliação. Essas são definidas por pixel, através de comparação com a região demarcada pelos especialistas.

A metodologia empregada, que utiliza classificação por textura, será chamada de MT-T. Esta representa o modelo clássico de VC, sendo construída com os melhores algoritmos encontrados nas seções anteriores. A aplicação dessa aos grupos de 32 e 12 imagens será denominada de MT-T32 e MT-T12.

Analogicamente, a aplicação da CNN nas imagens de entrada será chamada de MT-CNN. Os experimentos nos dois grupos serão denominados de MT-CNN32 e MT-CNN12, que representam a aplicação de um modelo com aprendizado profundo. Outro experimento realizado é um híbrido com a CNN partindo das imagens segmentadas. Esse será denominado de MT-CNNH e os testes nos grupos serão denominados de MT-CNNH32 e MT-CNNH12. Nesses experimentos as entradas da CNN são as imagens obtidas após a etapa de segmentação inicial. As médias de acurácia, sensibilidade e especificidade obtidas são apresentadas na Tabela 6.

Tabela 6 – Médias de Ac, Se e Es das metodologias.

Metodologias experimentadas	Ac(%)	Se(%)	Es(%)
MT-T32	95,64	43,68	95,71
MT-T12	96,70	67,41	97,74
MT-CNN32	98,94	73,65	98,97
MT-CNN12	98,90	57,86	98,93
MT-CNNH32	98,89	48,30	98,60
MT-CNNH12	98,86	69,14	98,98

Fonte: o autor.

Todos os valores de acurácia foram maior que 95%, sendo que os resultados da MT-T foram menores que a MT-CNN e a MT-CNNH. Isso demonstrando que a CNN identifica melhor as regiões císticas. Nota-se que ambas aplicações da CNN tiveram resultados bem próximos, porém os experimentos com a MT-CNN tiveram melhor acurácia. Em relação aos grupos de 12 e 32 imagens, é possível observar diferenças mais significativas para os métodos de textura, com o MT-T12 com todas as métricas superiores ao MT-T32. Em relação à CNN, os experimentos obtiveram pouca variação entre os diferentes grupos.

Verifica-se também que a sensibilidade do MT-T é inferior a maioria das demais.

Isso indica que essa metodologia, principalmente no teste MT-T32, não apresenta boa capacidade de marcar corretamente um *pixel* de cisto. De forma geral, as taxas apresentaram sensibilidade baixa, que é explicada pelas altas taxas de FN's ou baixa de VP's. Essa análise demonstra que a acurácia esta sendo mais influenciado pelos VN's que os VP's, já que estes apresentam taxas muito elevadas de acurácia e baixas de sensibilidade.

Dessa forma, há a falsa percepção de que as indicações realizadas pelos experimentos estão muito próximas das marcações feitas pelos especialistas, quando na realidade estão elevadas por causa dos VN's. Entretanto, isso representa que as metodologias são capazes de indicar as regiões sem cistos. Para mensurar o quão próximo do GT os experimentos conseguem marcar regiões é utilizado o coeficiente Dice.

Outras medidas consideradas são: quantidade de regiões marcadas que apresentam interseção com cistos e a taxa de regiões FP's por exame. A primeira representa a quantidade de cistos que possui alguma interseção com as indicações das metodologias. A segunda é a taxa de regiões conectadas indicadas pelos experimentos que não apresentam interseção com as indicações de cistos. As médias do coeficiente Dice são apresentadas na Tabela 7, juntamente com outras avaliações sobre as regiões marcadas.

Tabela 7 – Médias obtidas com o coeficiente Dice e outras avaliações.

Metodologias experimentadas	Dice	Interseção	Reg. FP/exame
MT-T32	0,05	29/43	36,47
MT-T12	0,16	13/14	34,75
MT-CNN32	0,70	43/43	1,11
MT-CNN12	0,58	11/14	0,80
MT-CNNH32	0,56	32/43	1,37
MT-CNNH12	0,71	14/14	0,46

Fonte: o autor.

