

UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
BRUNA BONATTI SIVIERI

**Aplicabilidade da inteligência artificial no diagnóstico,
planejamento e aquisição de imagens em Tomografia
Computadorizada de Feixe Cônico: Revisão de literatura**

CAMPO GRANDE
2025

BRUNA BONATTI SIVIERI

**Aplicabilidade da inteligência artificial no diagnóstico,
planejamento e aquisição de imagens em Tomografia
Computadorizada de Feixe Cônico: Revisão de literatura**

Trabalho de Conclusão de Curso da
Faculdade de Odontologia da Universidade
Federal de Mato Grosso do Sul, para
obtenção do título de Cirurgiã-Dentista.

Orientador: Prof. Dr. Yuri Nejaim

CAMPO GRANDE

2025

**Aplicabilidade da inteligência artificial no diagnóstico,
planejamento e aquisição de imagens em Tomografia
Computadorizada de Feixe Cônico: Revisão de literatura**

Trabalho de Conclusão de Curso da
Faculdade de Odontologia da Universidade
Federal de Mato Grosso do Sul, para
obtenção do título de Cirurgiã-Dentista.

Resultado: _____

Campo Grande (MS), 14 de novembro de 2025.

BANCA EXAMINADORA

Prof. Dr. Yuri Nejaim

Faculdade de Odontologia da Universidade Federal de Mato Grosso do Sul / UFMS

Prof. Dr. Victor de Aquino Wanderley

Faculdade de Odontologia da Universidade Federal de Mato Grosso do Sul / UFMS

Prof. Dr. Carmen Regina Coldebella

Faculdade de Odontologia da Universidade Federal de Mato Grosso do Sul / UFMS

AGRADECIMENTOS

Agradeço à minha família, cujo apoio foi essencial ao longo de toda minha trajetória acadêmica. Aos meus pais, agradeço pelo exemplo de integridade, dedicação e perseverança, que sempre nortearam minhas escolhas e me incentivaram a seguir em frente, mesmo diante dos desafios.

Aos amigos que o curso me proporcionou, pela troca de experiências e desafios superados em conjunto. Compartilhar essa jornada com pessoas tão comprometidas e inspiradoras foi, sem dúvida, um dos maiores privilégios dessa etapa.

À Elisabeth D 'Ávila Alcântara de Oliveira, com quem compartilhei não apenas trabalhos e responsabilidades acadêmicas, mas também aprendizados e conquistas. Sua parceria constante, dedicação, companheirismo e comprometimento foram fundamentais para que conseguíssemos superar cada etapa desse percurso.

Ao meu orientador, Prof. Dr. Yuri Nejaim, pela generosidade em compartilhar seu conhecimento e pela confiança depositada ao longo do desenvolvimento deste trabalho. Agradeço pela paciência, pelas valiosas contribuições e pela escuta atenta, essenciais à concretização deste trabalho.

À Faculdade de Odontologia da Universidade Federal de Mato Grosso do Sul, que foi essencial para meu crescimento pessoal e profissional. Registro aqui minha gratidão pelos cinco anos de ensinamentos aos quais tantos professores se dedicaram. Sou grata a cada um.

RESUMO

Sivieri, BB. Aplicabilidade da inteligência artificial no diagnóstico, planejamento e aquisição de imagens em Tomografia Computadorizada de Feixe Cônico: Revisão de literatura. Campo Grande; 2025. [Trabalho de conclusão de curso – Universidade Federal de Mato Grosso do Sul].

Nos últimos anos, a inteligência artificial (IA) tem se consolidado como um tema central na área da saúde, especialmente na odontologia, em que suas aplicações em tomografias computadorizadas de feixe cônico (TCFC) vêm sendo amplamente estudadas. Diante do crescente volume de pesquisas, o presente trabalho realizou uma revisão de literatura, analisando 101 artigos científicos indexados na base PubMed, publicados entre 2014 e 2025. A seleção priorizou estudos relacionados ao diagnóstico odontológico, ao planejamento de tratamentos orais e ao aprimoramento da qualidade das imagens tomográficas. Os resultados evidenciam uma integração progressiva da IA na interpretação de exames de TCFC, com ganhos significativos em acurácia diagnóstica e na elaboração de planos terapêuticos mais precisos. Embora ainda incipiente em contextos clínicos, a literatura indica que a IA pode aumentar a segurança e a agilidade das intervenções, mesmo sob supervisão profissional. Apesar do potencial para aprimorar a qualidade e a eficiência dos cuidados odontológicos, persistem desafios como os altos custos de implementação, a necessidade de capacitação profissional, a escassez de estudos com ampla aplicabilidade clínica e as limitações relacionadas à diversidade dos equipamentos utilizados. Pesquisas futuras devem concentrar-se na mitigação de vieses, na diversificação dos modelos de IA e na validação clínica dessas tecnologias, a fim de consolidar sua utilização prática. A adoção da IA nas tomografias odontológicas mostra-se promissora para a promoção de diagnósticos mais individualizados, precisos e seguros, representando um avanço transformador na prática odontológica.

Palavras-chave: Tomografia Computadorizada de Feixe Cônico. Inteligência Artificial. Odontologia.

ABSTRACT

Sivieri, BB. Applicability of artificial intelligence in the diagnosis, planning, and image acquisition in cone-beam computed tomography: a literature review; 2025. [Trabalho de conclusão de curso – Universidade Federal de Mato Grosso do Sul].

In recent years, artificial intelligence (AI) has become a central topic in healthcare, particularly in dentistry, where its applications in cone-beam computed tomography (CBCT) have been extensively investigated. Given the growing volume of research, the present study conducted a literature review analyzing 101 scientific articles indexed in the PubMed database, published between 2014 and 2025. The selection prioritized studies related to dental diagnosis, treatment planning, and the enhancement of tomographic image quality. The findings reveal a progressive integration of AI into the interpretation of CBCT examinations, resulting in significant improvements in diagnostic accuracy and in the development of more precise therapeutic plans. Although still incipient in clinical settings, the literature indicates that AI can enhance the safety and agility of interventions, even under professional supervision. Despite its potential to improve the quality and efficiency of dental care, challenges remain, such as high implementation costs, the need for professional training, the limited number of studies with broad clinical applicability, and constraints related to the diversity of imaging equipment. Future research should focus on bias mitigation, diversification of AI models, and clinical validation of these technologies to strengthen their practical implementation. The adoption of AI in dental tomography appears promising for promoting more individualized, accurate, and safe diagnoses, representing a transformative advancement in dental practice.

Keywords: Computed Cone-Beam Tomography. Artificial Intelligence. Dentistry.

SUMÁRIO

1 INTRODUÇÃO	6
2 MATERIAIS E MÉTODOS	9
3 REVISÃO DE LITERATURA	11
4 DISCUSSÃO	60
5 CONCLUSÃO	64

1. INTRODUÇÃO

A segunda década do século XXI consolidou a inteligência artificial (IA) como uma ferramenta de grande relevância para a odontologia. Nesse cenário, observa-se um movimento crescente da comunidade acadêmico-científica em compreender e promover o uso dessa tecnologia como recurso coadjuvante em diferentes especialidades odontológicas, com destaque para a radiologia e a imaginologia (VASEY et al., 2023). Entre os exames de imagem mais utilizados, a tomografia computadorizada de feixe cônico (TCFC) ocupa posição de destaque, apresentando grande potencial de aprimoramento quando associada a sistemas de deep learning ou machine learning (ABESI et al., 2023). Assim, a TCFC configura-se como um campo fértil para o desenvolvimento de aplicações baseadas em IA, que ampliam sua utilidade clínica e científica.

A TCFC, por si só, já representou um marco no avanço da odontologia ao possibilitar a visualização tridimensional dos tecidos calcificados craniofaciais. Essa inovação introduziu a dimensão da profundidade, inexistente nos exames bidimensionais convencionais, permitindo diagnósticos mais precisos e planejamentos clínicos e cirúrgicos mais seguros. Por essas razões, o exame é frequentemente considerado o padrão-ouro em diversas situações clínicas (OUSEHAL et al., 2024). Com a consolidação desse papel central, torna-se compreensível a busca pela integração da IA em sua rotina, a fim de otimizar a interpretação das imagens e elevar ainda mais a confiabilidade dos laudos, o que também abre espaço para refletir sobre a aplicação mais ampla da inteligência artificial na saúde.

A IA, diferentemente da TCFC, representa uma inovação mais recente no contexto da tecnologia aplicada à odontologia. Estruturada a partir de algoritmos interpretados por redes neurais artificiais alimentadas por grandes bases de dados, demonstra especial eficácia quando associada ao deep learning, modelo capaz de deduzir hipóteses diagnósticas e propor soluções clínicas com elevado grau de acurácia (PIANYKH et al., 2020). Essa característica desperta grande interesse na área médico-odontológica, pois o uso da IA em exames imaginológicos abre caminho para diagnósticos mais rápidos, consistentes e acessíveis. Nesse sentido,

sua integração com a TCFC representa uma das contribuições mais promissoras para a prática clínica contemporânea.

A literatura científica recente apresenta inúmeras aplicações da IA voltadas especificamente à análise de exames de TCFC. Estudos relatam o uso de algoritmos de deep learning para segmentação de diferentes estruturas maxilofaciais, com valores de Dice superiores a 0,90 (CUI et al., 2022; FONTENELE et al., 2023), comparáveis à segmentação manual realizada por especialistas. Outras pesquisas demonstram o emprego da IA para detecção de canais radiculares desobturados, medição de altura e espessura óssea, identificação das articulações temporomandibulares e até mesmo classificação de patologias do seio maxilar (SANTOS-JUNIOR et al., 2025; VINAYAHALINGAM et al., 2023; ALTUN et al., 2024). Além dessas aplicações, estudos recentes destacam a capacidade de segmentação de implantes dentários em menos de 30 segundos, mesmo em situações de presença de artefatos metálicos (ELGARBA et al., 2023). Essa diversidade de usos reforça o potencial da IA em expandir os limites da interpretação e da aplicabilidade da TCFC, criando perspectivas de diagnósticos mais ágeis e confiáveis.

Esses avanços, entretanto, também trazem desafios e limitações. Diversos estudos apontam dificuldades relacionadas à necessidade de grandes volumes de dados de alta qualidade para o treinamento adequado dos modelos, além da heterogeneidade das bases de imagens, que pode comprometer a generalização dos resultados (CHEN, 2025; GÜRSES et al., 2023). Destacam-se, ainda, variações metodológicas significativas e riscos de viés nas pesquisas já publicadas (HUNG et al., 2020), evidenciando a necessidade de padronização e rigor metodológico. Questões éticas e regulatórias também emergem nesse contexto, especialmente no que se refere à transparência dos algoritmos, à validação clínica robusta e à utilização da IA como ferramenta assistiva — e não substitutiva — do julgamento humano (DIPALMA et al., 2023). Tais limitações não invalidam o potencial da tecnologia, mas indicam que sua adoção clínica deve ocorrer de forma progressiva, acompanhada de constante revisão crítica da literatura.

Dessa forma, ao considerar tanto os avanços quanto as limitações atuais, torna-se evidente a relevância de um estudo que busque analisar e descrever as aplicações contemporâneas da inteligência artificial em exames de tomografia computadorizada de feixe cônico. A síntese desse conhecimento é fundamental para os profissionais da odontologia, que encontram nesse recurso uma possibilidade concreta de aperfeiçoamento dos diagnósticos e planejamentos, com benefícios diretos à prática clínica e, principalmente, aos pacientes.

2. MATERIAIS E MÉTODOS

Estratégia de busca

A busca por artigos foi realizada exclusivamente na base de dados PubMed, abrangendo publicações no período de janeiro de 2014 a setembro de 2025. Como critério de idioma, foram incluídos apenas estudos publicados em inglês. Os termos de busca utilizados foram: “Artificial intelligence”, “Cone-beam computed tomography” e “Dentistry”, aplicados de forma combinada por meio de operadores booleanos, com o objetivo de refinar os resultados e garantir a relevância dos estudos recuperados. A pesquisa resultou em 429 artigos, que foram posteriormente triados conforme os critérios de inclusão e exclusão estabelecidos.

Critérios de inclusão e exclusão

Foram considerados elegíveis os estudos originais que apresentavam abordagens relacionadas ao uso da inteligência artificial — incluindo machine learning e deep learning — aplicadas diretamente a processos de diagnóstico, planejamento clínico ou à aquisição e interpretação de imagens obtidas por tomografia computadorizada de feixe cônico (TCFC) em contextos odontológicos.

Foram excluídos da revisão: estudos de revisão da literatura (sistemáticas ou narrativas); artigos editoriais, cartas ao editor ou comunicações breves; pesquisas indisponíveis na íntegra; estudos que tratavam da fusão multimodal de imagens (por exemplo, TCFC associada à ressonância magnética ou a outros exames médicos complementares); e trabalhos cujo foco não se concentrava diretamente na aplicação da IA ou que se distanciavam da temática central proposta nesta revisão.

A seleção dos estudos foi realizada em duas etapas. Na primeira, procedeu-se à triagem dos títulos e resumos, com o intuito de identificar preliminarmente os artigos potencialmente relevantes. Em seguida, os textos completos dos estudos considerados adequados foram acessados e analisados em profundidade.

Organização do material

Os 101 estudos selecionados foram organizados em uma planilha do Microsoft Excel, dispostos em ordem alfabética. Nessa planilha, foram agrupadas informações referentes à autoria, objetivos, materiais e métodos utilizados, resultados quantitativos e conclusões de cada trabalho. Esses dados subsidiaram a elaboração da revisão, na qual os artigos com princípios metodológicos e temáticos semelhantes foram reunidos em grupos para análise comparativa e discussão integrada.

3. REVISÃO DE LITERATURA

3.1 Qualidade da Imagem

No estudo de Oliveira et al. (2025), foi analisado o desempenho de um modelo de deep learning aplicado à correção de artefatos metálicos decorrentes da exomassa em imagens de tomografia computadorizada de feixe cônico (TCFC) obtidas ex vivo. O modelo utilizado foi um *conditional denoising diffusion probabilistic model* (DDPM), técnica de *image-to-image translation* com arquitetura baseada em U-Net. As imagens resultantes foram avaliadas por três radiologistas, que realizaram comparações pareadas entre: (1) imagens originais com artefatos e imagens corrigidas por inteligência artificial (IA); e (2) imagens corrigidas por IA e o *ground truth* (imagens obtidas sem implantes). A análise subjetiva foi conduzida por meio de uma escala de cinco pontos (1 a 5), indicando o grau de superioridade visual entre as imagens comparadas. De modo geral, as imagens corrigidas pela IA apresentaram pontuação média superior a 4, o que indica que foram frequentemente consideradas possivelmente ou definitivamente superiores às imagens originais. A confiabilidade intraexaminador foi classificada como quase perfeita (coeficiente Kappa $\approx 0,94$), enquanto a confiabilidade interexaminador foi considerada substancial (Kappa $\approx 0,79$). Como parâmetro quantitativo, utilizou-se o *Contrast-to-Noise Ratio* (CNR), que demonstrou valores significativamente mais altos nas imagens corrigidas em comparação às originais — e, em diversos casos, até superiores ao *ground truth*. Esses resultados mostraram-se consistentes independentemente do equipamento de TCFC, do tipo de material dos implantes ou do número de implantes utilizados, indicando que o modelo de IA foi eficaz em restaurar o contraste e reduzir o ruído decorrente dos artefatos metálicos.

No estudo conduzido por Kim et al. (2024), foi avaliado o impacto da aplicação de IA na qualidade de imagens de TCFC obtidas com protocolos de baixa dose, visando reduzir a exposição à radiação sem comprometer a qualidade diagnóstica. O modelo utilizado foi uma rede neural convolucional denominada EDCNN (*Edge Enhancement–Densely Connected CNN*), que incorpora um módulo específico para realce de bordas. As imagens analisadas (nove por dente) foram obtidas a partir de

um único dente de fantoma, utilizando quatro protocolos experimentais com IA e um protocolo controle (alta resolução, sem IA). Onze parâmetros anatômicos e diagnósticos foram avaliados por examinadores cegos quanto ao protocolo empregado, abrangendo estruturas como raízes, canais radiculares, lâmina dura, padrão trabecular e córtex mandibular, além da qualidade geral da imagem para o diagnóstico de lesões periapicais. A confiabilidade entre avaliadores foi considerada substancial (Fleiss' kappa $\approx 0,621$), enquanto a confiabilidade intraexaminador foi quase perfeita (ICC $\approx 0,995$). Os protocolos que incorporaram IA apresentaram qualidade subjetiva comparável ou superior à do protocolo controle, dependendo do parâmetro avaliado, indicando que o uso de IA pode compensar a perda de qualidade resultante da redução da dose de radiação.

Por fim, o estudo de Minnema et al. (2019) teve como foco a segmentação óssea em imagens de TCFC contendo artefatos metálicos, comparando o desempenho de diferentes redes neurais convolucionais. A arquitetura proposta, denominada *Mixed-Scale Dense Network* (MS-D), foi composta por 100 camadas convolucionais com o uso de convoluções dilatadas, o que permite capturar simultaneamente detalhes finos e estruturas de maior escala, com maior eficiência paramétrica. As performances das arquiteturas MS-D, U-Net, ResNet e de um algoritmo tradicional semi-automático (snake evolution) foram comparadas. Os resultados mostraram desempenho semelhante entre as redes modernas, com coeficiente de similaridade de Dice (DSC) médio de aproximadamente 0,87 para MS-D e U-Net, e 0,86 para ResNet. O método snake evolution apresentou DSC inferior ($\sim 0,78$). No que se refere à eficiência, os modelos baseados em IA realizaram a segmentação de cada TCFC em menos de 5 minutos, enquanto o método tradicional demandou entre 20 minutos e 1 hora. A arquitetura MS-D destacou-se por atingir desempenho competitivo com um número menor de parâmetros treináveis, o que pode reduzir o risco de *overfitting* e acelerar o processo de treinamento.

3.2 Dentes

A revisão identificou 24 estudos que abordaram o uso de inteligência artificial (IA) aplicada à análise de dentes em exames de tomografia computadorizada de feixe cônico (TCFC). Os trabalhos se distribuíram em diferentes subtemas:

segmentação de dentes inteiros ($n = 7$), segmentação de estruturas dentais internas ($n = 3$), avaliação de reabsorções radiculares ($n = 4$), identificação de canais radiculares ($n = 2$), proximidade entre raízes e canal mandibular ($n = 2$), patologias dentárias ($n = 3$), identificação de fraturas dentárias ($n = 2$), além de um estudo combinando dentes e segmentação óssea.

3.2.1 Segmentação de Dentes Inteiros e Classificação

No estudo de Cui et al. (2022), foi utilizado um dos maiores conjuntos de dados já relatados na literatura, composto por 4.938 exames de tomografia computadorizada de feixe cônico (TCFC) provenientes de 4.215 pacientes atendidos em 15 centros distintos. O modelo proposto apresentou desempenho robusto na segmentação de dentes individuais, alcançando média de *Dice Similarity Coefficient* (DSC) de 94,1%, sensibilidade de 93,9% e *Average Surface Distance* (ASD) de 0,17 mm. Com o auxílio da inteligência artificial (IA), o tempo necessário para correção manual foi reduzido em mais de 95%, passando a ser inferior a cinco minutos por exame.

Em um contexto clínico pediátrico, Ozudogru et al. (2025) aplicaram o modelo *nnU-Net v2* em 49 exames de dentição mista. Apesar da complexidade anatômica característica dessa faixa etária, o modelo demonstrou acurácia de 0,99, precisão de 0,86 e *recall* de 0,84. O coeficiente *Dice* ($\sim 0,81$) indicou desempenho satisfatório mesmo em casos considerados desafiadores.

Gerhardt et al. (2022) utilizaram a ferramenta automatizada *Virtual Patient Creator*, desenvolvida para detecção, rotulagem e segmentação dentária automática. Em uma amostra de 175 exames de TCFC, a ferramenta atingiu precisão de 99,7% em arcadas totalmente dentadas e 99% mesmo na presença de pequenas regiões edêntulas. A segmentação apresentou *Intersection over Union* (IoU) entre 0,96 e 0,97, com redução expressiva no tempo de anotação — de 98 segundos no método manual para apenas 1,5 segundo com o uso da IA.

Já Ayidh Alqahtani et al. (2023) propuseram uma abordagem multiclasse utilizando múltiplos modelos *U-Net* em diferentes resoluções. Em 215 exames, totalizando 1.780 dentes, a segmentação e classificação simultânea das 32 classes

dentárias foi realizada com *IoU* e *DSC* médios de aproximadamente 0,99. A *Hausdorff Distance* 95% foi de apenas 0,12 mm, e o tempo de processamento foi reduzido de 43,6 para 13,7 segundos por exame completo.

Swaity et al. (2024) concentraram-se na segmentação de caninos impactados, empregando dois modelos *3D U-Net* aplicados em etapas sucessivas. Em uma amostra de 200 exames de TCFC, o modelo alcançou *DSC* médio de aproximadamente 0,99 e *IoU* também próximo de 0,99, com tempo médio de segmentação de 21 segundos. O método automático demonstrou ser cerca de 24 vezes mais rápido que o método semiautomático de referência.

No estudo conduzido por Hsu et al. (2022), foram comparadas diferentes variações da *U-Net* (2D, 2.5D, 3D e combinações baseadas em votação majoritária). A melhor performance foi obtida com o modelo *3.5D v5*, associado a um processo de pós-processamento envolvendo operações morfológicas de erosão e dilatação. Esse modelo atingiu *DSC* de aproximadamente 0,961 e acurácia de 98,6%, evidenciando que o uso de arquiteturas híbridas e estratégias de *ensemble* pode aprimorar ainda mais a segmentação dentária.

Elsonbaty et al. (2025) aplicaram uma arquitetura dupla *3D U-Net* para segmentação de 402 dentes decíduos em 37 exames de TCFC pediátricos. Os resultados foram expressivos, com *DSC* = $95 \pm 2\%$, *IoU* = $91 \pm 3\%$, acurácia = $98 \pm 1\%$ e *Hausdorff Distance* 95% $\approx 0,27$ mm. O tempo médio de segmentação por dente foi reduzido de aproximadamente 13 minutos (manual) para cerca de 24 segundos com IA. Pequenos ajustes manuais foram necessários em cerca de 40% dos casos, geralmente localizados nas regiões apicais.

Por fim, Li et al. (2024) utilizaram um modelo baseado em *regional level set* (*Chan–Vese – CV*) otimizado por *Local Binary Fitting* (*CV-LBF*), aplicado a 82 exames pediátricos. A acurácia de segmentação alcançou 89,9%, superando métodos convencionais como *region growing* (78,6%) e *Distance Regularized Level Set Evolution* (*DRLSE*) (68,2%). As menores taxas de subsegmentação e as menores distâncias de superfície observadas indicaram maior fidelidade ao contorno anatômico real.

3.2.2 Segmentação de estruturas dentais internas

No estudo conduzido por Dogan et al. (2024), foram analisadas 236 imagens de tomografia computadorizada de feixe cônico (TCFC) de indivíduos com idades entre 18 e 70 anos, resultando em um total de 1.416 medições da profundidade da translucidez radicular (*PTR*) realizadas em seis dentes selecionados por indivíduo (de um mesmo lado da arcada dentária). O objetivo foi classificar os indivíduos em cinco faixas etárias por meio de três algoritmos de *machine learning* (*ML*): *Support Vector Machine* (*SVM*), *Classification and Regression Tree* (*CART*) e *Random Forest* (*RF*). Os resultados evidenciaram limitações na acurácia dos modelos, com desempenho variando em torno de 0,25 para *CART* e *SVM*, e 0,30 para *RF*, indicando que a translucidez radicular isoladamente pode não ser um preditor suficientemente robusto da idade cronológica.

Por outro lado, Slim et al. (2024) investigaram a aplicabilidade de redes neurais convolucionais (*Convolutional Neural Networks – CNNs*) em um estudo retrospectivo com 66 exames de TCFC, focando na segmentação automática de molares inferiores. Foi empregada uma arquitetura do tipo *3D U-Net* em duas etapas — uma de segmentação grosseira, seguida de refinamento —, treinada com segmentações manuais detalhadas realizadas por especialistas em endodontia e radiologia. O desempenho do modelo foi avaliado por meio de métricas como *Dice Similarity Coefficient* (*DSC*), *Intersection over Union* (*IoU*), precisão, sensibilidade (*recall*) e *Hausdorff Distance*. Para primeiros molares mandibulares, o modelo alcançou $DSC \approx 88\% \pm 7\%$ e $IoU \approx 80\% \pm 12\%$; já para segundos molares, os valores foram $DSC \approx 90\% \pm 6\%$ e $IoU \approx 82\% \pm 10\%$. Não houve diferença estatisticamente significativa entre o desempenho da IA e o da segmentação refinada por humanos (*Refined-AI*) para a maioria dos parâmetros avaliados. Destaca-se, contudo, que a IA superou o desempenho humano em algumas métricas, como *recall* e *Hausdorff Distance*, além de apresentar tempo de segmentação substancialmente menor — cerca de $4,3 \pm 2$ segundos por molar, em comparação a 139 ± 93 segundos do *Refined-AI* e 2.349 ± 444 segundos da segmentação manual completa.

Por fim, Merdietio Boedi et al. (2024) realizaram um estudo voltado à predição da idade cronológica com base em segmentações volumétricas de 80 dentes anteriores superiores (incisivos centrais, laterais e caninos). A segmentação foi subdividida em

cinco regiões anatômicas — esmalte, dentina, câmara pulpar, raiz inferior e volume total —, sendo as razões volumétricas utilizadas como variáveis preditoras em diferentes modelos de regressão. Os algoritmos testados incluíram *Support Vector Regression (SVR)* com núcleos linear e polinomial, árvore de regressão e regressão linear múltipla. O melhor desempenho foi obtido com o *SVR* polinomial aplicado ao incisivo lateral superior, apresentando erro médio absoluto (*Mean Absolute Error – MAE*) de 4,86 anos, erro quadrático médio (*Root Mean Square Error – RMSE*) de 6,05 anos e coeficiente de determinação (R^2) de 0,73. Apesar dos resultados promissores, o processo de segmentação manual ainda demandava aproximadamente 45 minutos por dente, representando um obstáculo à aplicabilidade clínica imediata.

3.2.3 Reabsorções radiculares

Pirayesh et al. (2024) conduziram um estudo com 50 exames de tomografia computadorizada de feixe cônico (TCFC) de incisivos maxilares ($n = 176$), incluindo dentes saudáveis ($n = 142$) e com reabsorção radicular induzida por caninos impactados ($n = 34$). As imagens foram segmentadas manualmente por dois radiologistas, que também realizaram a rotulagem dos casos. Foram testados cinco modelos distintos, com destaque para arquiteturas híbridas que combinavam segmentação e classificação tridimensional. Os melhores resultados foram obtidos pelos modelos C (*máscaras geradas por 3D U-Net + classificação com 3D ResNet*) e E (*U-Net com transfer learning e fine-tuning*), ambos alcançando acurácia de aproximadamente 82% na detecção de reabsorção radicular induzida por caninos impactados (*CIRR*). A segmentação automática, realizada com um modelo *3D U-Net*, apresentou *Dice Similarity Coefficient (DSC)* médio de $\sim 0,901$ e *mean Intersection over Union (mIoU)* de $\sim 0,641$, demonstrando elevada capacidade de identificação anatômica. O processo automático mostrou-se extremamente eficiente, exigindo apenas 7 a 30 milissegundos por volume, em contraste com os 11 a 12 minutos por imagem necessários para a segmentação manual.

Em um estudo clínico prospectivo, Estrella et al. (2025) avaliaram a perda de estrutura radicular associada à aplicação de diferentes forças ortodônticas. Foram comparados dois grupos de pacientes submetidos a fios convencionais e fios com forças individualizadas, com acompanhamento médio de 142 dias. A análise foi

realizada por meio de TCFC, com segmentações manuais e automáticas das raízes dos incisivos superiores. A IA foi implementada via software *Diagnocat*, e os resultados indicaram que, na maioria das regiões radiculares, não houve diferença estatisticamente significativa entre os métodos de segmentação. Entretanto, na região mais apical (1 mm do ápice), o método automático detectou uma perda volumétrica significativamente maior no grupo experimental ($p = 0,011$), sugerindo maior sensibilidade do algoritmo nessa área crítica.

No estudo experimental de Reduwan et al. (2024), 88 pré-molares extraídos foram utilizados para simular reabsorções radiculares externas (*External Root Resorptions – ERR*) de diferentes profundidades (0,5 mm, 1,0 mm e 2,0 mm). Foram testadas combinações de redes convolucionais pré-treinadas (*VGG16* e *EfficientNetB4*) com classificadores como *Random Forest (RF)* e *Support Vector Machine (SVM)*, com e sem o uso de técnicas de seleção de características (*feature selection*). O modelo *VGG16 + RF* com *feature selection (FST)* apresentou o melhor desempenho geral, com acurácia de 81,9%, *F1-score* de 81,9% e *Area Under the Curve (AUC)* de 96%. De modo consistente, os modelos que utilizaram *feature selection* superaram significativamente os que não a empregaram, reforçando a importância dessa etapa na otimização de modelos preditivos. Além disso, observou-se que a acurácia aumentava proporcionalmente à profundidade da reabsorção, sendo as lesões mais profundas (2,0 mm) mais facilmente detectadas.

Por fim, Xu et al. (2025) realizaram um estudo em larga escala com 2.146 fatias de TCFC contendo incisivos submetidos ou não a tratamento ortodôntico, com e sem sinais de reabsorção radicular. Foram testadas seis arquiteturas de redes convolucionais pré-treinadas (*EfficientNet-B1 a B5* e *MobileNet-V3*), sendo o modelo *EfficientNet-B1* aquele de melhor desempenho, com acurácia de 97%, precisão de 98%, sensibilidade de 97%, especificidade de 98% e *AUC* de 0,99. Esse modelo superou significativamente o desempenho de dois ortodontistas com aproximadamente dois anos de experiência (acurácia média de 86%, *F1-score* de 87%), com diferença estatisticamente significativa ($p < 0,01$). O uso de mapas de ativação (*Grad-CAM*) demonstrou que a rede neural concentrou sua atenção nas regiões apicais das raízes ao classificar casos positivos, o que corrobora a

localização típica das lesões de reabsorção radicular externa. O tempo médio de processamento também foi altamente eficiente, com 4,66 milissegundos por fatia.

3.2.4 Canais radiculares

Santos-Junior et al. (2025) investigaram a performance de um modelo de segmentação automática baseado na arquitetura *3D U-Net* aplicado a imagens de tomografia computadorizada de feixe cônico (TCFC) de dentes unirradiculares. Foram analisados 69 exames obtidos em diferentes equipamentos e protocolos de aquisição, com o intuito de incluir variabilidade no conjunto de dados e avaliar a robustez do modelo. A segmentação manual, realizada por dois operadores e revisada em consenso, serviu como *ground truth* para o treinamento e validação. No conjunto de teste, foi também conduzida uma etapa de refinamento por especialista após a segmentação automática (*Refined-AI – R-AI*). O modelo automático apresentou excelente desempenho, com *Dice Similarity Coefficient (DSC)* variando entre ~89% e 93% e *Hausdorff Distance (HD95)* entre ~0,10 mm e 0,13 mm. Outras métricas, como *Intersection over Union (IoU)*, precisão e acurácia, também mostraram valores elevados. O tipo de dente (incisivo, canino ou pré-molar) não influenciou significativamente a maioria das métricas, embora incisivos laterais superiores tenham apresentado leve decréscimo no *recall*, sugerindo que o modelo é robusto frente à variação anatômica dentro da categoria de dentes unirradiculares. A segmentação manual demandou cerca de 2.262 segundos ($\approx$ 37–38 minutos) por dente, enquanto a segmentação automática foi substancialmente mais rápida ($\sim$ 41,8 segundos). Quando associada ao refinamento humano (*R-AI*), o tempo aumentou para aproximadamente 94 segundos, ainda representando redução expressiva no tempo total de processamento, sendo que os ajustes necessários após a segmentação automática foram mínimos.

No estudo de Shetty et al. (2025), o foco foi a detecção do canal *MB2* em molares superiores utilizando modelos clássicos de classificação aplicados a imagens de TCFC. Foram avaliados 2.500 exames, com 277 casos rotulados por especialistas experientes (160 com presença do canal *MB2* e 117 sem). Seis algoritmos foram comparados: *Logistic Regression (LR)*, *Naïve Bayes (NB)*, *Support*

Vector Machine (SVM), *K-Nearest Neighbours (KNN)*, *Random Forest (RF)* e uma *Neural Network (NN)* de arquitetura simples. Os modelos que apresentaram melhor desempenho foram a rede neural ($AUC = 0,903$; acurácia = 0,838), a regressão logística ($AUC = 0,885$; acurácia = 0,841) e o *SVM* ($AUC = 0,886$; acurácia = 0,819), com métricas semelhantes entre si em termos de *F1-score*, precisão e sensibilidade. O modelo *KNN* apresentou o pior desempenho, sendo superado por todos os demais. Além dos resultados promissores dos modelos de IA, o estudo também evidenciou alta reprodutibilidade, com coeficientes de confiabilidade interobservador de 0,85 e intraobservador de 0,92 e 0,96 para os dois examinadores, respectivamente.

3.2.5 Localização de raízes em relação ao canal mandibular

No estudo de Yasin et al. (2025), foi proposto um sistema automatizado para classificar a relação entre as raízes dos molares mandibulares e o canal mandibular em exames de tomografia computadorizada de feixe cônico (TCFC). Foram analisados 305 cortes parasagitais, divididos em três categorias clínicas: sem contato, quase contato e em contato, com base na relação espacial entre a raiz e o canal mandibular. As imagens foram anotadas por especialistas em radiologia bucomaxilofacial com dupla verificação, assegurando alto nível de confiabilidade no *ground truth*. Diversas arquiteturas de redes neurais convolucionais (*Convolutional Neural Networks – CNNs*) foram avaliadas, incluindo *MobileNet*, *Xception*, *DenseNet201*, *InceptionV3*, *VGG16* e *VGG19*. O modelo *MobileNet* apresentou o melhor desempenho global, alcançando acurácia de 99,44% sob as condições ideais de treinamento (taxa de aprendizado de 0,0001 e *batch size* de 16). Outras arquiteturas, como *Xception* e *DenseNet201*, também obtiveram resultados elevados, com acurácias de 98,74% e 98,73%, respectivamente. Além das altas taxas de acurácia, as métricas de precisão, *recall* e *F1-score* também apresentaram valores elevados, indicando excelente capacidade discriminativa. A análise de interpretabilidade, conduzida por meio de *Grad-CAM*, demonstrou que os modelos concentraram sua atenção em regiões anatômicas clinicamente relevantes — como a raiz molar, a cortical óssea e o canal mandibular — em consonância com o raciocínio diagnóstico humano.

Por sua vez, Picoli et al. (2023) avaliaram o desempenho de uma ferramenta tridimensional baseada em inteligência artificial (IA) para estimar o risco de lesão do nervo alveolar inferior (*Inferior Alveolar Nerve – IAN*) antes da remoção dos terceiros molares inferiores. O estudo incluiu 25 pacientes submetidos à remoção bilateral dos terceiros molares, dos quais apenas um lado apresentou lesão do *IAN* no pós-operatório. Foram comparadas três modalidades de imagem: panorâmica digital (*PANO*), TCFC e um modelo 3D gerado por IA (plataforma *Virtual Patient Creator AI*). Cinco examinadores, cegos quanto ao desfecho, atribuíram o risco de lesão em cada lado como baixo, médio ou alto. A sensibilidade — capacidade de identificar corretamente os casos com lesão — foi de 0,87 para o modelo 3D-IA, 0,89 para a TCFC e 0,73 para a *PANO*. A área sob a curva *Receiver Operating Characteristic* (*AUC-ROC*) foi maior para o modelo 3D-IA (0,63), em comparação à TCFC (0,58) e à *PANO* (0,57), indicando leve superioridade na capacidade global de discriminação. As especificidades foram relativamente baixas para todas as modalidades, sendo 0,39 (3D-IA), 0,28 (TCFC) e 0,41 (*PANO*), o que sugere uma tendência dos avaliadores à superestimação do risco. Entretanto, as diferenças entre as modalidades não foram estatisticamente significativas ($p > 0,05$).

3.2.6 Patologias dentárias

No estudo de Ezhov et al. (2021), foram utilizados 1.346 exames de tomografia computadorizada de feixe cônico (TCFC) para o treinamento de um sistema automatizado voltado à detecção de múltiplas alterações dentárias e periodontais. A avaliação clínica envolveu 24 cirurgiões-dentistas, que analisaram 30 exames em dois grupos experimentais: com e sem o auxílio da inteligência artificial (IA). O uso da IA resultou em melhora significativa do desempenho diagnóstico dos profissionais: a sensibilidade média aumentou de 0,7672 (sem IA) para 0,8537 (com IA), enquanto a especificidade passou de 0,9616 para 0,9672 ($p = 0,032$). Além disso, a IA reduziu o tempo médio de avaliação por exame, automatizando tarefas como a detecção preliminar de lesões e a geração de mapas dentários.