Nessa tabela, observa-se que as metodologias MT-T apresentam taxas Dice muito inferiores, indicando que as máscaras finais obtidas possuem grandes divergências em relação ao *GT*. Isso ocorre pela elevada quantidade de regiões FP's, com taxa de 36,47/exame para a MT-T32 e 34,75/exame para a MT-T12. Apesar disso, as regiões selecionadas pela MT-T32 intercedem com 29 de 43 cistos e a MT-T12 com 13 de 14 cistos.

Percebe-se portanto, que a metodologia por textura consegue marcar regiões císticas que apresentam considerável número de interseções. No entanto, essa metodologia realiza indicações de diversas regiões não císticas, apresentando alta taxa de regiões FP's por exame, que é um fator prejudicial para a métrica Dice.

Em relação a CNN, foram obtidos valores Dice entre 0,56 e 0,71. Comparando os resultados da MT-CNN e MT-CNNH, percebe-se que o modelo híbrido apresenta redução da métrica para o grupo de 32 imagens e aumento para o grupo de 12 imagens. Isso indica que as etapas de segmentação inicial da metodologia prejudicam a CNN no grupo de 32 imagens, reduzindo o Dice em 0,14, e auxiliam no grupo de 12 imagens, aumentando o Dice em 0,13. As altas taxas de Dice com a CNN refletem resultados próximos às do *GT*, pois apresentam interseção com quase todos os cistos e possuem baixas taxas de regiões FP's por exame.

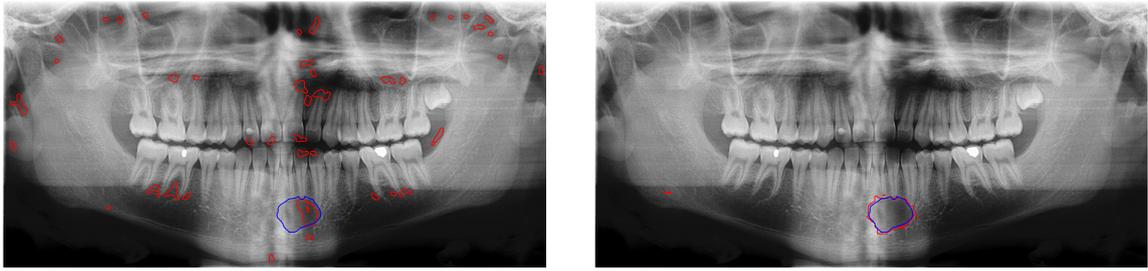
Comparando os resultados dos grupos de 32 e 12 imagens, percebe-se que o de 12 apresenta maiores métricas de Dice na MT-T e MT-CNNH, com redução na taxa de FP's por exame. Isso demonstra que as imagens de baixa qualidade prejudicaram significativamente os experimentos testados nas metodologias com as etapas de segmentação inicial. No caso da MT-CNN, a remoção destas imagens prejudicaram a classificação.

De forma geral e considerando o banco completo (com as 32 imagens), verifica-se que dentre as duas metodologias propostas nessa pesquisa, a MT-CNNH apresenta Dice de 0,56 sendo resultados superiores a MT-T. No entanto, aplicando a CNN diretamente na imagem (MT-CNN) os resultados obtidos são maiores, com taxa de 0,70. Essa investigação demonstra que a *U-net* aplicada nessa base de exames apresenta melhores métricas com a imagem completa e que a segmentação realizada reduz informações visuais dos exames e prejudica o resultado dessa CNN. Esse fato é evidenciado na própria etapa de segmentação inicial, que reduz a área cística e envia para a rede imagens com a região patológica prejudicada. Exemplos de marcações realizadas por essas três metodologias no grupo de 32 imagens são apresentadas na Figura 30.

A Figura 30 ilustra um dos resultados no grupo de 32 imagens. A indicação em azul corresponde ao *GT* e em vermelho às marcações preditas pelos métodos. A Figura 30a apresenta as marcações pela análise de textura. Apesar de parte da região cística ser encontrada, há alta incidência de FP's. Nas predições da CNN sem segmentação (Figura 30b) percebe-se que as marcações persistem na área patológica. O resultado obtido com o modelo híbrido (Figura 30c) encontra a região cística parcialmente, encontrando principalmente as regiões com tonalidade mais escura. A Figura 31 demonstra os resultados obtidos pelas três metodologias com o grupo de 12 imagens.