Quanto à performance técnica do sistema, os resultados foram igualmente expressivos: para lesões periapicais, sensibilidade = 0,903 e especificidade = 0,979; para periodontite/alveólise, sensibilidade = 0,878 e especificidade = 0,963; e para cáries moderadas a avançadas, sensibilidade $\approx$ 0,85 e especificidade $>$ 0,95. A

segmentação dentária automática alcançou acurácia de 96,7% na identificação e numeração de dentes, mesmo em casos com restaurações ou reabsorções. Entretanto, o sistema apresentou limitações em situações mais complexas, como lesões iniciais de cárie, fraturas radiculares finas e reabsorções atípicas, além de sofrer impacto negativo de artefatos metálicos na acurácia das predições.

Em outro estudo clínico, Kazimierczak et al. (2024) aplicaram a plataforma *Diagnocat* para avaliar parâmetros técnicos de tratamentos endodônticos em 55 pacientes. A IA foi comparada ao julgamento de dois especialistas experientes, considerados padrão-ouro por consenso. O sistema apresentou desempenho excelente na detecção de obturações radiculares, com 100% de acurácia, precisão, sensibilidade e *F1-score* na identificação da presença de material obturador. Para parâmetros mais específicos — como densidade da obturação, sobrepreenchimento e presença de vacúolos —, o desempenho manteve-se elevado, com *F1-scores* variando entre 76% e 97%, dependendo do critério avaliado. A concordância com os especialistas humanos também foi alta na identificação do número de canais radiculares, demonstrando a confiabilidade da ferramenta.

Complementando essa perspectiva, Amasya et al. (2023) investigaram o impacto do uso de IA na detecção de cáries em 500 volumes de TCFC, abrangendo mais de 19.000 superfícies dentárias, das quais cerca de 6.000 apresentavam lesões cariosas. As imagens foram avaliadas por radiologistas dentomaxilofaciais em dois momentos: sem e com o auxílio da IA (*Diagnocat*). O uso da IA resultou em aumento significativo na acurácia diagnóstica — a área sob a curva *Receiver Operating Characteristic* (*AUC-ROC*) dos observadores passou de 0,747–0,863 (sem IA) para 0,903–0,920 (com IA). A sensibilidade aumentou de forma consistente, enquanto a especificidade melhorou na maioria dos casos, apresentando queda discreta apenas em um observador. A concordância interobservador também foi positivamente impactada, com aumento do coeficiente *Fleiss' kappa* de aproximadamente 0,325 (sem IA) para 0,468 (com IA), tanto na escala de confiança quanto nas decisões binárias de presença ou ausência de cárie.

3.2.7 Fraturas dentárias

No estudo conduzido por Johari et al. (2017), foram avaliados 240 dentes pré-molares com raiz única, distribuídos igualmente entre dentes com fraturas verticais induzidas artificialmente e dentes controle. Em cada grupo, metade dos dentes havia recebido tratamento endodôntico, enquanto a outra metade permaneceu intacta. Para a análise, foi utilizada uma rede neural do tipo *Multilayer Perceptron (MLP)*, que apresentou desempenho elevado, com acurácia máxima de 96,6%, sensibilidade de 93,3% e especificidade de 100%. Esses resultados demonstram que a abordagem baseada em inteligência artificial (IA) é altamente eficaz na identificação de fraturas, inclusive em dentes endodonticamente tratados — um cenário clínico tipicamente mais desafiador devido à presença de materiais obturadores e alterações morfológicas pós-tratamento.

De forma complementar, Yang et al. (2023) aplicaram redes neurais convolucionais profundas (*Deep Convolutional Neural Networks – CNNs*) para a detecção de fraturas radiculares verticais (*Vertical Root Fractures – VRFs*) em imagens de tomografia computadorizada de feixe cônico (TCFC), utilizando dois conjuntos de dados distintos: um *in vivo*, composto por 1.641 cortes de TCFC de 28 dentes de pacientes (14 com VRF e 14 controles), e outro *in vitro*, com 3.665 cortes de TCFC de 60 dentes humanos secos com fraturas induzidas. Foram comparadas três variantes da arquitetura *ResNet* (*ResNet-18*, *ResNet-50* e *ResNet-101*), incluindo avaliação comparativa com radiologistas experientes. O modelo *ResNet-50* apresentou o melhor desempenho, alcançando, no conjunto misto (*in vivo* + *in vitro*), sensibilidade de aproximadamente 95,9%, especificidade de 74,1%, acurácia de 84,0% e área sob a curva *Receiver Operating Characteristic (AUC)* de 0,936. Quando avaliado exclusivamente com dados clínicos (*in vivo*), os resultados permaneceram semelhantes, com sensibilidade de 94,5%, especificidade de 73,2% e *AUC* de 0,929, indicando boa capacidade de generalização do modelo entre diferentes contextos de aplicação.

3.3 Ossos

Foram selecionados 22 estudos que abordaram o uso de inteligência artificial (IA) aplicada à análise de ossos em exames de tomografia computadorizada de feixe cônico (TCFC). Os subtemas são: fendas palatinas (n = 2), identificação de osteonecrose (n = 1), lesões de furca (n = 1), segmentação do osso alveolar (n = 6),

segmentação do osso hioide ($n = 1$), avaliação da sutura palatina ($n = 2$), lesões periapicais ($n = 4$), densidade óssea ($n = 3$), segmentação da mandíbula ($n = 1$), além de um estudo combinando dentes e segmentação óssea.

3.3.1 Fendas palatinas

No estudo conduzido por Wang et al. (2021), cerca de 60 arquivos *DICOM* de pacientes portadores de fissura palatina foram utilizados para o desenvolvimento e avaliação de um modelo de segmentação automática baseado na arquitetura *3D U-Net*. Os casos foram divididos em dois grupos, sendo o Grupo 1 segmentado manualmente e empregado como conjunto de treinamento do algoritmo. Os resultados evidenciaram alta concordância entre as segmentações automáticas e as referências manuais, com coeficiente de similaridade (*Dice*) de $0,92 \pm 0,01$ para a maxila e $0,77 \pm 0,06$ para o defeito ósseo. Além da elevada acurácia, observou-se uma redução expressiva no tempo de processamento, passando de aproximadamente 10 horas por imagem no processo manual para cerca de 1 minuto na etapa automática, acrescido de um refinamento manual médio de 5 minutos.

Por sua vez, Miranda et al. (2023) empregaram 194 exames de tomografia computadorizada de feixe cônico (TCFC) de pacientes com fissura labiopalatina (*Cleft Lip and Palate – CLP*) para desenvolver um classificador *multi-view* voltado à avaliação da gravidade do defeito ósseo alveolar. As superfícies tridimensionais do maxilar foram segmentadas automaticamente, possibilitando a geração de modelos 3D do defeito, a partir dos quais foram extraídas características de forma, altura e largura. Com base nesses parâmetros, foi estabelecido um índice de gravidade variando de 0 a 3. O modelo *multi-view* foi composto por renderizações bidimensionais (*snapshots*) dos modelos tridimensionais obtidas de múltiplos ângulos, processadas por uma rede neural convolucional 2D (*2D CNN*) equipada com camadas de atenção responsáveis pela integração das informações provenientes de cada vista.

O desempenho do classificador foi considerado satisfatório, apresentando acurácia geral de 0,816, precisão de 0,823, *recall* de 0,816, *F1-score* de 0,817 e área sob a curva *Receiver Operating Characteristic (AUC)* de 0,948. A concordância entre as classificações automáticas e as avaliações de referência (*ground truth*)

mostrou-se elevada, variando entre 97,4% e 100% quando admitida discrepância de um nível no índice de gravidade. Além disso, os mapas de explicação (*SurfGradCAM*) demonstraram que o modelo concentrou sua atenção em regiões anatômicas compatíveis com a patologia, como a pré-maxila, o contorno palatal e as paredes bucal e palatina do defeito, reforçando a interpretabilidade clínica da abordagem.

3.3.2 Segmentação de osso alveolar

O estudo de Cui et al. (2022) avaliou um sistema de segmentação automática do osso alveolar utilizando 4.938 exames de tomografia computadorizada de feixe cônico (TCFC) provenientes de 4.215 pacientes, distribuídos em 15 centros clínicos distintos. Foram conduzidos testes em conjuntos internos (dados provenientes de hospitais incluídos no treinamento) e externos (novos centros e perfis populacionais diferentes), além de uma comparação direta com radiologistas experientes para validação clínica. O método apresentou coeficiente *Dice* médio de aproximadamente 94,5% e distância média de superfície (*Average Surface Distance – ASD*) de 0,33 mm, indicando alta precisão na delimitação do osso alveolar. O desempenho manteve-se robusto mesmo em situações desafiadoras, como ausência ou desalinhamento dentário, presença de artefatos metálicos (implantes e restaurações) e variações entre protocolos e equipamentos de TCFC. Em termos de eficiência, a diferença em relação à prática manual foi expressiva: enquanto os radiologistas demandavam entre 147 e 169 minutos por exame para a segmentação completa, o sistema automático realizava o mesmo processo em cerca de 17 segundos, com ajustes manuais inferiores a 5 minutos — correspondendo a uma redução média de 96–97% no tempo total de anotação.

No trabalho de Palkovics et al. (2025), uma rede *SegResNet* foi treinada com 70 exames de TCFC (57 para treinamento e 13 para validação) e testada em 10 pares de exames pré e pós-operatórios de pacientes submetidos à regeneração óssea guiada (*Guided Bone Regeneration – GBR*) horizontal da mandíbula. A segmentação semiautomática (*Semi-Automatic Segmentation – SA*) realizada por avaliador calibrado serviu como referência (*ground truth*). Para a mandíbula completa, os resultados indicaram *Dice Similarity Coefficient (DSC)* $\approx 0,96 \pm 0,01$, *Intersection over Union (IoU)* $\approx 0,92 \pm 0,02$ e *Hausdorff Distance 95% (HD95)* $\approx 0,62\text{--}0,77$ mm,

tanto no período pré quanto no pós-operatório. A diferença volumétrica entre as segmentações por *Deep Learning* (DL) e SA não foi estatisticamente significativa, demonstrando boa equivalência entre os métodos. Para o tecido ósseo regenerado, a sobreposição foi menor ($DSC \approx 0,85 \pm 0,08$; $IoU \approx 0,78 \pm 0,07$; $HD95 \approx 0,91 \pm 0,92$ mm), mas sem diferença significativa no volume obtido ($p \approx 0,06$). Em relação ao tempo, a segmentação por DL foi consideravelmente mais rápida (52–54 s) em comparação à SA (≈ 42 min).

Kurt-Bayrakdar et al. (2025) utilizaram 502 volumes de TCFC provenientes de duas instituições, aplicando a arquitetura *nnU-Net v2*, ajustável automaticamente ao conjunto de dados. O modelo foi posteriormente treinado com 251 volumes de pacientes com doença periodontal, com 10% dos dados reservados para validação. As áreas sob a curva (AUCs) variaram conforme o tipo de lesão: perda óssea alveolar total $\approx 0,85$; lesões *perio-endo* $\approx 0,89$; defeitos supraósseos $\approx 0,51$; infraósseos $\approx 0,56$; defeitos em furca $\approx 0,63$; e defeitos bucais $\approx 0,68$. Na classificação geral entre pacientes saudáveis e com doença periodontal, o modelo apresentou precisão média de 80% e 76%, respectivamente, com sensibilidade, especificidade e *F1-score* variando entre 0,76 e 0,80.

Em Fontenele et al. (2023), 141 exames de TCFC foram analisados para testar uma *Convolutional Neural Network* (CNN) de segmentação automática, com refinamento posterior por especialista (*Refined-AI*). Em cerca de 30% dos casos, também foi realizada segmentação manual para comparação. O método automático atingiu *Hausdorff* $\approx 0,27 \pm 0,03$ mm, $IoU \approx 92\% \pm 1\%$ e $DSC \approx 96\% \pm 1\%$, resultados ligeiramente inferiores aos obtidos manualmente (*Hausdorff* $\approx 0,20 \pm 0,05$ mm; $IoU \approx 95\% \pm 3\%$; $DSC \approx 97\% \pm 2\%$). O principal ganho, entretanto, foi em eficiência: o processamento automático consumiu aproximadamente 51,5 segundos por exame, contra cerca de 1 h 40 min (≈ 5.973 s) na segmentação manual — uma aceleração de aproximadamente 116 vezes.

O estudo de Al-Asali et al. (2024) empregou 150 volumes de TCFC para treinar duas redes *U-Net*: uma voltada à segmentação do osso na região de dente ausente e outra à predição da região de interesse (*Region of Interest – ROI*) para implante. O desempenho foi elevado, com *Dice* $\approx 0,93$, precisão $\approx 0,94$ e *recall* $\approx 0,93$, além de erro volumétrico médio de apenas 1%.

De forma semelhante, Moufti et al. (2023) utilizaram 43 exames de TCFC de mandíbulas parcialmente desdentadas, aplicando uma *U-Net* para segmentar o espaço ósseo entre dentes remanescentes. O modelo apresentou *DSC* médio de 0,89 no treinamento e 0,78 no teste, com melhor desempenho nos casos unilaterais (*DSC* $\approx$ 0,91) em comparação aos bilaterais (*DSC* $\approx$ 0,73).

Por fim, Al-Sarem et al. (2022) processaram 500 imagens de TCFC, convertendo os volumes tridimensionais em projeções bidimensionais (*2D panorâmicas*). Após anotações manuais por especialista, foram testados diferentes modelos *CNN* pré-treinados (*AlexNet*, *VGG16*, *VGG19*, *ResNet50*, *DenseNet169* e *MobileNetV3*), com e sem segmentação prévia via *U-Net*. O *DenseNet169*, quando combinado à segmentação, apresentou o melhor desempenho, com acurácia de aproximadamente 93,3% e *F1-score* próximo de 0,94 na detecção de dentes ausentes. O uso da segmentação prévia melhorou significativamente as métricas de precisão, *recall* e *F1-score* em todos os modelos, evidenciando o impacto positivo da pré-segmentação automatizada na classificação de estruturas dentárias em imagens de TCFC.

3.3.3 Osteonecrose dos maxilares associada a medicamentos (OMAM)

No estudo de Gürses et al. (2023), foram analisados exames de tomografia computadorizada de feixe cônico (TCFC) de sete pacientes com osteonecrose dos maxilares associada a medicamentos (OMAM) em estágio 0 — caracterizado pela ausência de exposição óssea clínica — e de oito indivíduos saudáveis, utilizados como grupo controle. A amostra totalizou 55 fatias tomográficas (28 de pacientes com OMAM e 27 de controles). O processo de rotulagem foi conduzido por três radiologistas bucomaxilofaciais experientes, que identificaram manualmente quatro categorias anatômicas: osso saudável (controles), esclerose óssea (OMAM), necrose óssea (OMAM) e osso de aparência normal em pacientes com OMAM.

A análise estatística dos momentos de textura revelou diferenças significativas entre os grupos saudáveis e doentes (esclerose + necrose), evidenciando alterações microestruturais detectáveis nas imagens mesmo em estágios iniciais da doença. Observou-se também diferença estatisticamente significativa entre o osso visualmente normal dos pacientes com OMAM e o osso saudável dos controles,

indicando que mudanças sutis, ainda não perceptíveis visualmente, já se manifestam nas propriedades estatísticas da imagem.

O desempenho do modelo de classificação foi notavelmente elevado: utilizando apenas três medidas estatísticas simples (como média, moda e mediana), o sistema alcançou acurácia de aproximadamente 0,969 e área sob a curva *Receiver Operating Characteristic* (*ROC AUC*) de 0,995. Quando ampliado para oito parâmetros estatísticos, o desempenho atingiu acurácia, *AUC*, *F1-score* e sensibilidade de 0,99, demonstrando altíssimo poder discriminativo entre tecidos ósseos normais e alterados — inclusive em fases subclínicas da osteonecrose.

3.3.4 Identificação de lesões de furca

No estudo de Shetty et al. (2024), foram analisadas 285 tomografias computadorizadas de feixe cônico (TCFC), das quais 143 não apresentavam comprometimento de furca e 142 exibiam lesões nessa região. O objetivo foi desenvolver e avaliar um modelo de classificação automática do envolvimento de furca em imagens tomográficas. Foram testadas diferentes variantes da arquitetura *ResNet* — *ResNet50*, *ResNet101* e *ResNet101V2* —, sendo esta última selecionada como modelo final, empregando *transfer learning* e *fine-tuning* para otimização do desempenho.

A rotulagem das imagens foi conduzida por dois periodontistas experientes, que analisaram as fatias axiais dos exames de TCFC e determinaram a presença ou ausência de comprometimento de furca. O modelo apresentou acurácia de aproximadamente 98% no treinamento, 97% na validação e 91% no conjunto de teste (*dados não vistos*). As demais métricas de desempenho também foram elevadas, incluindo precisão de 0,98, *F1-score* de 0,98, *AUC* de 0,98 e *loss* no teste de 0,2170.

Em termos de desempenho clínico, o modelo demonstrou sensibilidade de 89%, identificando corretamente a maioria dos casos com lesão de furca (baixo índice de falsos negativos), e especificidade de 94%, reconhecendo adequadamente os casos sem envolvimento de furca (baixo índice de falsos positivos).

3.3.5 Segmentação do osso hioide

No estudo de Gümüşsoy et al. (2025), foram analisados 190 exames retrospectivos de tomografia computadorizada de feixe cônico (TCFC) com o objetivo de desenvolver um modelo de segmentação automática do osso hioide (OH). Foram incluídas apenas imagens que apresentavam margens anatômicas bem definidas do osso hioide, sendo excluídos casos com trauma cervical, artefatos metálicos ou processos estilóides fundidos ao hioide, bem como outras condições que pudessem comprometer a acurácia anatômica da estrutura.

As imagens originais foram convertidas para o formato *DICOM* e submetidas à segmentação manual no software *3D Slicer* por dois radiologistas experientes. As anotações divergentes foram resolvidas por consenso, estabelecendo o *ground truth* utilizado para o treinamento e validação do modelo. A rede *nnU-Net* v2 foi empregada, aproveitando sua capacidade de autoajuste aos parâmetros do conjunto de dados e sua arquitetura tridimensional baseada em *U-Net*.

Os resultados demonstraram alto desempenho de segmentação, com coeficiente *Dice* (*DC*) de aproximadamente 0,94, *Intersection over Union* (*IoU*) de 0,89 e *F1-score* de 0,94. A distância *Hausdorff* a 95% (*HD95*) foi de cerca de 2,00 mm, indicando excelente correspondência espacial entre as segmentações automáticas e o *ground truth*. Além disso, a área sob a curva *Receiver Operating Characteristic* (*AUC*) atingiu 0,98, evidenciando forte capacidade discriminativa entre *voxels* pertencentes e não pertencentes ao osso hioide. O modelo também apresentou alta precisão (0,93) e *recall* (0,96), demonstrando baixo índice de falsos positivos e falsos negativos.

3.3.6 Maturação da sutura palatina

No estudo de Milani et al. (2025), foram analisados 618 cortes axiais de tomografia computadorizada de feixe cônico (TCFC), previamente rotulados por especialistas segundo os cinco estágios de maturação da sutura palatina média definidos pela classificação de Angelieri et al. (A–E), em que o estágio A representa o início do desenvolvimento e o estágio E indica a sutura completamente fusionada. Foram comparadas arquiteturas de redes neurais profundas amplamente utilizadas — *ResNet18* e *EfficientNet* — com uma nova proposta denominada *MFCRAN* (*Multi-Filter Convolutional Residual Attention Network*). Essa arquitetura incorpora

camadas de atenção residual, filtros convolucionais múltiplos e uma camada de *Discrete Cosine Transform (DCT)*, projetada para capturar componentes de frequência que auxiliam na diferenciação entre os estágios suturais.

O modelo proposto (*MFCRAN* com camada *DCT* e função de perda marginada baseada em *ranking*) apresentou o melhor desempenho geral, alcançando acurácia de 79,02%, precisão de 78,95%, *recall* de 78,87% e *F1-score* de 78,79%. Os modelos comparativos apresentaram resultados inferiores: *ResNet18* obteve acurácia de 73,33% e *EfficientNet_b0*, 71,87%. As principais confusões de classificação ocorreram entre estágios adjacentes (A–B e D–E), o que reflete a transição morfológica gradual entre essas fases e a dificuldade diagnóstica, mesmo em avaliações clínicas humanas.

A análise de interpretabilidade das redes, realizada por meio da técnica *Grad-CAM*, revelou que o modelo concentrou sua atenção principalmente na sutura palatina média e em estruturas ósseas adjacentes, confirmando coerência com as regiões de interesse utilizadas por ortodontistas durante a avaliação manual.

De forma complementar, Tang et al. (2024) desenvolveram um modelo de classificação baseado em *Vision Transformer (ViT_B/16)* aprimorado com blocos convolucionais multiescala e sobrepostos, com o objetivo de combinar a capacidade global de atenção dos *Transformers* com a extração local de características espaciais típica das *CNNs*. Foram utilizadas 2.518 imagens de TCFC do plano palatino contendo a região da sutura palatina média. O modelo híbrido proposto apresentou acurácia de aproximadamente 95,75% no conjunto de teste, *AUC* média global (*Macro-AUC*) de 97,89% e *Micro-AUC* de 98,36%, superando o desempenho dos modelos puramente convolucionais.

Entre as arquiteturas *CNN* testadas — *MobileNetV2*, *ResNet50*, *ResNet18*, *ResNet101*, *InceptionV3*, *EfficientNetV2_S*, *EfficientNet_B4* e *VGG16* —, o melhor desempenho foi obtido pela *EfficientNetV2_S*, com acurácia de cerca de 93,76%, ainda assim inferior ao modelo baseado em *Transformer*. Em comparação, a avaliação clínica manual por ortodontistas atingiu acurácia média de 89,10%, demonstrando que o modelo *ViT* aprimorado superou tanto as *CNNs* convencionais

quanto o desempenho humano na classificação dos estágios de maturação da sutura palatina.

3.3.7 Lesões periapicais

No estudo de Calazans et al. (2022), foram utilizados 1.000 pares de cortes sagitais e coronais de imagens de tomografia computadorizada de feixe cônico (TCFC), classificados em três categorias: dentes saudáveis, dentes com lesões pequenas (0,5–1,9 mm) e dentes com lesões grandes ($\geq 2,0$ mm). Foram testados dois modelos de redes neurais profundas — *VGG-16* e *DenseNet-121* —, ambos com *transfer learning*, integrados em uma arquitetura *Siamese Concatenation Network*, que combina as duas projeções (sagital e coronal) de um mesmo dente para extrair características tridimensionais conjuntas.

O desempenho foi avaliado em três cenários distintos. No cenário da base completa (todas as categorias), o modelo *DenseNet-121* alcançou acurácia em torno de 70,00%, com métricas complementares (*F1-score*, especificidade e precisão) de valores semelhantes, enquanto o *VGG-16* apresentou resultados ligeiramente inferiores. No cenário “saudável vs. grandes lesões”, o desempenho foi superior, atingindo acurácia de 81,25% com o *VGG-16* e especificidade de 92,39% com o *DenseNet-121*, indicando boa capacidade para identificar corretamente dentes sem lesão. Já no cenário “saudável vs. pequenas lesões”, o desempenho foi mais modesto (acurácia $\approx 66,67\%$ com *DenseNet-121*), com *F1-score* e *recall* menores, refletindo a maior dificuldade do modelo em distinguir lesões sutis.

O estudo de Fu et al. (2024) propôs a rede *PAL-Net*, uma *convolutional neural network* (CNN) tridimensional desenvolvida para detecção e segmentação de lesões periapicais em TCFC, avaliada por meio de validação cruzada de cinco dobras (*5-fold cross-validation*) e validação externa em três regiões geográficas da China (Central, Leste e Norte). O conjunto de validação contou com 100 imagens de TCFC, avaliadas também por 20 dentistas de diferentes níveis de experiência, com e sem o auxílio da *PAL-Net*. O modelo apresentou *AUC* interna de 0,98 e coeficiente de similaridade (*Dice*) médio de 0,87 (variação 0,85–0,88), demonstrando excelente desempenho de segmentação.

O uso assistido do modelo resultou em melhora significativa do desempenho diagnóstico humano: dentistas juniores tiveram *AUC* aumentada de 0,89 para 0,94, e seniores, de 0,91 para 0,93. Além disso, observou-se redução expressiva no tempo de diagnóstico, com média de 69,3 minutos a menos para profissionais juniores e 32,4 minutos a menos para seniores. A validação externa demonstrou boa generalização do modelo entre diferentes regiões.

Em um estudo retrospectivo, Allihaibi et al. (2025) analisaram 134 molares (327 raízes) obtidos por TCFC, comparando os resultados do software *Diagnocat* com a avaliação de dois endodontistas experientes, considerada o padrão de referência. Para molares não tratados, a ferramenta apresentou alta sensibilidade (93,9% por dente e 86,2% por raiz), especificidade moderada (65,2% e 79,9%, respectivamente) e acurácia geral de 79,1% (por dente) e 82,6% (por raiz). O valor preditivo positivo (*VPP*) foi de 71,8% (dente) e 75,8% (raiz), enquanto o valor preditivo negativo (*VPN*) atingiu 91,8% (dente) e 88,8% (raiz). O *F1-score* foi de 81,3% por dente e 80,7% por raiz, e a *AUC-ROC* foi de 0,76 e 0,79, respectivamente. Para dentes tratados (pós-operatórios), o desempenho diminuiu, especialmente nos valores de *VPP* (54,2% por dente; 46,9% por raiz) e *F1-score* (67,2% e 59,2%), sugerindo maior dificuldade na detecção de lesões residuais após o tratamento endodôntico.

Por sua vez, Kazimierczak et al. (2024) avaliaram o desempenho do *Diagnocat* em 49 pacientes (1.223 dentes) que possuíam radiografias panorâmicas e TCFC obtidas com intervalo de até 30 dias. A avaliação humana, conduzida por um ortodontista e um radiologista, estabeleceu o padrão de verdade (*ground truth*) a partir do consenso entre ambas as modalidades de imagem. A inteligência artificial analisou as imagens panorâmicas e TCFC separadamente, classificando como “positivas” as detecções com probabilidade superior a 50% de lesão periapical. Os resultados mostraram que o desempenho da inteligência artificial foi substancialmente superior na TCFC, com sensibilidade de 77,78%, especificidade acima de 98%, acurácia geral de 99,35%, *VPP* de 91,30% e *F1-score* de aproximadamente 84,00%.

3.3.8 Determinação da densidade óssea

No estudo de Namatevs et al. (2023), foram incluídas 188 mulheres pós-menopausa (54–87 anos; média de $69,1 \pm 8,1$ anos), edentadas ou parcialmente edentadas, submetidas à tomografia computadorizada de feixe cônico (TCFC) no contexto de planejamento para implantes. A densidade mineral óssea (DMO) de referência foi obtida por *dual-energy X-ray absorptiometry (DEXA)* nas regiões lombar (L2–L4) e colo femoral, sendo o pior *T-score* utilizado para classificar as pacientes como normais, osteopênicas ou osteoporóticas. O modelo proposto baseou-se na *ResNet-101*, utilizada para três tarefas sequenciais: (1) classificação dos *slices* corretos, (2) detecção de pontos anatômicos de referência e (3) estimação da espessura cortical mandibular.

No Estágio 1 (classificação dos *slices*), o modelo alcançou acurácia de treinamento de 98,85% (no 39º *epoch*) e acurácia de validação de 93,99% (no 35º *epoch*), demonstrando alta consistência entre as fases. No Estágio 2 (detecção de pontos de referência) — incluindo cinco pontos mandibulares e dois pontos nos canais do nervo alveolar inferior —, a métrica de perda *L1 (mean absolute error)* foi de aproximadamente 1,02 *pixel* na validação, indicando precisão adequada na regressão espacial. Já no Estágio 3 (estimação da espessura cortical mandibular), considerando 180 observações rotuladas, a espessura média verdadeira foi de $2,97 \pm 0,96$ mm, e o erro médio quadrático (*MSE*) do algoritmo foi de 0,84, demonstrando boa correspondência entre os valores preditos e os dados de referência.

No estudo de Xiao et al. (2022), foram avaliadas imagens de TCFC de 70 pacientes com defeitos dentários mandibulares, rotuladas manualmente por quatro cirurgiões-dentistas experientes e revisadas por um especialista com mais de 20 anos de prática clínica. As regiões de interesse foram classificadas em cinco tipos de densidade óssea, de acordo com os valores de unidades *Hounsfield (HU)*: Tipo 1 (1000–2000), Tipo 2 (700–1000), Tipo 3 (400–700), Tipo 4 (100–400) e Tipo 5 (100–200). Foi empregada a arquitetura *Nested U-Net* para segmentação semântica, escolhida por sua capacidade de extrair e integrar informações em múltiplas escalas espaciais.

No conjunto de teste (68 casos), o modelo alcançou coeficiente de similaridade *Dice* médio de aproximadamente 0,75 para as diferentes categorias. Os valores médios de *HU* estimados automaticamente mostraram-se muito próximos aos

valores rotulados pelos especialistas e dentro dos intervalos esperados para cada tipo. Por exemplo, para o Tipo 1, os especialistas indicaram média de cerca de 1519 *HU*, enquanto o modelo previu aproximadamente 1520 *HU*; para o Tipo 5, a média foi de 195 *HU* (especialistas) e 136 *HU* (modelo). Além disso, os desvios-padrão intra-categoria foram semelhantes entre a rotulagem humana e a automática, indicando que o modelo conseguiu reproduzir adequadamente a variabilidade anatômica e de densidade observada nos tecidos ósseos mandibulares.

O estudo de Yong et al. (2021) propôs o *QCBCT-Net*, uma arquitetura híbrida voltada à quantificação precisa da densidade mineral óssea (DMO) em TCFC. O experimento utilizou dois crânios-fantomas de acrílico — um sem restaurações metálicas (para treinamento e validação) e outro com restaurações (para teste). Foram adquiridas imagens por tomografia computadorizada convencional e por TCFC em diferentes condições de dose (80 kVp/8 mA e 90 kVp/10 mA).

O *QCBCT-Net* combina um *Cycle-GAN* com blocos residuais, responsável por converter imagens de TCFC em representações semelhantes às *QCT*, e uma *U-Net* de múltiplos canais, que recebe simultaneamente a imagem original e a saída do *Cycle-GAN* para refinamento, redução de ruído e correção de artefatos. Quando comparado aos métodos de calibração tradicionais (*CAL_CBCT*, *CYC_CBCT* e *U_CBCT*), o *QCBCT-Net* apresentou ganhos substanciais de desempenho. Sob as condições de 80 kVp/8 mA, observaram-se, na maxila, aumentos de 38% em *MAD*, 20% em *PSNR*, 45% em *SSIM*, 40% em *NCC*, 80% em *SNU* e 84% em *Slope*, enquanto na mandíbula os incrementos foram de 39%, 20%, 50%, 40%, 47% e 102%, respectivamente. Em comparação ao modelo *U_CBCT*, as melhorias foram ainda mais expressivas (por exemplo, até +112% em *SSIM* e +167% em *Slope*). Em 90 kVp/10 mA, o *QCBCT-Net* manteve desempenho superior em todos os parâmetros, demonstrando robustez frente a diferentes condições de aquisição.

3.3.9 Segmentação da mandíbula

No estudo de Hernandez et al. (2025), foram avaliadas 12 mandíbulas humanas secas, com imagens adquiridas por tomografia computadorizada de feixe cônico (TCFC) para obtenção dos arquivos *DICOM*, e por *structured light scanner* (*GoSCAN Spark*) para geração dos modelos de superfície de referência, considerados o

padrão-ouro. Três protocolos de segmentação foram comparados: (1) manual, realizada por operador experiente; (2) limiarização global, baseada na seleção de um limiar de intensidade para separar os tecidos ósseos do restante da imagem; e (3) segmentação automática baseada em inteligência artificial (*Diagnocat*).

A avaliação consistiu na sobreposição dos modelos tridimensionais obtidos por cada método ao modelo de referência do *scanner* óptico, com cálculo do *Root Mean Square* (*RMS*) de erro entre as superfícies, utilizando o software de comparação de malhas *Geomagic Design X*. Valores menores de *RMS* indicam maior fidelidade da segmentação em relação ao padrão-ouro. As análises estatísticas foram realizadas por meio de *ANOVA* de medidas repetidas, considerando diferentes regiões anatômicas da mandíbula, incluindo corpo, ramo, côndilos, processo coronóide, sínfise e ângulo mandibular, entre outras.

Os resultados mostraram que as segmentações baseadas em inteligência artificial apresentaram valores de *RMS* significativamente mais altos que as segmentações manuais para a mandíbula inteira ($p < 0,001$), processo alveolar ($p < 0,001$) e corpo mandibular ($p < 0,001$), indicando menor precisão nessas regiões. Por outro lado, as segmentações por IA foram mais precisas que as manuais nos côndilos ($p = 0,018$) e no ramo mandibular ($p = 0,013$). Não foram observadas diferenças estatisticamente significativas entre IA e manual no processo coronóide ($p = 0,275$), sínfise ($p = 0,346$) e ângulo mandibular ($p = 0,344$).

Em relação à limiarização global, os valores de *RMS* foram significativamente mais altos que os obtidos manualmente em praticamente todas as regiões avaliadas — incluindo processo alveolar, corpo, ramo, côndilos, coronóide, sínfise e mandíbula inteira ($p < 0,05$ em todos os casos) —, evidenciando precisão inferior desse método. Na comparação entre limiarização global e IA, a segmentação baseada em IA apresentou valores de *RMS* significativamente menores nas mesmas regiões, exceto nas áreas em que IA e manual não diferiram estatisticamente, demonstrando a superioridade da IA em relação à limiarização global, ainda que seu desempenho varie conforme a região anatômica.

3.4 Forames e canais intraósseos

A revisão identificou 14 estudos que empregaram a inteligência artificial (IA) para a identificação e análise de forames e canais intraósseos em exames de tomografia computadorizada de feixe cônico (TCFC). Os trabalhos foram classificados nos seguintes subtemas: óstio acessório ($n = 1$), forame lingual ($n = 1$), canal mandibular ($n = 8$), forame infraorbitário ($n = 1$), canal nasolacrimonial ($n = 1$) e canal incisivo mandibular ($n = 2$).

3.4.1 Segmentação do canal mandibular

No estudo de Issa et al. (2025), 150 exames de tomografia computadorizada de feixe cônico (TCFC) retrospectivos, totalizando 300 canais mandibulares, foram analisados. A segmentação de referência foi realizada de modo semiautomático no software *Romexis*, enquanto a segmentação automática foi gerada pela plataforma *Diagnocat*. A comparação entre os modelos tridimensionais foi feita por meio do *Cloud Compare*, calculando-se as distâncias *surface-to-surface*. A discrepância mediana entre a segmentação da IA e a referência foi de 0,29 mm (desvio-padrão 0,25–0,37 mm), com 88% dos casos dentro do limite clínico aceitável ($\leq 0,5$ mm). A confiabilidade *inter-rater* foi de 84,5% e a *intra-rater* de 95,5%, indicando boa consistência entre avaliadores. A presença e o estado do terceiro molar influenciaram levemente a precisão: canais com terceiros molares impactados apresentaram discrepância mediana superior (0,32 mm) em relação aos erupcionados (0,28 mm) e ausentes (0,27 mm).

Gümüşsoy et al. (2025) utilizaram 69 exames de TCFC contendo variações bifurcadas do canal mandibular. As anotações manuais foram realizadas no *3D Slicer* por radiologistas experientes, e o modelo aplicado foi o *nnU-Net v2*. Para canais normais (*MaC*), o desempenho foi elevado (*Dice* $\approx 0,82$; *IoU* $\approx 0,70$; *AUC* $\approx 0,90$), enquanto para canais bifurcados (*BMaC*) observou-se queda significativa (*Dice* $\approx 0,46$; *IoU* $\approx 0,33$; *AUC* $\approx 0,71$), revelando maior dificuldade da IA em detectar ramificações anatômicas.

Em Kwak et al. (2020), 102 exames de TCFC foram anotados manualmente a cada 1 mm, com interpolação tridimensional. Foram comparadas arquiteturas bidimensionais (*U-Net*, *SegNet*) e tridimensionais (*3D U-Net*). O modelo tridimensional apresentou os melhores resultados, com acurácia global de

aproximadamente 0,99 e *class accuracy* de 0,96, superando substancialmente os modelos bidimensionais, especialmente na continuidade estrutural do canal.

De modo semelhante, Yang et al. (2025) testaram múltiplas arquiteturas (*2D-ResUNet*, *2D-AttUNet*, *2.5D*, *3D-UNet* e *3D-Swin UNETR*) em bases públicas (153 volumes de TCFC da Itália) e externas (30 volumes da Coreia). O *3D-UNet* apresentou o melhor desempenho global (*DSC* $\approx$ 0,72; *Jaccard* $\approx$ 0,57; precisão $\approx$ 0,66; sensibilidade $\approx$ 0,81), mantendo boa robustez em dados externos. O uso de estratégias como *Image Cropping* e *Multi-Planar Dice Loss* aprimorou a conectividade e reduziu falsos negativos. Os modelos *2.5D* apresentaram precisão elevada, porém inferior em continuidade estrutural quando comparados aos modelos tridimensionais.