Na Figura 31a são demonstradas as marcações obtidas com a análise de textura, cujo padrão é semelhante ao apresentado na Figura 30a. A Figura 31b apresenta o resultado obtido com a CNN e a Figura 31c a indicação com o modelo híbrido. Percebe-se para os dois últimos

Figura 30 – Imagens geradas pelas metodologias com o grupo de 32 imagens.



(a) MT-T

(b) MT-CNN



(c) MT-CNNH

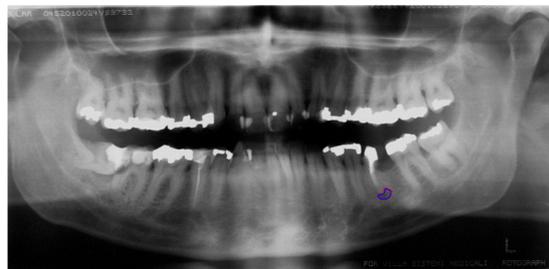
Fonte: Imagens obtidas durante os experimentos.

Figura 31 – Imagens geradas pelas metodologias com o grupo de 12 imagens.



(a) MT-T

(b) MT-CNN



(c) MT-CNNH

Fonte: Imagens obtidas durante os experimentos.

métodos que as indicações obtidas estão próximas do *GT*.

Os resultados, evidenciam que os modelos de aprendizado profundo indicam as

regiões císticas com pequena incidência de FP's. O método híbrido com 32 imagens apresenta redução da média Dice, devido à etapa de segmentação inicial ser atrapalhada pelas imagens de baixa qualidade, ocorrência que não é evidenciada no grupo de 12 imagens.

Na análise de textura, houve alta incidência de FP's, porém as regiões indicadas coincidem com diversos cistos, sendo que no grupo de 12 imagens coincidem com quase todos. Assim, ele também demonstra a capacidade de reconhecer os padrões semelhantes a cistos, embora isso contribua para o aumento de FP's. Além disso, do ponto de vista clínico, o sistema cumpre o papel de auxiliar o profissional de odontologia a verificar com melhor atenção regiões que apresentam maior possibilidade da patologia.

4.5 Considerações Finais

Este capítulo, apresentou os resultados obtidos com a metodologia desenvolvida, através das diversas investigações intermediárias e das abordagens completas, com textura, CNN e o modelo híbrido. Os resultados apresentados nesse capítulo são promissores, uma vez que as marcações obtidas ao final da proposta são relevantes em meio clínico, com potencial para fornecer mais características sobre o perfil da patologia encontrada.

O capítulo seguinte apresenta as conclusões do trabalho obtidas na pesquisa, suas contribuições e sugestões de trabalhos futuros.

5 CONCLUSÕES E TRABALHOS FUTUROS

Nessa pesquisa foi apresentada uma metodologia para detecção automática de cistos em imagens de radiografia panorâmica odontológicas, através do desenvolvimento de um software para a formação de bancos de imagens e da aplicação de técnicas clássicas e modernas de visão computacional. Diversos ensaios e experimentos foram realizados para auxiliar na escolha dos métodos utilizados. Um banco de imagens de radiografia panorâmica foi obtido em parceria com o curso de Odontologia da Universidade Federal do Ceará - UFC do *Campus Sobral*. Nessa pesquisa, foi proposto um algoritmo para obtenção da mandíbula, e realizados experimentos com técnicas de segmentação, descrição de texturas, classificação, similaridade dentro da marcação do cisto e aplicação de uma CNN. A metodologia possibilitou três resultados finais, que foram avaliados em relação a acurácia, sensibilidade, especificidade e coeficiente Dice, sendo o melhor obtido com a CNN aplicada diretamente nas imagens. Assim, a pesquisa cumpre o principal objetivo: desenvolver uma metodologia que detecta automaticamente cistos nas imagens de radiografia panorâmicas. Além disso, foram analisadas diferentes técnicas de visão computacional, na qual o melhor método para a segmentação da região interna da mandíbula foi o multinível de Otsu. A análise de textura obteve melhores resultados através do descritor de Haralick combinado aos classificadores ELM e RNA. Além disso, os estudos de segmentação a partir da região patológica mostraram que a técnica de contornos ativos com o filtro morfológico é a melhor combinação para a segmentação final.