No estudo de Pan et al. (2025), 836 exames de TCFC de um fabricante e 220 de outros quatro fabricantes foram utilizados para testar a capacidade de generalização do modelo. Uma rede do tipo *3D U-Net-like* com perda híbrida (*Dice* + *Cross-Entropy*) atingiu *ASSD* $\approx$ 0,49 mm e *SMCD* $\approx$ 0,30 mm no teste interno. Em dados externos, 86,8% dos exames receberam pontuação visual $\geq$ 4/5 por radiologistas, indicando boa acurácia clínica. O tempo médio para segmentação automática foi de 8,5 s, contrastando com aproximadamente 10 min na segmentação manual, representando um ganho expressivo de eficiência.

Usman et al. (2022) desenvolveram um *pipeline* em duas etapas. A primeira etapa, de localização grosseira, utilizou uma *3D Attention U-Net*, enquanto a segunda, de segmentação fina, aplicou uma *Residual U-Net* com entrada multiescala (*MSiR-UNet*). O método alcançou *Dice* $\approx$ 0,75 e *mIoU* $\approx$ 0,80, com melhora consistente ao incorporar conexões residuais e entradas multiescala. Os resultados foram considerados clinicamente aceitáveis e robustos em diferentes conjuntos de dados.

Em On et al. (2025), foram testadas abordagens *single-label*, *pair-label* e *multi-label* com o modelo *3D nnU-Net* em exames de TCFC. O modelo *multi-label* apresentou desempenho superior (*DSC* $\approx$ 90% no teste interno e $\approx$ 86% no externo para o nervo alveolar inferior). Além disso, o uso de *active learning* reduziu o tempo de correção manual de 285 minutos para 20 minutos usando o *multi-label*.

Por fim, Järnstedt et al. (2023) avaliaram 165 exames de TCFC de 72 pacientes, abrangendo diferentes situações clínicas. O modelo convolucional tridimensional (*DLS*) obteve *SMCD* mediana de 0,64 mm, *ASSD* mediana de 0,35 mm e *DSC* médio de 0,55, com coeficiente de repetibilidade (*RC*) de 0,97 mm. Apesar do desempenho moderado em sobreposição, o modelo apresentou boa repetibilidade e avaliação qualitativa próxima à do radiologista especialista (pontuação média $3,84 \pm 0,65$ vs. $3,94 \pm 0,27$).

3.4.2 Identificação do óstio acessório

No estudo de Shetty et al. (2025), foram analisados 856 exames de tomografia computadorizada de feixe cônico (TCFC) com campo de visão amplo (*Large Field of View – FOV*), com o objetivo de desenvolver um método de identificação automática de óstios acessórios. A partir desse conjunto, foram selecionadas 454 imagens de seções coronais, distribuídas equitativamente entre casos com óstio acessório ($n = 227$) e sem óstio acessório ($n = 227$).

Para a classificação automática, os autores empregaram modelos de redes neurais convolucionais pré-treinados, incluindo *VGG16*, *MobileNetV2* e *ResNet101V2*. Após testes comparativos, o modelo *ResNet101V2* foi selecionado por apresentar o melhor desempenho, sendo posteriormente ajustado por meio de *fine-tuning* e aplicação de regularização *L1* (*Lasso*), com o intuito de reduzir o risco de *overfitting*.

A análise descritiva indicou uma prevalência de óstio acessório de aproximadamente 24,18% entre os exames avaliados. No conjunto de teste, o modelo *ResNet101V2* alcançou acurácia média de cerca de 0,81, com precisão (*precision*) de 0,82, sensibilidade (*recall*) de 0,81 e *F1-score* de 0,81, demonstrando desempenho equilibrado entre detecção e classificação correta das estruturas.

Além disso, a avaliação da confiabilidade entre observadores humanos revelou boa concordância *inter* e *intra-avaliador*, indicando consistência na identificação manual dos óstios acessórios e reforçando a confiabilidade das anotações utilizadas como referência para o treinamento e validação do modelo.

3.4.3 Análise do forame lingual

No estudo de Mahabob et al. (2025), foi conduzida uma análise retrospectiva de 166 exames de tomografia computadorizada de feixe cônico (TCFC), com o objetivo de investigar variações anatômicas relacionadas ao forame mandibular e avaliar o potencial de métodos de inteligência artificial (IA) na estratificação de risco anatômico. Foram coletadas diversas medidas morfológicas, incluindo o número de foramina, a posição em relação ao tubérculo geniano, a distância até a crista alveolar e o diâmetro do forame, entre outros parâmetros.

Dois modelos de IA foram empregados de maneira complementar. Inicialmente, utilizou-se o algoritmo de *clustering K-means*, que permitiu agrupar os casos com base nas características anatômicas, formando três categorias de risco distintas: *Low*, *Moderate* e *High*. Em seguida, aplicou-se um modelo de árvore de decisão (*decision tree*), cujo objetivo foi determinar quais variáveis anatômicas apresentavam maior poder preditivo para a classificação de risco estabelecida pelo agrupamento anterior.

O modelo de árvore de decisão apresentou acurácia global de aproximadamente 92,6%, demonstrando alta capacidade de previsão das categorias de risco. Quando comparadas às classificações manuais realizadas por especialistas, as predições do modelo mostraram concordância de cerca de 89,4%, evidenciando bom alinhamento entre a análise automatizada e a avaliação humana.

3.4.4 Identificação do canal incisivo mandibular

No estudo de Jindanil et al. (2025), foi conduzida uma análise de 10 casos clínicos de pacientes do *University Hospitals Leuven* submetidos a cirurgias de implantes dentários, com documentação de exames de tomografia computadorizada de feixe cônico (TCFC) pré e pós-operatórios, além de registros clínicos de lesões nervosas decorrentes dos procedimentos. A segmentação foi realizada por meio de uma arquitetura de rede neural convolucional do tipo *U-Net*, previamente treinada para ambas as estruturas anatômicas.

Dois radiologistas orais compararam visualmente os canais segmentados pela IA com aqueles observados nas imagens originais sem processamento, enquanto cinco observadores (cirurgiões, protodontistas e residentes) avaliaram a identificação do

canal incisivo mandibular e a detecção de lesões nervosas em imagens pré e pós-operatórias.

A ferramenta de IA demonstrou desempenho significativamente superior no pré-operatório, com acurácia média de 95%, em comparação aos 70% obtidos nas análises sem IA ($p = 0,025$). Além disso, observou-se um aumento de 25% na detecção correta do canal incisivo mandibular e uma elevação de aproximadamente 8% no nível de confiança dos observadores ($p = 0,038$). No pós-operatório, a acurácia na detecção de lesões nervosas também foi maior com o uso da IA (92,5% versus 87,5%), embora sem significância estatística. Em termos qualitativos, a IA permitiu a visualização clara e bilateral dos canais incisivos, evidenciando relações anatômicas entre implantes e canais condizentes com os sintomas clínicos de dor e alterações neurossensoriais relatadas pelos pacientes.

De forma semelhante, Da Andrade-Bortoletto et al. (2025) avaliaram o desempenho de um modelo aprimorado de IA para segmentação automática do canal mandibular em 140 exames de TCFC, com 40 exames adicionais utilizados para validação externa. O modelo baseou-se em uma arquitetura tridimensional *U-Net* com múltiplas camadas convolucionais e uso de técnicas de *data augmentation* para aumentar a robustez do treinamento. Os resultados da segmentação foram revisados por radiologistas experientes, com refinamentos manuais aplicados em casos de subsegmentação e validação final conduzida por um radiologista sênior.

O modelo aprimorado (*enhanced AI*) apresentou desempenho quantitativo de destaque: $IoU \approx 93\%$, *Dice Similarity Coefficient (DSC)* $\approx 93\%$, *recall* $\approx 94\%$, precisão $\approx 93\%$, acurácia global $\approx 99\%$ e erro médio de superfície (*RMSE*) $\approx 0,23$ mm. Todas as métricas foram significativamente superiores às do modelo anterior e, em diversos aspectos, também às segmentações manuais realizadas por especialistas — especialmente em *IoU*, *DSC* e acurácia global. O tempo médio de processamento foi de 17,6 segundos por exame, contrastando com o tempo substancialmente maior das segmentações manuais, tornando o método aproximadamente 125 vezes mais rápido.

3.4.5 Segmentação do canal infraorbitário

No estudo de Gümüşsoy et al. (2025), foram analisadas 220 imagens de tomografia computadorizada de feixe cônico (TCFC), correspondentes aos canais infraorbitários de 110 pacientes (considerando ambos os lados). As imagens foram anotadas manualmente por radiologistas experientes utilizando o software *3D Slicer*, com verificação independente por um segundo e, quando necessário, um terceiro especialista, garantindo alta confiabilidade nas segmentações de referência.

Para a segmentação automática, os autores empregaram o modelo *nnU-Net v2*, uma rede neural profunda tridimensional amplamente reconhecida por sua capacidade de adaptação a diferentes tarefas de segmentação em imagens médicas. O desempenho do modelo foi avaliado por meio de múltiplas métricas quantitativas, demonstrando resultados promissores: *Dice Coefficient (DC)* = 0,7792, *Intersection over Union (IoU)* = 0,6402, *F1-score* $\approx$ 0,7837 e *Hausdorff Distance a 95% (HD95)* = 0,7661 (em milímetros ou *voxels*, conforme a parametrização do estudo). Além disso, a área sob a curva *ROC (AUC)* foi de 0,91, indicando excelente poder discriminativo do modelo na identificação e delimitação do canal infraorbitário.

3.4.6 Segmentação do canal nasolacrimal

No estudo de Haylaz et al. (2025), foi analisada uma amostra composta por 100 pacientes, cujos exames de tomografia computadorizada de feixe cônico (TCFC) foram selecionados aleatoriamente. As imagens foram convertidas para o formato *DICOM* e processadas no software *3D Slicer*, onde o canal nasolacrimal foi anotado manualmente por meio de um método poligonal, servindo como referência (*ground truth*) para o treinamento e validação do modelo de inteligência artificial.

A arquitetura de aprendizado profundo utilizada foi o *nnU-Net v2*, uma rede neural convolucional tridimensional adaptável, amplamente empregada em tarefas de segmentação automática de imagens médicas. O modelo apresentou desempenho quantitativo consistente e de alta qualidade, com os seguintes resultados médios: precisão (*precision*) $\approx$ 0,79, sensibilidade (*recall*) $\approx$ 0,92, *Dice Coefficient (DC)* $\approx$ 0,85, *Intersection over Union (IoU)* $\approx$ 0,73, *F1-score* $\approx$ 0,85 e

Hausdorff Distance a 95% (*HD95*) $\approx 0,95$ mm. A área sob a curva ROC (*AUC*) foi de aproximadamente 0,96, refletindo excelente capacidade discriminativa do modelo na diferenciação entre as estruturas segmentadas e os tecidos adjacentes.

3.5 Seios paranasais

Foram identificados, na revisão, 10 estudos que empregaram a inteligência artificial (IA) para a detecção e análise dos seios paranasais em exames de tomografia computadorizada de feixe cônico (TCFC). Esses trabalhos foram classificados nos seguintes subtemas: segmentação dos seios maxilares ($n = 7$) e determinação de sexo ($n = 3$).

3.5.1 Segmentação dos seios maxilares

No estudo de Bayrakdar et al. (2024), a base de dados foi composta por 101 exames de tomografia computadorizada de feixe cônico (TCFC) com seios maxilares saudáveis, sem achados inflamatórios. A rotulação (*ground truth*) foi realizada por dois especialistas em radiologia bucomaxilofacial e posteriormente revisada por radiologistas mais experientes, assegurando elevada precisão nas anotações. O modelo de inteligência artificial (IA) utilizado foi o *nnU-Net* v2, uma rede convolucional voltada à segmentação automática. O desempenho obtido foi elevado, com *F1-score* $\approx 0,96$, acurácia $\approx 0,99$, sensibilidade $\approx 0,96$, precisão $\approx 0,96$, *AUC* $\approx 0,97$, *Dice Coefficient* (*DC*) $\approx 0,96$, *Intersection over Union* (*IoU*) $\approx 0,93$ e *Hausdorff Distance* a 95% (*HD95*) $\approx 1,19$. O modelo demonstrou alta confiabilidade e precisão na segmentação automática dos seios maxilares, com interseção superior a 50% entre as previsões e o *ground truth* em praticamente todos os casos de teste.

Em Altun et al. (2024), foram analisados 307 exames de TCFC anônimos, com cortes coronais selecionados em intervalos de 1 mm. As imagens foram classificadas em quatro categorias: seios maxilares saudáveis, espessamento de mucosa, cistos de retenção de muco e sinusite (opacificação total ou parcial). A segmentação manual foi realizada por dois radiologistas bucomaxilofaciais, apresentando excelente concordância interobservador ($\kappa \approx 0,96$). O modelo adotado foi uma *YOLOv5x* modificada, com *transfer learning* a partir de pesos pré-treinados. O desempenho para a segmentação do seio maxilar total foi excelente (*recall* = 1,00;

precisão = 0,99; $F1\text{-score}$ = 0,99; $mAP \approx 0,99$; $AUC \approx 0,98$). Entre as classes patológicas, a IA apresentou maior acurácia para sinusite ($F1\text{-score}$ = 0,97) e cistos de retenção mucosa ($F1\text{-score}$ = 0,92), e desempenho ligeiramente inferior para espessamento de mucosa ($F1\text{-score}$ = 0,89), o que foi atribuído à dificuldade de delimitação de contornos sutis e áreas reduzidas.

O estudo de Yoo et al. (2023) avaliou 67 exames de TCFC, com segmentações manuais realizadas por radiologista bucomaxilofacial. Foram comparadas redes bidimensionais (*U-Net*, *U-Net++*), 2.5D e tridimensionais (*U-Net*, *V-Net*). O modelo 2.5D, que integra informações simultâneas dos planos axial, sagital e coronal por meio de um *ensemble* de previsões, apresentou os melhores resultados. Para o seio maxilar (MS), obteve $Jaccard \approx 0,95$, $Dice \approx 0,97$, precisão $\approx 0,97$ e $recall \approx 0,97$; para a região MSL, os valores foram $Jaccard \approx 0,79$ e $Dice \approx 0,88$. As redes tridimensionais apresentaram limitações de memória e menor capacidade de generalização, enquanto o método de votação majoritária nas redes 2.5D mostrou desempenho superior.

No estudo de Chen (2025), a amostra foi composta por 300 imagens de TCFC de 100 pacientes chineses, analisadas nos planos axial, coronal e sagital. Utilizando a *U-Net* clássica, o modelo apresentou excelente desempenho em todos os planos: axial – $IoU \approx 0,94$; $F1\text{-score} \approx 0,97$; acurácia $\approx 0,99$; sagital – $IoU \approx 0,94$; $F1\text{-score} \approx 0,97$; acurácia $\approx 0,99$; coronal – $IoU \approx 0,92$; $F1\text{-score} \approx 0,96$; acurácia $\approx 0,99$. Embora o plano coronal tenha apresentado desempenho ligeiramente inferior, o modelo demonstrou robustez tanto para seios regulares quanto para aqueles com septos internos.

Em Morgan et al. (2022), foram avaliados 132 exames de TCFC (264 seios maxilares) utilizando uma 3D *U-Net*. A segmentação automática apresentou $Dice \approx 98,4\%$ em comparação à segmentação semiautomática de referência, com tempo médio de processamento de apenas 24,4 segundos por seio, em contraste com 60,8 minutos no método semiautomático. O refinamento manual, quando necessário, levou cerca de 7,1 minutos, mantendo $DSC \approx 99,6\%$ e $RMS \approx 0,21$ mm. Clinicamente, cerca de 70% das segmentações automáticas foram classificadas como “perfeitas”, evidenciando alto grau de aplicabilidade prática.

O estudo de Choi et al. (2022) analisou 19.350 imagens de TCFC de 90 seios maxilares (34 claros e 56 turvos). A segmentação manual foi realizada em cortes axiais por meio do software *3D Slicer*, e a rede *U-Net* foi aplicada para segmentação automática. Antes do pós-processamento, o desempenho médio foi de $DSC = 0,91 \pm 0,19$ e $HD = 2,70 \pm 4,62$; após o pós-processamento, observou-se melhora para $DSC = 0,91 \pm 0,19$ e $HD = 2,15 \pm 2,28$. O tempo de segmentação automática foi significativamente menor (≈ 46 s) em comparação ao método manual ($\approx 48,7$ min), sem diferença relevante de desempenho entre seios claros e turvos.

Por fim, Serindere et al. (2022) avaliaram 296 imagens (148 de seios maxilares saudáveis e 148 com sinusite), obtidas por meio de TCFC e radiografia panorâmica (*panoramic radiograph – PR*). Utilizando uma *convolutional neural network (CNN)* implementada em *PyTorch*, observaram desempenho excepcional com TCFC (acurácia $\approx 99,7\%$; sensibilidade $\approx 100\%$; especificidade $\approx 99,3\%$; $AUC \approx 0,99–1,00$), significativamente superior ao obtido com *PR*.

3.5.2 Determinação de sexo

No estudo conduzido por Şenol et al. (2024), foram analisados 160 exames de tomografia computadorizada de feixe cônico (TCFC) de indivíduos adultos. Os autores excluíram da amostra participantes que apresentavam patologias ou histórico de intervenções cirúrgicas no ramo da mandíbula, assegurando a integridade anatômica da região de interesse. Foram realizadas medições bilaterais (direita e esquerda) de oito parâmetros antropométricos relacionados à língua mandibular, incluindo distâncias entre o ponto superior da língua e referências anatômicas como a incisura mandibular, as bordas anterior e posterior da mandíbula, a base mandibular e o gônio, além da altura da própria língua.

Para a estimativa do sexo, diferentes modelos de *machine learning* foram avaliados, com destaque para os algoritmos *Random Forest* e *Gaussian Naive Bayes*, que apresentaram as maiores acurácias observadas, atingindo 0,88 (88%). Outras combinações e parâmetros testados obtiveram acurácias variando entre 0,78 e 0,88, demonstrando desempenho consistente. Os autores também reportaram valores equilibrados de sensibilidade, especificidade e *F1-score*, indicando boa capacidade discriminatória dos modelos empregados.

Por sua vez, Da Silva et al. (2024) utilizaram uma base composta por 310 exames de TCFC, propondo uma arquitetura de rede neural profunda denominada *SDetNet*, estruturada em duas etapas principais: *FSNet*, uma rede bidimensional de segmentação dos seios frontais, responsável por delinear precisamente o contorno anatômico dessa estrutura; e *SDNet*, uma rede tridimensional de classificação, que utiliza as regiões de interesse (*ROIs*) segmentadas para prever o sexo. O modelo faz uso de entradas multicanais (*MSI/MCI*), combinando imagens originais e mapas de máscara, e incorpora o *Anatomy-Guided Attention Module (AGAM)*, um módulo de atenção projetado para direcionar o foco da rede a contextos anatômicos relevantes dos seios frontais.

O desempenho da etapa de segmentação foi expressivo, com a *backbone DenseNet201* alcançando *Jaccard Index (JI)* = $0,878 \pm 0,042$, *F1-score* = $0,935 \pm 0,024$ e *recall* = $0,930 \pm 0,038$. Na etapa de determinação do sexo, a arquitetura completa (*DenseNet201 + AGAM + MSI*) apresentou *AUC* $\approx 0,98$, acurácia $\approx 0,92$ e especificidade $\approx 0,96$, além de *Brier Score* e sensibilidade igualmente favoráveis. Estudos de ablação demonstraram que a remoção dos mapas de máscara ou do *AGAM* reduziu significativamente o desempenho, reforçando a importância desses componentes na extração de informações morfológicas discriminantes entre os sexos.

Já no estudo de Hamidi et al. (2024), foram analisadas 240 imagens de TCFC de pacientes iranianos, com foco na mensuração de parâmetros lineares e volumétricos dos seios frontal e maxilar (altura, largura, comprimento e volume). Os autores desenvolveram um método denominado *GADNN (Genetic Algorithm-based Deep Neural Network)*, que combina redes neurais profundas (*DNN*) com um algoritmo genético (*GA*) para seleção de características e otimização do modelo, além de empregar a técnica *SMOTE* para balanceamento das classes.

A variável idade foi abordada de duas formas: como variável contínua, para regressão, e em faixas etárias, para classificação multiclasse. O modelo *GADNN* apresentou acurácia de 86% na determinação do sexo, superando modelos comparativos como *Logistic Regression*, *Random Forest* e *Multilayer Perceptron (MLP)*, e atingiu 68% de acurácia na classificação etária.

3.6 Articulação temporomandibular

A revisão identificou sete estudos que empregaram a inteligência artificial (IA) na análise das condições da articulação temporomandibular (ATM) e de seus componentes em exames de tomografia computadorizada de feixe cônico (TCFC). Esses trabalhos foram agrupados em três subtemas principais: osteoartrite ($n = 3$), segmentação do côndilo mandibular ($n = 3$) e articulação temporomandibular propriamente dita ($n = 1$).

3.6.1 Identificação de osteoartrite na ATM

No estudo de De Dumast et al. (2018), foram empregados modelos de superfície tridimensionais (*meshes*) de côndilos mandibulares obtidos a partir de exames de tomografia computadorizada de feixe cônico (TCFC). O conjunto de treinamento foi composto por 259 côndilos, sendo 105 provenientes de indivíduos saudáveis e 154 de pacientes com diagnóstico de osteoartrite da articulação temporomandibular (ATM). Para o grupo de teste, foram aplicados questionários clínicos padronizados e coletadas amostras de sangue e saliva. Dois sistemas integrados foram desenvolvidos: o *Shape Variation Analyzer* (SVA), módulo voltado à análise da variabilidade tridimensional da forma condilar e à classificação automática do grau de degeneração, e o *Data Storage for Computation and Integration* (DSCI), plataforma *web* destinada ao armazenamento e integração de dados clínicos, biológicos e de imagem, além da execução remota de tarefas computacionais complexas.

As características morfológicas tridimensionais foram descritas a partir de vetores, curvaturas, distâncias e posições de vértices, além do índice de forma. A rede neural empregada possuía camadas ocultas com função *softmax* na saída, permitindo a classificação dos côndilos em cinco estágios distintos de degeneração estrutural. O sistema apresentou 91% de concordância entre a classificação automática do SVA e o consenso clínico obtido por avaliadores experientes, indicando alta confiabilidade do modelo. Além disso, as análises integradas revelaram correlações significativas entre a morfologia condilar, níveis de dor e marcadores biológicos, demonstrando o potencial da integração multimodal de dados.

No estudo conduzido por Talaat et al. (2023), uma base composta por 2.737 imagens de TCFC provenientes de 943 pacientes foi utilizada para o treinamento e validação de um modelo de rede neural convolucional com detecção de objetos (*object detection*), empregando regressão para localizar sinais patológicos. Como referência diagnóstica (*gold standard*), dois avaliadores experientes aplicaram os critérios *DC/TMD* em 350 imagens de um conjunto de teste independente. O consenso entre os avaliadores foi utilizado para comparação com os resultados do modelo de IA e com o diagnóstico fornecido por um radiologista humano.

Os resultados demonstraram que a IA apresentou maior concordância com o padrão de referência em relação ao radiologista, com coeficiente *Kappa* de Cohen de 0,815 (versus 0,709). Essa diferença foi estatisticamente significativa para os sinais patológicos em conjunto ($p = 0,0079$) e, em especial, para cistos subcorticais ($p = 0,0214$). Para outros achados, como achatamento, erosão e osteófitos, não houve diferença estatística relevante. A IA também apresentou elevados valores de sensibilidade e especificidade, variando entre 86% e 95% e 88% e 93%, respectivamente, dependendo do tipo de alteração avaliada.

Por fim, o estudo de Eşer et al. (2023) utilizou 2.000 cortes sagitais de TCFC de 290 pacientes, empregando a arquitetura *YOLOv5*, reconhecida por sua alta velocidade e precisão na detecção de imagens. O modelo foi configurado para realizar simultaneamente a segmentação da região da ATM e a classificação da condição óssea condilar em quatro categorias: normal, erosão, osteófitos e achatamento.

O modelo alcançou excelente desempenho na segmentação, com sensibilidade de 100%, precisão de 99,53%, *F1-score* de 99,76% e *AUC* de 0,9723, evidenciando praticamente ausência de falsos positivos e negativos. Na classificação das condições ósseas, o desempenho geral foi satisfatório, com acurácia média de 86,86%, apresentando melhores resultados para os grupos “normal” (88%) e “erosão” (95%), e desempenho ligeiramente inferior para “achatamento” (70%). A sensibilidade geral de 100% confirmou a capacidade do modelo em detectar corretamente todos os casos de osteoartrite, sem ocorrência de falsos negativos.

3.6.2 Segmentação de côndilos mandibulares

No estudo de Orhan et al. (2023), foram analisados 432 exames retrospectivos de tomografia computadorizada de feixe cônico (TCFC) provenientes de quatro universidades, totalizando 864 côndilos mandibulares. Seis radiologistas especialistas em imagem dentomaxilofacial avaliaram, de forma independente e às cegas, os exames originais no formato *DICOM*, classificando cinco tipos de alterações ósseas: achatamento, erosão, formação de osteófitos, osteosclerose e côndilo bífido. As divergências diagnósticas foram resolvidas em reunião de consenso, estabelecendo o diagnóstico de referência (*ground truth* – *GT*). Paralelamente, o software de inteligência artificial (*Diagnocat*) foi utilizado para gerar modelos tridimensionais (arquivos *STL*) a partir dos exames de TCFC, os quais foram avaliados por dois outros radiologistas com o objetivo de verificar se as mesmas alterações ósseas poderiam ser identificadas a partir desses modelos tridimensionais.

Do total de 864 côndilos avaliados, foram observados 372 casos de achatamento, 185 de erosão, 70 de osteófitos, 117 de osteosclerose e 15 de côndilo bífido. A confiabilidade interobservador entre os radiologistas na leitura dos TCFCs originais foi excelente, com coeficientes de correlação intraclassa (*ICC*) variando de 0,927 a 0,971 para todas as alterações ósseas. Quando comparado o diagnóstico obtido nos modelos *STL* com o *ground truth*, os resultados indicaram concordância perfeita (*ICC* = 1,000) para achatamento, formação de osteófitos e côndilo bífido, e boa concordância (*ICC* = 0,782) para erosão. Apenas o diagnóstico de osteosclerose apresentou baixa confiabilidade (*ICC* = 0,000), sugerindo limitação do modelo tridimensional gerado pela IA para detectar esse tipo de alteração.

No estudo de Kim et al. (2021), foram utilizadas 12.800 imagens de TCFC obtidas de 25 indivíduos sem alterações patológicas, com o objetivo de desenvolver um modelo de segmentação automática da cortical e da medular óssea do côndilo mandibular. As imagens foram manualmente anotadas por um radiologista bucomaxilofacial experiente, servindo como padrão de referência (*gold standard*). O sistema proposto combinou uma *U-Net* modificada, responsável pela separação entre os ossos cortical e medular, e uma *convolutional neural network* (*CNN*) de classificação, empregada para identificar se cada corte continha ou não a cabeça do

côndilo mandibular. Os resultados demonstraram alto desempenho na segmentação, com *IoU* médio de 0,87 e *Hausdorff Distance (HD)* de 0,93 mm para o osso medular, e *IoU* de 0,73 e *HD* de 1,25 mm para o osso cortical. As distribuições de espessura determinadas automaticamente apresentaram forte correlação com as medidas manuais, principalmente nas regiões de cortical mais delgada, e os mapas de cores tridimensionais gerados pelo modelo reproduziram de forma semelhante o padrão visual observado nas anotações de referência. O tempo médio de processamento foi de 10 a 15 segundos por exame completo em *GPU* de alto desempenho, demonstrando significativa eficiência computacional.

Já o estudo de Jha et al. (2022) avaliou o desempenho de duas arquiteturas distintas de redes neurais tridimensionais na segmentação automática do côndilo mandibular. Foram utilizadas 234 imagens de TCFC provenientes de 117 indivíduos, coletadas em duas instituições diferentes. As segmentações manuais serviram como *ground truth* para o treinamento e validação dos modelos. Os autores compararam uma *Basic 3D U-Net*, responsável pela segmentação direta do côndilo na imagem completa, e uma *Cascaded 3D U-Net*, composta por duas etapas: uma primeira *U-Net* para detecção grosseira da região de interesse (*ROI*) e uma segunda, aplicada sobre o recorte da *ROI*, para refinar a segmentação.

Em testes progressivos com diferentes tamanhos de conjunto de dados, observou-se melhora consistente do desempenho com o aumento do número de amostras de treinamento. No estágio com maior volume de dados, o coeficiente *Dice (DSC)* foi de aproximadamente $0,922 \pm 0,021$ para a *Basic 3D U-Net* e $0,932 \pm 0,023$ para a *Cascaded 3D U-Net*, indicando segmentação mais precisa no modelo em cascata. A *Hausdorff Distance* também foi menor na arquitetura *cascaded* (~2,45–2,60 mm) em comparação à básica (~2,56–3,10 mm). O tempo médio de segmentação manual foi de $14,75 \pm 3,63$ minutos por exame, enquanto o processamento automático reduziu esse tempo para $4,13 \pm 1,94$ minutos com a rede básica e $2,31 \pm 1,54$ minutos com a *cascaded*, demonstrando expressiva economia de tempo e alta eficiência computacional.

3.6.3 Segmentação da ATM

No estudo de Vinayahalingam et al. (2023), foi desenvolvida uma abordagem automatizada de três etapas baseada em rede neural tridimensional *U-Net* para análise da articulação temporomandibular (ATM) em exames de tomografia computadorizada de feixe cônico (TCFC). O método compreendeu: (1) determinação da região de interesse (*ROI*), com o objetivo de localizar os côndilos mandibulares e as fossas glenoides nas imagens; (2) segmentação óssea propriamente dita das estruturas localizadas; e (3) classificação da ATM (*TMJ classification*), distinguindo as estruturas segmentadas como pertencentes ou não à articulação.

O conjunto de dados foi composto por 154 exames de TCFC, previamente segmentados manualmente por especialistas, utilizados para o treinamento e validação dos modelos. O desempenho do algoritmo foi avaliado pelo índice de interseção sobre união (*Intersection over Union – IoU*), que alcançou 0,955 para os côndilos mandibulares e 0,935 para as fossas glenoides, indicando alta precisão na segmentação das estruturas.

Quando comparado ao desempenho de dois observadores humanos realizando a segmentação manual dos côndilos, os valores de *IoU* foram inferiores (0,895 e 0,928, respectivamente), demonstrando que o modelo de IA obteve desempenho igual ou superior ao de especialistas humanos.

O tempo médio de segmentação também evidenciou expressiva vantagem da abordagem automatizada: o algoritmo executou a tarefa em aproximadamente $3,6 \pm 0,9$ segundos por exame, enquanto os observadores humanos demandaram, em média, 379 ± 205 segundos e 572 ± 257 segundos, respectivamente. Esses resultados reforçam a eficiência e a robustez do modelo, capaz de produzir segmentações de alta qualidade em tempo significativamente reduzido.

3.7 Landmarks

No estudo de Tang et al. (2025), foi proposto um método semissupervisionado baseado na arquitetura *Mean Teacher*, utilizando 192 tomografias computadorizadas de feixe cônico (TCFC) tridimensionais de pacientes candidatos à cirurgia ortognática. Essa estrutura envolve dois modelos — *teacher* (professor) e *student* (aluno) —, em que o aluno aprende tanto com dados rotulados quanto com dados

não rotulados, supervisionado pelo modelo professor, que gera *pseudo-labels*. A consistência entre ambos, mesmo sob perturbações e ruídos, garante melhor generalização e estabilidade de aprendizado.

Para a segmentação da maxila e da mandíbula, foi utilizada uma rede tridimensional *V-Net*, enquanto a detecção dos *landmarks* anatômicos foi realizada por meio de regressão de *heatmaps* 3D, que indicam as probabilidades espaciais de cada ponto anatômico. O sistema incorporou ainda um módulo de estimação de incerteza, permitindo ao modelo professor filtrar apenas os *pseudo-labels* considerados confiáveis. O desempenho obtido foi expressivo, com coeficiente *Dice* de 93,41% para a maxila e 96,89% para a mandíbula. Na detecção de 18 *landmarks* anatômicos, o erro médio foi de $1,908 \pm 1,166$ mm, desempenho superior ao de um modelo puramente supervisionado, evidenciando a eficácia da abordagem semissupervisionada para análise crânio-facial.

No estudo de Deng et al. (2023), foram avaliados 61 conjuntos de TCFC de pacientes também indicados para cirurgia ortognática. O sistema desenvolvido, denominado *SkullEngine*, baseia-se em uma rede neural convolucional multiestágio voltada à segmentação e detecção automática de marcos anatômicos tridimensionais. O modelo executa a segmentação automática da face média, mandíbula e arcadas dentárias, além de identificar automaticamente marcos anatômicos de referência utilizados no planejamento cirúrgico. Em comparação com o padrão-ouro (segmentações e marcações manuais realizadas por especialistas), o sistema alcançou coeficiente *Dice* médio de 96%, com distância média simétrica de 0,1 mm, demonstrando elevada precisão geométrica.

A segmentação automática foi considerada clinicamente utilizável sem necessidade de ajustes manuais em 98,4% da face média, 70,5% da mandíbula, 98,4% dos dentes superiores e 93,4% dos dentes inferiores. A diferença média entre as posições dos marcos automáticos e manuais foi de 2,3 mm para a face média e 2,4 mm para a mandíbula, valores dentro dos limites clínicos aceitáveis. O tempo médio de segmentação automática foi de 4 minutos por exame, sendo que eventuais ajustes manuais adicionaram cerca de 10 minutos, representando expressiva redução do tempo total de processamento em comparação com o método manual.

O estudo de Park et al. (2024) investigou a detecção automática de *landmarks* cefalométricos tridimensionais em 80 exames de TCFC, divididos em três grupos: pacientes sem cirurgia ortognática ($n = 39$), pós-cirurgia sem *hardware* ($n = 9$) e pós-cirurgia com *hardware* ($n = 32$). Foram identificados 65 *landmarks* e realizadas 53 medições cefalométricas (27 lineares, 21 angulares e 5 proporcionais), comparando resultados manuais e automáticos. O algoritmo, baseado em uma abordagem *coarse-to-fine*, apresentou erro médio de $1,7 \pm 0,1$ mm entre os *landmarks* automáticos e manuais, com taxas de detecção de 88,16% (erro ≤ 3 mm) e 94,35% (erro ≤ 4 mm), ambas dentro da faixa de tolerância clínica.

Das 53 medidas avaliadas, apenas seis apresentaram diferenças estatisticamente significativas entre os métodos. O grupo pós-cirurgia sem *hardware* apresentou os menores erros, indicando melhor concordância geral. Mesmo em casos com placas e parafusos metálicos, o algoritmo manteve desempenho robusto. O tempo de marcação foi substancialmente reduzido: a identificação manual de *landmarks* exigia entre 40 e 60 minutos por exame, enquanto o método automático executou a tarefa em cerca de 10,9 segundos por volume, evidenciando grande eficiência e aplicabilidade clínica.

No estudo de Tanikawa et al. (2025), foram analisados 185 exames de TCFC de adultos japoneses, com o objetivo de desenvolver um modelo de identificação automática de *landmarks* tridimensionais e reconstrução padronizada da morfologia craniofacial. As superfícies do crânio e da mandíbula foram reconstruídas a partir das imagens TCFC e validadas com 64 *landmarks* manuais (19 cranianos e 45 mandibulares) definidos por especialistas. O modelo proposto utilizou a arquitetura *PointNet++*, que processa nuvens de pontos e vetores normais das superfícies, aprendendo a prever as coordenadas dos *landmarks*.

Após a marcação automática, aplicou-se uma etapa de *homologous mesh fitting*, ajustando um modelo *template* às posições preditas para gerar malhas tridimensionais padronizadas, úteis para comparações interindividuais. Os erros médios de localização dos *landmarks* foram de 3,07 mm para a maxila e 2,15 mm para a mandíbula, com maiores discrepâncias na direção vertical. Após o ajuste de malha, os erros foram substancialmente reduzidos ($0,80 \pm 0,57$ mm para a maxila e $1,45 \pm 0,34$ mm para a mandíbula), demonstrando o potencial da técnica em refinar

a acurácia geométrica. O tempo médio de identificação automática foi de aproximadamente 2 segundos, confirmando alto desempenho e aplicabilidade prática da abordagem.

No estudo de Chung et al. (2022), foram analisados 30 indivíduos para comparar medições obtidas em três tipos de imagens: cefalogramas convencionais, cefalogramas reconstruídos de TCFC (software *OnDemand 3D*) e cefalogramas reconstruídos de TCFC (software *Invivo5*). O programa *WebCeph* foi utilizado para marcação automática de 17 *landmarks* e cálculo de 11 medidas cefalométricas (distâncias e ângulos). Não foram observadas diferenças estatisticamente significativas entre os três métodos, e as pequenas variações encontradas — como na medida *upper lip–E line* — permaneceram dentro de limites clínicos aceitáveis, confirmando a equivalência diagnóstica entre radiografias convencionais e reconstruções TCFC processadas por IA.

Em outro estudo de Chung et al. (2024), com 16 participantes, foram comparadas radiografias cefalométricas convencionais, imagens reconstruídas de TCFC padrão e TCFC de baixa dose processadas por IA. As médias dos parâmetros angulares (SNA, SNB, ANB, entre outros) foram muito próximas entre os grupos, sem diferenças estatísticas ($p > 0,05$), exceto para a medida linear *nasion–menton* ($p < 0,05$). Os resultados mostraram que as imagens de baixa dose processadas por IA mantêm acurácia diagnóstica e qualidade suficiente para análise cefalométrica, com a vantagem de reduzir significativamente a dose de radiação.