Dentre as investigações realizadas, surgiram as seguintes propostas de trabalhos futuros:

- Avaliar o algoritmo de obtenção da mandíbula deste trabalho, através da formação de outro *ground truth* pelos especialistas, visto que a curvatura da mandíbula também apresenta interesse clínico;
- Aumentar o número de classes durante a extração de textura, assumindo a hipótese de que rotular todas as estruturas não císticas em uma mesma classe esteja atrapalhando a classificação;
- Propor uma técnica que evite perda na área cística na etapa de segmentação inicial;
- Implementar testes com outras técnicas de validação para avaliar a CNN;
- Investigar mais técnicas de pós-processamento que, a partir das regiões obtidas com a análise de textura, reduza o número de FP's;

- Investigar métodos para a identificação de cistos em imagens de tomografia, que teriam análises tridimensionais e mais informações a respeito da patologia;
- Realizar a identificação sobre o tipo de cisto e suas características clínicas.

Como contribuições deste trabalho, destacamos os seguintes artigos publicados em anais de congressos nacionais ligados à sociedades científicas:

- OLIVEIRA, D. A.; PINTO JÚNIOR, R. F.; PAULA JÚNIOR, I. C. Detecção automática de lesões periapicais em imagens de radiografia. XVI Congresso Brasileiro de Informática em Saúde, Fortaleza, p.231-244, 2018.
- OLIVEIRA, D. A.; SOUSA, L. M. R.; SAMPIERE, M. B. d. S.; CHAVES, F. N.; PAULA JÚNIOR, I. C. Identificação de regiões císticas dentárias em imagens de radiografia panorâmica utilizando algoritmos de segmentação e classificação por textura. 14º Simpósio Brasileiro de Automação Inteligente, Ouro Preto, v. 1, 2019.

REFERÊNCIAS

- ALSMADI, M. K. A hybrid fuzzy c-means and neutrosophic for jaw lesions segmentation. **Ain Shams Engineering Journal**, Elsevier, 2016.
- AMBIKA, D.; NARENDER, S.; RISHABH, K.; RAJAN, R. History of x-rays in dentistry. **Annals of Dental Research**, v. 2, n. 1, p. 21–25, 2012.
- ANGELINA, S.; SURESH, L. P.; VENI, S. K. Image segmentation based on genetic algorithm for region growth and region merging. In: IEEE. **2012 international conference on computing, electronics and electrical technologies (ICCEET)**. [S.l.], 2012. p. 970–974.
- ARIJI, Y.; YANASHITA, Y.; KUTSUNA, S.; MURAMATSU, C.; FUKUDA, M.; KISE, Y.; NOZAWA, M.; KUWADA, C.; FUJITA, H.; KATSUMATA, A. *et al.* Automatic detection and classification of radiolucent lesions in the mandible on panoramic radiographs using a deep learning object detection technique. **Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology**, Elsevier, 2019.
- BACKES, A. R.; JUNIOR, J. J. d. M. S. **Introdução à visão computacional usando Matlab**. [S.l.]: Alta Books Editora, 2016.
- BANUMATHI, A.; KANNAMMAL, A.; ARTHEE, R.; RAJU, S.; ABHAIKUMAR, V. Automated diagnosis and severity measurement of cysts in dental x-ray images using neural network. **International Journal of Biomedical Soft Computing and Human Sciences: the official journal of the Biomedical Fuzzy Systems Association**, Biomedical Fuzzy Systems Association, v. 11, n. 1, p. 15–19, 2006.
- BILODEAU, E. A.; COLLINS, B. M. Odontogenic cysts and neoplasms. **Surgical pathology clinics**, Elsevier, v. 10, n. 1, p. 177–222, 2017.
- BIRDAL, R. G.; GUMUS, E.; SERTBAS, A.; BIRDAL, I. S. Automated lesion detection in panoramic dental radiographs. **Oral Radiology**, Springer, v. 32, n. 2, p. 111–118, 2016.
- BISHOP, C. M. **Pattern recognition and machine learning**. [S.l.]: springer, 2006.
- BOUSQUET, O.; LUXBURG, U. von; RÄTSCH, G. **Advanced Lectures on Machine Learning: ML Summer Schools 2003, Canberra, Australia, February 2-14, 2003, Tübingen, Germany, August 4-16, 2003, Revised Lectures**. [S.l.]: Springer, 2011. v. 3176.
- BRADLEY, D.; ROTH, G. Adaptive thresholding using the integral image. **Journal of graphics tools**, Taylor & Francis, v. 12, n. 2, p. 13–21, 2007.
- BRASIL. **SB Brasil 2010: Pesquisa Nacional de Saúde Bucal: resultados principais**. Brasília: Ministério da Saúde, 2012. v. 1.
- CAVALCANTE, T. d. S. **Análise Comparativa de Técnicas para Segmentação Automática de Vias Aéreas em Imagens de Tomografia Computadorizada do Tórax**. Dissertação (Mestrado Acadêmico em Engenharia de Teleinformática) — Departamento de Engenharia de Teleinformática: Programa de Pós Graduação em Engenharia de Teleinformática, Universidade Federal do Ceará, Sobral, 2010.
- CHAN, T. F.; VESE, L. A. Active contours without edges. **IEEE Transactions on image processing**, IEEE, v. 10, n. 2, p. 266–277, 2001.