Por fim, o estudo de Chen et al. (2020) utilizou o algoritmo *LINKS* (*Learning-based Multi-source Integration Framework for Segmentation*) para realizar a segmentação automática do osso maxilar e a detecção de *landmarks* anatômicos (Basión, Násion e Espinha Nasal Anterior). A amostra incluiu 60 pacientes (30 com canino impactado e 30 controles). O modelo alcançou coeficiente *Dice* médio de $0,800 \pm 0,029$ na segmentação tridimensional e diferença média de cerca de dois *voxels* entre os *landmarks* automáticos e manuais. A confiabilidade foi altíssima, com $ICC = 0,994$ para segmentação e $ICC = 0,999$ para detecção de *landmarks*. O tempo de processamento foi de aproximadamente 15 minutos por TCFC em uma CPU convencional.

3.8 Lesões orais

No estudo de MahdaviFar et al. (2025), foram analisados 1.134 laudos de tomografias computadorizadas de feixe cônico (TCFC) provenientes do banco de dados da *Shiraz University of Medical Sciences*. Três especialistas — um protesista, um radiologista bucomaxilofacial e um médico bucomaxilofacial — realizaram a rotulagem manual dos exames, classificando-os segundo diferentes níveis de severidade clínica, que variavam desde condições críticas com potencial risco à vida, como tumores malignos e cistos agressivos, até achados de menor relevância clínica ou ausência de alterações significativas.

Com o objetivo de ampliar a aplicabilidade do modelo, os autores também propuseram uma versão binária da classificação, agrupando os casos em duas categorias principais: alto risco, que incluía as condições mais graves ou que demandavam atenção clínica imediata, e baixo risco, que abrangia os achados de menor urgência.

Para a análise automatizada, foi desenvolvida uma arquitetura híbrida de aprendizado profundo, combinando redes neurais convolucionais (*Convolutional Neural Networks – CNN*) e redes recorrentes do tipo *Long Short-Term Memory (LSTM)*. Essa abordagem permitiu integrar a capacidade da *CNN* em extrair características espaciais complexas das imagens com a habilidade da *LSTM* em modelar dependências contextuais e temporais entre os dados. O desempenho do modelo foi comparado com diferentes algoritmos convencionais de aprendizado de máquina, como *Support Vector Machine (SVM)*, *Decision Tree (DT)*, *Random Forest* e *Logistic Regression*.

Os resultados demonstraram superioridade consistente do modelo *CNN–LSTM* em todas as métricas avaliadas, incluindo acurácia, precisão, sensibilidade (*recall*) e *F1-score*. Na classificação em quatro níveis de severidade, o modelo alcançou acurácia média de aproximadamente 97,3%, enquanto na tarefa de classificação binária obteve valores iguais ou superiores a 95% para todas as métricas. Além disso, a análise da área sob a curva *ROC (ROC–AUC)* confirmou a alta capacidade discriminativa do sistema, evidenciando excelente sensibilidade e especificidade na diferenciação entre casos de alto e baixo risco.

3.9 Marcadores radiográficos

No estudo de Alsomali et al. (2022), foram utilizados 34 conjuntos de dados de tomografia computadorizada de feixe cônico (TCFC) de pacientes submetidos a exames com o uso de estênceis radiográficos contendo marcadores de guta-percha (GP), empregados no planejamento de implantes dentários. O conjunto total de imagens axiais provenientes desses casos compreendeu 16.272 fatias, utilizadas para o treinamento e validação do modelo.

O sistema proposto baseou-se na arquitetura *Mask R-CNN*, uma rede neural convolucional profunda amplamente reconhecida por sua capacidade de realizar detecção e segmentação precisa de objetos em imagens médicas e não médicas. O desempenho do modelo foi avaliado em um conjunto de teste independente, composto por quatro casos clínicos, totalizando 193 imagens contendo marcadores GP e 2.284 imagens sem marcadores.

Os resultados demonstraram uma taxa de verdadeiros positivos de 83%, indicando que o modelo foi capaz de identificar corretamente a maioria dos marcadores presentes nas imagens. A taxa de falsos positivos foi relativamente baixa, correspondendo a 2,8% quando consideradas todas as fatias sem marcador como potenciais fontes de erro. Por outro lado, observou-se que 17% dos marcadores reais não foram detectados (falsos negativos), representando os casos em que o algoritmo não reconheceu corretamente a presença do marcador.

3.10 Segmentação das vias aéreas faríngeas

No estudo de Sin et al. (2021), foram incluídas retrospectivamente 306 imagens de tomografia computadorizada de feixe cônico (TCFC) contendo a via aérea faríngea, com o objetivo de avaliar a segmentação automática utilizando um algoritmo baseado em *Convolutional Neural Network (CNN)*. A segmentação manual, realizada com o software semiautomático *ITK-SNAP*, serviu como referência (padrão-ouro). A arquitetura aplicada foi baseada em *U-Net*, um modelo do tipo *encoder-decoder* amplamente utilizado em segmentação médica, no qual o caminho *encoder* realiza o *downsampling* para capturar o contexto global, enquanto o *decoder* realiza o *upsampling* para restaurar os detalhes espaciais. Os resultados

demonstraram excelente desempenho do modelo: o volume médio da via aérea estimado manualmente foi de 18,08 cm³, enquanto a IA estimou 17,32 cm³. O coeficiente *Dice Similarity Coefficient (DSC)* atingiu 0,919 e o *Intersection over Union (IoU)* ponderado foi de 0,993, evidenciando alta precisão da segmentação automática. A reprodutibilidade entre as medições humanas e automáticas também foi elevada, com coeficiente de correlação intraclassa (*ICC*) variando entre 0,985 e 0,986.

No estudo de Orhan et al. (2022), foram analisadas imagens de 200 pacientes — 100 com apneia obstrutiva do sono (*Obstructive Sleep Apnea, OSA*) e 100 controles — obtidas em três diferentes aparelhos de TCFC, com resolução isotrópica de *voxels* entre 0,1 e 0,2 mm³. O algoritmo automático (*Diagnocat* associado a uma rede neural convolucional) utilizou uma abordagem *coarse-to-fine*, composta por duas fases: a fase *coarse*, que processa a imagem inteira com *voxel* de aproximadamente 1 mm para estimativa geral, e a fase *fine*, que processa *patches* com *voxel* de cerca de 0,25 mm, utilizando como referência os resultados da fase inicial.

As comparações entre medições manuais, semiautomáticas e automáticas incluíram parâmetros como volume total da via aérea, ponto mais estreito e área da seção transversal mínima. Não foram observadas diferenças estatisticamente significativas entre os métodos em nenhum dos subgrupos de OSA (mínima, leve, moderada e grave) ou nos controles. A concordância entre medições foi elevada: *ICC* entre segmentação manual e automática $\approx 0,954$; *ICC* entre *Diagnocat* e automática $\approx 0,956$; e *ICC* entre *Diagnocat* e manual $\approx 0,972$. Nos pacientes sem OSA, o volume médio da via aérea foi de aproximadamente 17,95 cm³ (manual) e 18,50 cm³ (*Diagnocat*), enquanto, nos pacientes com OSA, a diferença média entre as medições automática e manual foi de cerca de 1,36 cm³.

3.11 Implantes dentários

No estudo de Elgarba et al. (2023), foi utilizado um conjunto de 280 exames de tomografia computadorizada de feixe cônico (TCFC) de maxila e mandíbula de pacientes com implantes dentários, incluindo casos com e sem restauração coronal. A segmentação de referência (*ground truth*) foi realizada por especialistas utilizando

método semiautomático. A segmentação automática, baseada em *Convolutional Neural Network* (CNN), apresentou tempo médio inferior a 30 segundos — aproximadamente 60 vezes mais rápida que o método de referência. O *Dice Similarity Coefficient* (DSC) foi de aproximadamente $0,92 \pm 0,02$ para implantes sem coroa e $0,91 \pm 0,03$ para implantes com coroa, enquanto o desvio médio das superfícies (*Root Mean Square – RMS*) manteve-se baixo, sendo $0,08 \pm 0,09$ mm para implantes simples e $0,11 \pm 0,07$ mm para implantes com restauração.

No estudo de Sakai et al. (2023), foram analisados exames TCFC anonimizados de 60 pacientes, a partir dos quais foram extraídas 1.200 imagens da região óssea destinada à colocação do implante. Cada imagem foi classificada segundo três protocolos de perfuração distintos utilizados na cirurgia: convencional com broca *tapping*, convencional sem broca *tapping* e *undersized*. A classificação automática foi realizada por meio de uma rede *LeNet-5*, que obteve acurácia global de aproximadamente 93,8%. As sensibilidades variaram entre 85,0% e 97,5%, as precisões entre 86,7% e 100%, e os valores de *F1-score* entre 91,8% e 93,8%. As áreas sob a curva *Receiver Operating Characteristic* (ROC-AUC) foram elevadas para todos os protocolos, variando de 98,6% a 99,4%, indicando excelente capacidade discriminativa do modelo.

Em Kurt Bayrakdar et al. (2021), foram incluídas 75 imagens de TCFC, totalizando 508 regiões com indicação para implante. As medições manuais de altura e espessura óssea foram comparadas às avaliações automatizadas realizadas pelo sistema *Diagnocat*. Não foram observadas diferenças estatisticamente significativas na altura óssea em regiões específicas da maxila e mandíbula ($p > 0,05$), enquanto a espessura óssea apresentou diferenças significativas em todas as regiões avaliadas ($p < 0,001$). A detecção de estruturas anatômicas apresentou taxas de acerto variando de 66,4% (seios/fossas) a 95,3% (regiões com dentes ausentes), com 72,2% de acerto na identificação dos canais mandibulares.

O estudo de Roongruangsilp et al. (2025) analisou 332 imagens de posições de implante derivadas de 184 exames de TCFC, processadas por quatro softwares distintos. A anotação de referência foi realizada por um implantologista experiente, que definiu *bounding boxes* para cada local de implante. A comparação entre os modelos *Faster R-CNN* e *YOLOv7* demonstrou que, em imagens seccionais obtidas

com o *DentiPlan Pro 3.7*, o *YOLOv7* apresentou taxa de detecção mais alta ($\approx 93,75\%$) em relação ao *Faster R-CNN* ($\approx 59,38\%$), porém com menor precisão (maior número de falsos positivos), enquanto a acurácia geral foi superior para o *Faster R-CNN* ($\approx 84,21\%$ vs. $46,67\%$). O desempenho dos modelos variou conforme o software de visualização, refletindo diferenças na renderização, contraste e algoritmos proprietários, embora não tenham sido observadas diferenças estatisticamente significativas entre os modelos ($p > 0,05$) nas análises globais.

3.12 Múltiplas segmentações

No estudo de Nogueira-Reis et al. (2024), foram avaliados 30 exames de tomografia computadorizada de feixe cônico (TCFC) contendo múltiplas estruturas anatômicas. A segmentação automatizada integrada, realizada por meio de um *pipeline* composto por redes neurais convolucionais (CNNs) previamente validadas, apresentou tempo médio de 1,1 minuto, enquanto a abordagem semiautomatizada, baseada em *threshold* e refinamento manual, demandou aproximadamente 48,4 minutos. A similaridade em relação ao referencial refinado foi de 99,6% para a abordagem automática e 88,3% para a semiautomatizada, evidenciando alta precisão e consistência da segmentação automática. Pequenas subsegmentações e sobre-segmentações foram observadas em regiões complexas, como espessamento mucoso, raízes dentárias e canais mandibulares. Em outro estudo conduzido por Nogueira-Reis et al. (2023), envolvendo 40 exames de TCFC, a avaliação qualitativa realizada por radiologistas indicou que 85% dos exames receberam notas entre 7 e 10 quanto à qualidade da segmentação. O tempo médio de processamento automático foi de 1,7 minuto e o coeficiente *Dice Similarity Coefficient* (DSC) foi de aproximadamente 99,3% em relação às segmentações refinadas manualmente, demonstrando excelente concordância entre os avaliadores.

Liu et al. (2024) empregaram 451 exames de TCFC provenientes de diferentes instituições e fabricantes, utilizando redes especializadas para segmentação óssea, segmentação individualizada de dentes e identificação do canal mandibular. O desempenho foi elevado, com DSC médio de aproximadamente 96,5% para dentes, 95,4% para osso alveolar, 93,6% para seio maxilar e 94,8% para o canal mandibular, além de *mean Intersection over Union* (mIoU) de 88,4% para dentes. O pré-processamento adaptativo melhorou significativamente os resultados em

comparação ao método genérico, e o modelo superou abordagens concorrentes, mantendo baixos valores de erro de superfície e de distância de *Hausdorff*.

No estudo de Wang et al. (2021), realizado com 30 exames de TCFC de pacientes ortodônticos, a rede *Multi-Scale Dense Network (MS-D)* alcançou *DSC* médio de $0,934 \pm 0,019$ para mandíbula e $0,945 \pm 0,021$ para dentes, com desvio médio absoluto de superfície (*Mean Absolute Deviation – MAD*) de $0,390 \pm 0,093$ mm e $0,204 \pm 0,061$ mm, respectivamente. O tempo médio de segmentação automática foi de 25 segundos, em contraste com as 5 horas necessárias para a segmentação manual, demonstrando expressivo ganho em eficiência sem perda de acurácia, mesmo na segmentação multiclasse em comparação à binária.

Em Huang et al. (2024), o modelo *Bayesian U-Net* com *Monte Carlo Dropout* foi aplicado a volumes de TCFC de 20 pacientes, segmentando múltiplas classes, incluindo lesões periapicais. O uso de aprendizado ativo (*Active Learning – AL*), com a função *Bayesian Active Learning by Disagreement (BALD)*, aumentou significativamente a sensibilidade na detecção de lesões para 84%, em comparação a 32% sem *AL* ($p < 0,0001$). O coeficiente *Dice* médio após *AL* foi de $0,703 \pm 0,166$ para todas as classes, enquanto, especificamente para a classe “lesão”, os métodos *BALD* e *Maximum Entropy (ME)* alcançaram valores próximos a 0,504 e 0,501, respectivamente, evidenciando melhoria substancial no desempenho da IA quando orientada por seleção ativa de amostras.

3.13 Comparação entre modelos de inteligência artificial

Fernandes et al. (2025) avaliaram a aplicação de modelos de inteligência artificial (IA) para classificação em imagens radiográficas derivadas de tomografias computadorizadas de feixe cônico (TCFC), totalizando 3.600 imagens distribuídas entre quatro diferentes tarefas. As tarefas incluíram: classificação da aparência dos seios maxilares, análise do formato de incisivos maxilares e mandibulares, detecção da presença do forame mentual em seções pré-molares mandibulares e avaliação da relação posicional do terceiro molar mandibular com o canal do nervo alveolar inferior (*Inferior Alveolar Nerve Canal – IANC*).

Foram comparadas três arquiteturas de aprendizado profundo: uma *Convolutional Neural Network* (CNN) clássica, um *Vision Transformer* (ViT) e um *gated Multi-Layer Perceptron* (gMLP). A CNN apresentou acurácia variando aproximadamente entre 0,71 e 0,99, enquanto o ViT alcançou desempenho semelhante, entre 0,74 e 0,98, mostrando vantagem marginal em algumas tarefas — especialmente na detecção de contato entre o terceiro molar e o IANC —, com área sob a curva (AUC) de aproximadamente 0,80–0,83, em comparação com 0,77–0,79 da CNN. O gMLP apresentou desempenho ligeiramente inferior, com acurácia variando entre 0,65 e 0,98.

As métricas de discriminação, incluindo as curvas *Receiver Operating Characteristic* (ROC) e *precision-recall*, foram elevadas para a maioria das tarefas: CNN (~0,77–1,00), ViT (~0,80–1,00) e gMLP (~0,73–1,00). O reconhecimento de incisivos foi realizado com alto desempenho por todos os modelos, sendo o ViT marginalmente superior em algumas métricas de AUC. Em contraste, a tarefa de determinar o contato entre o terceiro molar mandibular e o canal do nervo alveolar inferior apresentou os menores valores de desempenho geral, refletindo a maior complexidade dessa análise e a dificuldade dos modelos em captar detalhes anatômicos sutis necessários para essa avaliação.

4. DISCUSSÃO

Os resultados do presente estudo reforçam a tendência crescente observada na literatura quanto à eficácia das abordagens baseadas em inteligência artificial (IA) para a segmentação automática de imagens obtidas por tomografia computadorizada de feixe cônico (TCFC). Protocolos que combinam arquiteturas tridimensionais, como a 3D U-Net, com etapas de refinamento manual, têm demonstrado alta precisão na delimitação da maxila e de defeitos ósseos (WANG et al., 2021), confirmando o potencial dessas ferramentas para aplicações clínicas em larga escala. A identificação de hipoplasia significativa no lado da fissura, especialmente nas regiões da abertura piriforme e da crista alveolar, corrobora achados prévios que apontam a assimetria estrutural como característica marcante em pacientes com fissura labiopalatina. Além disso, verificou-se que parâmetros morfológicos — como altura, largura e volume do defeito — influenciam diretamente a variabilidade anatômica da maxila (MIRANDA et al., 2023), sugerindo que modelos tridimensionais automatizados podem fornecer métricas objetivas e padronizadas para o diagnóstico e acompanhamento desses casos.

De maneira geral, os estudos revisados evidenciam avanços expressivos no uso de IA e aprendizado profundo (deep learning) na análise e segmentação de estruturas ósseas da face, como maxila e mandíbula. Esse progresso, associado à expressiva redução no tempo de processamento, representa um salto qualitativo para a prática clínica, sobretudo em contextos que envolvem grandes volumes de dados (XU et al., 2025). O uso de arquiteturas tridimensionais mostra-se especialmente vantajoso para estruturas complexas — como seios maxilares, côndilos mandibulares e canais mandibulares —, pois permite capturar de forma mais eficiente a continuidade espacial e o contexto volumétrico.

Diversos estudos apontam que a aplicação da IA na análise de TCFC pode transformar significativamente o fluxo clínico em diferentes especialidades odontológicas, incluindo ortodontia, implantodontia, cirurgia bucomaxilofacial e periodontia (TANG et al., 2025; FU et al., 2024; AL-ASALI et al., 2024). Ferramentas automáticas baseadas em redes convolucionais têm apresentado resultados excelentes em tarefas como segmentação de dentes com brackets, caninos

impactados, fossas glenoides e seios maxilares, reduzindo o tempo de processamento de horas para segundos (AYIDH ALQAHTANI et al., 2023; SWAITY et al., 2024; BAYRAKDAR et al., 2024). Essa eficiência reforça o potencial das tecnologias para otimizar o planejamento de implantes, cirurgias ortognáticas e reconstruções ósseas, além de aprimorar o diagnóstico de defeitos periodontais, fraturas e patologias periapicais.