- CORTES, C.; VAPNIK, V. Support-vector networks. 1995. **Machine Learning**, v. 20, n. 3, p. 273, 2017.
- COSTA, F. R.; ESTEVES, C.; BACELAR, M. T. Lesões benignas da mandíbula: Uma revisão pictórica. **ACTA RADIOLÓGICA PORTUGUESA**, v. 28, n. 108, p. 25–35, 2016.
- DEVI, R. K.; BANUMATHI, A.; ULAGANATHAN, G. An automated and hybrid method for cyst segmentation in dental x-ray images. **Cluster Computing**, Springer, p. 1–13, 2018.
- DICE, L. R. Measures of the amount of ecologic association between species. **Ecology**, Wiley Online Library, v. 26, n. 3, p. 297–302, 1945.
- EDLUND, A.; SANTIAGO-RODRIGUEZ, T. M.; BOEHM, T. K.; PRIDE, D. T. Bacteriophage and their potential roles in the human oral cavity. **Journal of oral microbiology**, Taylor & Francis, v. 7, n. 1, p. 27423, 2015.
- ERIKSEN, H. M.; DIMITROV, V. The human mouth: oral functions in a social complexity perspective. **Acta Odontologica Scandinavica**, Taylor & Francis, v. 61, n. 3, p. 172–177, 2003.
- FARIAS, J.; SOUZA, R.; HASSAM, S.; CARDOSO, J.; RAMOS, T.; SANTOS, H. Epidemiological study of intraosseous lesions of the stomatognathic or maxillomandibular complex diagnosed by a reference centre in brazil from 2006–2017. **British Journal of Oral and Maxillofacial Surgery**, Elsevier, 2019.
- FEJERSKOV, O.; URIBE, S.; MARIÑO, R. J. Dentistry in a historical perspective and a likely future of the profession. In: **Career Paths in Oral Health**. [S.l.]: Springer, 2018. p. 3–19.
- GHOM, A. G.; GHOM, S. A. L. **Textbook of oral medicine**. [S.l.]: JP Medical Ltd, 2014.
- GONZALEZ, R. C.; WOODS, R. C. **Processamento digital de imagens**. [S.l.]: Pearson Educación, 2009.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep learning**. [S.l.]: MIT press, 2016.
- GORDÓN-NÚÑEZ, M.; NONAKA, C.; TORRES, T.; EMILIANO, G. *et al.* Odontogenic cysts: demographic profile in a brazilian population over a 38-year period. **Medicina oral, patologia oral y cirugía bucal**, v. 15, n. 4, p. e583–90, 2010.
- GROSSMANN, S. M.; MACHADO, V. C.; XAVIER, G. M.; MOURA, M. D.; GOMEZ, R. S.; AGUIAR, M. C. F.; MESQUITA, R. A. Demographic profile of odontogenic and selected nonodontogenic cysts in a brazilian population. **Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology**, Elsevier, v. 104, n. 6, p. e35–e41, 2007.
- GUENTHER, N.; SCHONLAU, M. Support vector machines. **The Stata Journal**, SAGE Publications Sage CA: Los Angeles, CA, v. 16, n. 4, p. 917–937, 2016.
- HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. H. Textural features for image classification. **IEEE Transactions on systems, man, and cybernetics**, Ieee, n. 6, p. 610–621, 1973.
- HAYKIN, S. S. *et al.* **Neural networks and learning machines/Simon Haykin**. [S.l.]: New York: Prentice Hall,, 2009.

- HSU, C.; CHANG, C.; LIN, C. **A practical guide to support vector classification, Taipei: National Taiwan University**. 2016.
- HUANG, G.-B.; ZHU, Q.-Y.; SIEW, C.-K. Extreme learning machine: a new learning scheme of feedforward neural networks. In: **2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541)**. [S.l.: s.n.], 2004. v. 2, p. 985–990 vol.2.
- IANNUCCI, J.; HOWERTON, L. **J. Dental Radiography-E-Book: Principles and Techniques**. [S.l.]: Elsevier Health Sciences, 2016.
- IHME. **Global Burden of Disease Study 2017**. 2018. Disponível em: <<http://ghdx.healthdata.org/gbd-2017>>. Acesso em: 17 jul. 2019.
- IMRAN, A.; JAYANTHI, P.; TANVEER, S.; GOBU, S. C. Classification of odontogenic cysts and tumors—antecedents. **Journal of oral and maxillofacial pathology: JOMFP**, Wolters Kluwer–Medknow Publications, v. 20, n. 2, p. 269, 2016.
- JAIN, A. K. Data clustering: 50 years beyond k-means. **Pattern recognition letters**, Elsevier, v. 31, n. 8, p. 651–666, 2010.
- JAMES, S. L.; ABATE, D.; ABATE, K. H.; ABAY, S. M.; ABBAFATI, C.; ABBASI, N.; ABBASTABAR, H.; ABD-ALLAH, F.; ABDELA, J.; ABDELALIM, A. *et al.* Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. **The Lancet**, Elsevier, v. 392, n. 10159, p. 1789–1858, 2018.
- KANG, W.-X.; YANG, Q.-Q.; LIANG, R.-P. The comparative research on image segmentation algorithms. In: IEEE. **2009 First International Workshop on Education Technology and Computer Science**. [S.l.], 2009. v. 2, p. 703–707.
- KASS, M.; WITKIN, A.; TERZOPOULOS, D. Snakes: Active contour models. **International journal of computer vision**, Springer, v. 1, n. 4, p. 321–331, 1988.
- KASSEBAUM, N.; SMITH, A.; BERNABÉ, E.; FLEMING, T.; REYNOLDS, A.; VOS, T.; MURRAY, C.; MARCENES, W.; COLLABORATORS, G. . O. H. Global, regional, and national prevalence, incidence, and disability-adjusted life years for oral conditions for 195 countries, 1990–2015: A systematic analysis for the global burden of diseases, injuries, and risk factors. **Journal of Dental Research**, v. 96, n. 4, p. 380–387, 2017. PMID: 28792274.
- KATSUMATA, A.; FUJITA, H. Progress of computer-aided detection/diagnosis (cad) in dentistry<runningtitle>cad in dentistry</runningtitle>. **Japanese Dental Science Review**, v. 50, n. 3, p. 63 – 68, 2014. ISSN 1882-7616. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1882761614000040>>.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **nature**, Nature Publishing Group, v. 521, n. 7553, p. 436, 2015.
- LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. *et al.* Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, Taipei, Taiwan, v. 86, n. 11, p. 2278–2324, 1998.
- 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.

LIAO, P.-S.; CHEN, T.-S.; CHUNG, P.-C. *et al.* A fast algorithm for multilevel thresholding. **J. Inf. Sci. Eng.**, v. 17, n. 5, p. 713–727, 2001.

LIMA, F. D. S. **Segmentação de lesões periapicais em imagens odontológicas radiográficas digitais utilizando modelos de contorno ativo com inicialização automática.** Dissertação (Mestrado Acadêmico em Engenharia Elétrica e de Computação) — Departamento de Engenharia Elétrica e de Computação: Programa de Pós-Graduação em Engenharia Elétrica e de Computação, Universidade Federal do Ceará, Sobral, 2017.

LISTL, S.; GALLOWAY, J.; MOSSEY, P.; MARCENES, W. Global economic impact of dental diseases. **Journal of dental research**, SAGE Publications Sage CA: Los Angeles, CA, v. 94, n. 10, p. 1355–1361, 2015.

LOUREDO, B. V. R.; FREITAS, C. T. S. de; CÂMARA, J.; LIBÓRIO-KIMURA, T. N. Estudo epidemiológico de lesões odontogênicas provenientes do departamento de patologia e medicina legal da universidade federal do amazonas. **Revista Brasileira de Odontologia**, v. 74, n. 2, p. 126, 2017.

MACQUEEN, J. *et al.* Some methods for classification and analysis of multivariate observations. In: OAKLAND, CA, USA. **Proceedings of the fifth Berkeley symposium on mathematical statistics and probability.** [S.l.], 1967. v. 1, n. 14, p. 281–297.

MAHMOUD, Y. E.; LABIB, S. S.; MOKHTAR, H. M. Teeth periapical lesion prediction using machine learning techniques. In: IEEE. **2016 SAI Computing Conference (SAI).** [S.l.], 2016. p. 129–134.

MARTIN, L. H.; SPEIGHT, P. M. Odontogenic cysts: an update. **Diagnostic Histopathology**, Elsevier, v. 23, n. 6, p. 260–265, 2017.

MARTINEZ, B. M. **In a New Light: Early X-Ray Technology in Dentistry, 1890-1955.** Dissertação (Mestrado) — Arizona State University, 2013.

MCDONALD, S. W.; MACFARLANE, N. G. The mouth, stomach and intestines. **Anaesthesia & Intensive Care Medicine**, Elsevier, v. 19, n. 3, p. 128–132, 2018.

MIKULKA, J.; KABRDA, M.; GESCHEIDTOVÁ, E.; PEŘINA, V. Classification of jawbone cysts via orthopantomogram processing. In: IEEE. **2012 35th International Conference on Telecommunications and Signal Processing (TSP).** [S.l.], 2012. p. 499–502.

NEVILLE, B. W.; DAMM, D. D.; ALLEN, C. M.; BOUQUOT, J. E. **Patologia Oral e Maxilofacial.** [S.l.]: Elsevier Brasil, 2009.

NURTANIO, I.; ASTUTI, E. R.; PURNAMA, I. K. E.; HARIADI, M.; PURNOMO, M. H. Classifying cyst and tumor lesion using support vector machine based on dental panoramic images texture features. **IAENG International Journal of Computer Science**, International Association of Engineers, v. 40, n. 1, p. 29–37, 2013.

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, 2011.

- OHASHI, Y.; ARIJI, Y.; KATSUMATA, A.; FUJITA, H.; NAKAYAMA, M.; FUKUDA, M.; NOZAWA, M.; ARIJI, E. Utilization of computer-aided detection system in diagnosing unilateral maxillary sinusitis on panoramic radiographs. **Dentomaxillofacial Radiology**, The British Institute of Radiology., v. 45, n. 3, p. 20150419, 2016.
- OJALA, T.; PIETIKAINEN, M.; MAENPAA, P. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 24, n. 7, p. 971–987, July 2002.
- OTSU, N. A threshold selection method from gray-level histograms. **IEEE transactions on systems, man, and cybernetics**, IEEE, v. 9, n. 1, p. 62–66, 1979.
- PEIXOTO, A.; VELHO, L. C. **Transformadas de distância**. [S.l.]: PUC, 2000.
- POEDJIASTOETI, W.; SUEBNUKARN, S. Application of convolutional neural network in the diagnosis of jaw tumors. **Healthcare informatics research**, v. 24, n. 3, p. 236–241, 2018.
- PONTES, C. G. C.; NETO, A. I. T.; RIBEIRO, I. L. H.; SARMENTO, V. A.; SANTOS, J. N. d.; AZEVEDO, R. A. Epidemiologia dos cistos e tumores odontogênicos tratados sob anestesia geral, em um hospital filantrópico de salvador: Bahia. **Revista de Cirurgia e Traumatologia Buco-maxilo-facial**, UPE/FO, v. 12, n. 1, p. 93–100, 2012.
- REDMON, J.; FARHADI, A. Yolov3: An incremental improvement. **arXiv**, 2018.
- RONCALLI, A. G. **Projeto SB Brasil 2010-pesquisa nacional de saúde bucal revela importante redução da cárie dentária no país**. [S.l.]: SciELO Public Health, 2011.
- RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: SPRINGER. **International Conference on Medical image computing and computer-assisted intervention**. [S.l.], 2015. p. 234–241.
- RUSSAKOVSKY, O.; DENG, J.; SU, H.; KRAUSE, J.; SATHEESH, S.; MA, S.; HUANG, Z.; KARPATY, A.; KHOSLA, A.; BERNSTEIN, M.; BERG, A. C.; FEI-FEI, L. ImageNet Large Scale Visual Recognition Challenge. **International Journal of Computer Vision (IJCV)**, v. 115, n. 3, p. 211–252, 2015.
- SANTOS, T. d. S.; ANTUNES, A. A.; AVELAR, R. L.; ANTUNES, A. P. Cistos odontogênicos: estudo epidemiológico de 72 casos. **Revista brasileira de cirurgia da cabeça e pescoço**, p. 30–32, 2007.
- SETHIAN, J. A. **Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science**. [S.l.]: Cambridge university press, 1999. v. 3.
- SØRENSEN, T. J. **A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons**. [S.l.]: I kommission hos E. Munksgaard, 1948.
- SOUSA, E. L. R. de; FERRAZ, C. C. R.; GOMES, B. P. F. de A.; PINHEIRO, E. T.; TEIXEIRA, F. B.; SOUZA-FILHO, F. J. de. 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.

- STELT, P. F. V. D. Filmless imaging: the uses of digital radiography in dental practice. **The Journal of the American Dental Association**, Elsevier, v. 136, n. 10, p. 1379–1387, 2005.
- SU, Y.; WANG, J.; ZHANG, T.; ZHANG, Z. Odontogenic tumors and odontogenic cysts: a clinical and pathological analysis of 4 181 cases. **Chinese journal of stomatology**, v. 54, n. 8, p. 546–552, 2019.
- TANG, J. A color image segmentation algorithm based on region growing. In: IEEE. **2010 2nd International Conference on Computer Engineering and Technology**. [S.l.], 2010. v. 6, p. V6–634.
- VARINAUSKAS, V.; GERVICKAS, A.; ROŽNOVA, O. Analysis of odontogenic cysts of the jaws. **Medicina**, Lietuvos gydytojų sąjunga, v. 42, n. 3, p. 201–207, 2006.
- VIJAYAKUMARI, B.; ULAGANATHAN, G.; BANUMATHI, A.; BANU, A. F. S.; KAYALVIZHI, M. Dental cyst diagnosis using texture analysis. In: IEEE. **2012 International Conference on Machine Vision and Image Processing (MVIP)**. [S.l.], 2012. p. 117–120.
- WHAITES, E.; DRAGE, N. **Essentials of dental radiography and radiology**. [S.l.]: Elsevier Health Sciences, 2013.
- WHITE, S. C.; PHAROAH, M. J. **Oral Radiology: principles and interpretation**. [S.l.]: ELSEVIER INDIA, 2014.
- WITTEN, I. H.; FRANK, E.; HALL, M. A.; PAL, C. J. **Data Mining: Practical machine learning tools and techniques**. [S.l.]: Morgan Kaufmann, 2016.
- YONG, T.-H.; LEE, S.-J.; YI, W.-J. Odontogenic cysts and tumors detection in panoramic radiographs using deep convolutional neural network (dcnn). 2019.
- ZAITOUN, N. M.; AQEL, M. J. Survey on image segmentation techniques. **Procedia Computer Science**, Elsevier, v. 65, p. 797–806, 2015.
- ZHANG, X.; ZOU, J.; HE, K.; SUN, J. Accelerating very deep convolutional networks for classification and detection. **IEEE transactions on pattern analysis and machine intelligence**, IEEE, v. 38, n. 10, p. 1943–1955, 2015.