No campo da implantodontia, a IA tem se mostrado eficaz na detecção e segmentação de implantes, na localização do canal mandibular e na delimitação de áreas edêntulas, promovendo maior segurança cirúrgica e padronização nos planejamentos virtuais (KURT BAYRAKDAR et al., 2021; KWAK et al., 2020). Modelos de aprendizado profundo vêm apresentando desempenho equivalente ou até superior ao de especialistas humanos, com tempos de execução até 60 vezes menores (ELGARBA et al., 2023).

Em termos diagnósticos, soluções comerciais como o Diagnocat e o DentalSegmentator demonstraram alta sensibilidade e especificidade na detecção de lesões periapicais, alterações condilares e lesões de cárie dentária, consolidando-se como sistemas de apoio à decisão clínica (ALLIHAIBI et al., 2025; ORHAN et al., 2023; EZHOV et al., 2021). Contudo, o desempenho desses modelos ainda depende fortemente da qualidade da imagem e da padronização dos protocolos de aquisição, reforçando a necessidade de supervisão e validação por especialistas (AMASYA et al., 2023). A aplicação da IA também tem se expandido para análises morfológicas, cefalométricas e forenses, com modelos híbridos — combinando Convolutional Neural Networks (CNNs) e Vision Transformers — superando especialistas humanos na classificação da maturação da sutura palatina média e na estimativa de idade e sexo por medidas cefalométricas, evidenciando o caráter multidisciplinar e crescente da tecnologia (HAMIDI et al., 2024).

Apesar dos avanços, persistem desafios importantes. A generalização dos modelos ainda é limitada, pois muitos estudos utilizam dados provenientes de uma única instituição ou equipamento de TCFC, com protocolos homogêneos e amostras restritas a determinadas faixas etárias ou populações (VINAYAHALINGAM et al., 2023; MOUFTI et al., 2023). Essa limitação reduz a robustez dos algoritmos diante

da variabilidade clínica real, caracterizada por diferenças anatômicas, artefatos metálicos, variações de densidade e condições patológicas diversas. Diversas pesquisas demonstram queda de desempenho dos modelos em imagens com restaurações metálicas, edentulismo ou anatomias atípicas, reforçando a necessidade de validações multicêntricas e de bancos de dados mais heterogêneos (ELSONBATY et al., 2025; KAZIMIERCZAK et al., 2024).

Outro aspecto relevante é a dependência de refinamentos manuais para atingir precisão ideal. Embora as segmentações automáticas sejam altamente confiáveis, ajustes ainda se fazem necessários em regiões de borda ou baixo contraste, demonstrando que a automação plena ainda não foi alcançada (FERNANDES et al., 2025; TANG et al., 2025). Modelos baseados em deformação de TCFCs de referência também enfrentam limitações em pacientes jovens, devido às variações anatômicas associadas ao crescimento. Além disso, diferenças entre imagens simuladas e radiografias reais exigem etapas adicionais de pré-processamento, como correção de alinhamento e equalização de histograma (REDUWAN et al., 2024).

A dependência de hardware e recursos computacionais de alta performance constitui outra barreira à adoção ampla dessas ferramentas (KWAK et al., 2020; YANG et al., 2025). Arquiteturas robustas, como Vision Transformers e modelos híbridos tridimensionais, demandam elevado poder de processamento e longos tempos de treinamento, o que limita sua implementação em clínicas de pequeno e médio porte. Nesse cenário, o desenvolvimento de modelos mais leves e otimizados, capazes de manter desempenho satisfatório em dispositivos com menor capacidade computacional, surge como uma prioridade emergente.

Embora os resultados atuais sejam altamente promissores, a consolidação da IA na odontologia requer esforços voltados à validação multicêntrica, à padronização de protocolos e à interoperabilidade entre plataformas. A criação de bancos de dados amplos, diversificados e devidamente anotados é essencial para reduzir vieses regionais e ampliar a generalização dos algoritmos. Além disso, a adoção clínica plena dessas tecnologias depende do treinamento dos profissionais, do

desenvolvimento de interfaces intuitivas e da integração com sistemas de planejamento e diagnóstico já consolidados.

Por fim, a literatura converge para a compreensão de que a IA deve ser encarada como uma ferramenta de suporte, e não como substituta do julgamento clínico. A supervisão humana permanece indispensável, sobretudo em casos complexos ou com artefatos intensos. Futuras pesquisas devem priorizar o aprimoramento das arquiteturas de rede, o uso de aprendizado multimodal e o desenvolvimento de algoritmos de pós-processamento mais avançados, de modo a ampliar a precisão, a confiabilidade e a aplicabilidade dos modelos em cenários clínicos reais. Assim, a inteligência artificial consolida-se como uma aliada promissora e cada vez mais indispensável à odontologia digital contemporânea.

5. CONCLUSÃO

A análise dos estudos revisados evidencia que a inteligência artificial (IA), especialmente por meio de arquiteturas de aprendizado profundo como a *3D U-Net*, representa um avanço substancial na segmentação automática e na análise tridimensional de imagens obtidas por tomografia computadorizada de feixe cônico (TCFC), alcançando desempenho comparável ao de especialistas humanos e reduzindo significativamente o tempo de processamento. Modelos tridimensionais demonstram alta eficácia na caracterização volumétrica de estruturas complexas e na identificação de assimetrias, enquanto a integração entre segmentação automática e refinamento manual equilibra precisão anatômica e eficiência operacional. Apesar de desafios persistentes — como a limitação de bases de dados heterogêneas, a necessidade de ajustes manuais e a dependência de hardware de alto desempenho —, a IA consolida-se como ferramenta indispensável de apoio à decisão clínica, promovendo diagnósticos mais precisos, padronizados e ágeis, e reafirmando que a integração entre tecnologia e expertise humana é essencial para o futuro da odontologia digital.

REFERÊNCIAS

ABESI, Farida; JAMALI, Atena Sadat; ZAMANI, Mohammad. Accuracy of artificial intelligence in the detection and segmentation of oral and maxillofacial structures using cone-beam computed tomography images: a systematic review and meta-analysis. *Polish Journal of Radiology*, v. 88, n. 1, p. e256–e263, 2023.

AL-ASALI, Mohammed et al. Deep learning-based approach for 3D bone segmentation and prediction of missing tooth region for dental implant planning. *Scientific Reports*, v. 14, n. 1, p. 13888, 2024.

ALLIHAIBI, Marwa; KOLLER, Garrit; MANNOCCI, Francesco. Diagnostic accuracy of an artificial intelligence-based platform in detecting periapical radiolucencies on cone-beam computed tomography scans of molars. *Journal of Dentistry*, v. 160, n. 105854, p. 105854, 2025.

AL-SAREM, Mohammed et al. Enhanced tooth region detection using pretrained deep learning models. *International Journal of Environmental Research and Public Health*, v. 19, n. 22, p. 15414, 2022.

ALSOMALI, Mona et al. Development of a deep learning model for automatic localization of radiographic markers of proposed dental implant site locations. *The Saudi Dental Journal*, v. 34, n. 3, p. 220–225, 2022.

ALTUN, Oğuzhan et al. Automatic maxillary sinus segmentation and pathology classification on cone-beam computed tomographic images using deep learning. *BMC Oral Health*, v. 24, n. 1, p. 1208, 2024.

AMASYA, Hakan et al. Evaluation of a decision support system developed with deep learning approach for detecting dental caries with cone-beam computed tomography imaging. *Diagnostics*, v. 13, n. 22, p. 3471, 2023.

AYIDH ALQAHTANI, Khalid et al. Deep convolutional neural network-based automated segmentation and classification of teeth with orthodontic brackets on

cone-beam computed-tomographic images: a validation study. *European Journal of Orthodontics*, v. 45, n. 2, p. 169–174, 2023.

BAYRAKDAR, Ibrahim Sevki et al. Artificial intelligence system for automatic maxillary sinus segmentation on cone beam computed tomography images. *Dentomaxillofacial Radiology*, v. 53, n. 4, p. 256–266, 2024.

CALAZANS, Maria Alice Andrade et al. Automatic classification system for periapical lesions in cone-beam computed tomography. *Sensors*, v. 22, n. 17, p. 6481, 2022.

CHEN, Jiayi. Convolutional neural network for maxillary sinus segmentation based on the U-Net architecture at different planes in the Chinese population: a semantic segmentation study. *BMC Oral Health*, v. 25, n. 1, p. 961, 2025.

CHEN, Si et al. Machine learning in orthodontics: introducing a 3D auto-segmentation and auto-landmark finder of CBCT images to assess maxillary constriction in unilateral impacted canine patients. *The Angle Orthodontist*, v. 90, n. 1, p. 77–84, 2020.

CHOI, Hanseung et al. Deep learning-based fully automatic segmentation of the maxillary sinus on cone-beam computed tomographic images. *Scientific Reports*, v. 12, n. 1, p. 14009, 2022.

CHUNG, Eun-Ji et al. Effectiveness of cone-beam computed tomography-generated cephalograms using artificial intelligence cephalometric analysis. *Scientific Reports*, v. 12, n. 1, p. 20585, 2022.

CHUNG, Eun-Ji et al. Validation of 2D lateral cephalometric analysis using artificial intelligence-processed low-dose cone beam computed tomography. *Heliyon*, v. 10, n. 21, p. e39445, 2024.

CUI, Zhiming et al. A fully automatic AI system for tooth and alveolar bone segmentation from cone-beam CT images. *Nature Communications*, v. 13, n. 1, p. 2096, 2022.

DA ANDRADE-BORTOLETTO, Maria Fernanda Silva et al. Comparison of AI-powered tools for CBCT-based mandibular incisive canal segmentation: a validation study. *Clinical Oral Implants Research*, v. 36, n. 9, p. 1086–1094, 2025.

DA SILVA, Renan Lucio Berbel et al. Automatic segmentation and classification of frontal sinuses for sex determination from CBCT scans using a two-stage anatomy-guided attention network. *Scientific Reports*, v. 14, n. 1, p. 11750, 2024.

DE DUMAST, Priscille et al. A web-based system for neural network based classification in temporomandibular joint osteoarthritis. *Computerized Medical Imaging and Graphics*, v. 67, p. 45–54, 2018.

DENG, H. H. et al. Clinical feasibility of deep learning-based automatic head CBCT image segmentation and landmark detection in computer-aided surgical simulation for orthognathic surgery. *International Journal of Oral and Maxillofacial Surgery*, v. 52, n. 7, p. 793–800, 2023.

DIPALMA, Gianna et al. Artificial intelligence and its clinical applications in orthodontics: a systematic review. *Diagnostics*, v. 13, n. 24, p. 3677, 2023.

DOGAN, Ozlem B.; BOYACIOGLU, Hatice; GOKSULUK, Dincer. Machine learning assessment of dental age classification based on cone-beam CT images: a different approach. *Dentomaxillofacial Radiology*, v. 53, n. 1, p. 67–73, 2024.

ELGARBA, Bahaaeldeen M. et al. Deep learning-based segmentation of dental implants on cone-beam computed tomography images: a validation study. *Journal of Dentistry*, v. 137, n. 104639, p. 104639, 2023.

ELSONBATY, Sara et al. Novel AI-based tool for primary tooth segmentation on CBCT using convolutional neural networks: a validation study. *International Journal of Paediatric Dentistry*, v. 35, n. 1, p. 97–107, 2025.

EŞER, Gözde et al. Classification of temporomandibular joint osteoarthritis on cone beam computed tomography images using artificial intelligence system. *Journal of Oral Rehabilitation*, v. 50, n. 9, p. 758–766, 2023.

ESTRELLA, Navarro-Fraile et al. AI-aided volumetric root resorption assessment following personalized forces in orthodontics: preliminary results of a randomized clinical trial. *The Journal of Evidence-Based Dental Practice*, v. 25, n. 2, p. 102095, 2025.

EZHOV, Matvey et al. Clinically applicable artificial intelligence system for dental diagnosis with CBCT. *Scientific Reports*, v. 11, n. 1, p. 15006, 2021.

FERNANDES, Fara A. et al. Preparing for downstream tasks in artificial intelligence for dental radiology: a baseline performance comparison of deep learning models. *Dentomaxillofacial Radiology*, v. 54, n. 2, p. 149–162, 2025.

FONTENELE, Rocharles Cavalcante et al. Convolutional neural network-based automated maxillary alveolar bone segmentation on cone-beam computed tomography images. *Clinical Oral Implants Research*, v. 34, n. 6, p. 565–574, 2023.

FU, W. T. et al. Clinically oriented CBCT periapical lesion evaluation via 3D CNN algorithm. *Journal of Dental Research*, v. 103, n. 1, p. 5–12, 2024.

GERHARDT, Maurício do Nascimento et al. Automated detection and labelling of teeth and small edentulous regions on cone-beam computed tomography using convolutional neural networks. *Journal of Dentistry*, v. 122, n. 104139, p. 104139, 2022.

GUMUSSOY, Ismail et al. AI-powered segmentation of bifid mandibular canals using CBCT. *BMC Oral Health*, v. 25, n. 1, p. 907, 2025a.

GÜMÜSSOY, Ismail et al. Automated 3D segmentation of the hyoid bone in CBCT using nnU-Net v2: a retrospective study on model performance and potential clinical utility. *BMC Medical Imaging*, v. 25, n. 1, p. 217, 2025.

GUMUSSOY, Ismail et al. Automatic segmentation of the infraorbital canal in CBCT images: anatomical structure recognition using artificial intelligence. *Diagnostics*, v. 15, n. 13, p. 1713, 2025b.

GÜRSES, Barış Oğuz et al. A support vector machine-based algorithm to identify bisphosphonate-related osteonecrosis throughout the mandibular bone by using

cone beam computerized tomography images. *Dentomaxillofacial Radiology*, v. 52, n. 4, p. 20220390, 2023.

HAMIDI, Omid; AFRASIABI, Mahlagha; NAMAKI, Marjan. GADNN: a revolutionary hybrid deep learning neural network for age and sex determination utilizing cone beam computed tomography images of maxillary and frontal sinuses. *BMC Medical Research Methodology*, v. 24, n. 1, p. 50, 2024.

HAYLAZ, Emre et al. Automatic segmentation of the nasolacrimal canal: application of the nnU-Net v2 model in CBCT imaging. *Journal of Clinical Medicine*, v. 14, n. 3, p. 778, 2025.

HERNANDEZ, Andrea Kristine T. et al. Trueness of artificial intelligence-based, manual, and global thresholding segmentation protocols for human mandibles. *Journal of Prosthodontics*, n. jopr.70008, 2025.

HSU, Kang et al. Improving performance of deep learning models using 3.5D U-Net via majority voting for tooth segmentation on cone beam computed tomography. *Scientific Reports*, v. 12, n. 1, p. 19809, 2022.

HUANG, Jiayu et al. Uncertainty-based active learning by Bayesian U-Net for multi-label cone-beam CT segmentation. *Journal of Endodontics*, v. 50, n. 2, p. 220–228, 2024.

HUNG, Kuofeng et al. Current applications, opportunities, and limitations of AI for 3D imaging in dental research and practice. *International Journal of Environmental Research and Public Health*, v. 17, n. 12, p. 4424, 2020.

ISSA, Julien et al. Assessing the accuracy of artificial intelligence in mandibular canal segmentation compared to semi-automatic segmentation on cone-beam computed tomography images. *Polish Journal of Radiology*, v. 90, p. e172–e179, 2025.

JÄRNSTEDT, Jorma et al. Reproducibility analysis of automated deep learning based localisation of mandibular canals on a temporal CBCT dataset. *Scientific Reports*, v. 13, n. 1, p. 14159, 2023.

JHA, Nayansi et al. Fully automated condyle segmentation using 3D convolutional neural networks. *Scientific Reports*, v. 12, n. 1, p. 20590, 2022.

JINDANIL, T. et al. Artificial intelligence-based incisive canal visualization for preventing and detecting post-implant injury, using cone beam computed tomography. *International Journal of Oral and Maxillofacial Surgery*, v. 54, n. 8, p. 769–776, 2025.

JOHARI, Masume et al. Detection of vertical root fractures in intact and endodontically treated premolar teeth by designing a probabilistic neural network: an ex vivo study. *Dentomaxillofacial Radiology*, v. 46, n. 2, p. 20160107, 2017.

KAZIMIERCZAK, Wojciech et al. Periapical lesions in panoramic radiography and CBCT imaging: assessment of AI's diagnostic accuracy. *Journal of Clinical Medicine*, v. 13, n. 9, p. 2709, 2024a.

KAZIMIERCZAK, Wojciech et al. Endodontic treatment outcomes in cone beam computed tomography images: assessment of the diagnostic accuracy of AI. *Journal of Clinical Medicine*, v. 13, n. 14, p. 4116, 2024b.

KIM, Na-Hyun et al. Preclinical and preliminary evaluation of perceived image quality of AI-processed low-dose CBCT analysis of a single tooth. *Bioengineering*, v. 11, n. 6, p. 576, 2024.

KIM, Young Hyun et al. Automated cortical thickness measurement of the mandibular condyle head on CBCT images using a deep learning method. *Scientific Reports*, v. 11, n. 1, p. 14852, 2021.

KURT BAYRAKDAR, Sevda et al. A deep learning approach for dental implant planning in cone-beam computed tomography images. *BMC Medical Imaging*, v. 21, n. 1, p. 86, 2021.

KURT-BAYRAKDAR, Sevda et al. Advancing periodontal diagnosis: harnessing advanced artificial intelligence for patterns of periodontal bone loss in cone-beam computed tomography. *Dentomaxillofacial Radiology*, v. 54, n. 4, p. 268–278, 2025.

KWAK, Gloria Hyunjung et al. Automatic mandibular canal detection using a deep convolutional neural network. *Scientific Reports*, v. 10, n. 1, p. 5711, 2020.

LI, Yan et al. Evaluation of children's oral diagnosis and treatment using imaging examination using AI based Internet of Things. *Technology and Health Care*, v. 32, n. 3, p. 1323–1340, 2024.

LIU, Yu et al. Fully automatic AI segmentation of oral surgery-related tissues based on cone beam computed tomography images. *International Journal of Oral Science*, v. 16, n. 1, p. 34, 2024.

MAHABOB, Nazargi et al. AI-driven risk stratification of the lingual foramen: a CBCT-based prevalence and morphological analysis. *Healthcare*, v. 13, n. 13, p. 1515, 2025.

MAHDAVIFAR, Sare; FAKHRAHMAD, Seyed Mostafa; ANSARIFARD, Elham. Estimating the severity of oral lesions via analysis of cone beam computed tomography reports: a proposed deep learning model. *International Dental Journal*, v. 75, n. 1, p. 135–143, 2025.

MERDIETIO BOEDI, R. *et al.* Machine learning assisted 5-part tooth segmentation method for CBCT-based dental age estimation in adults. *The journal of forensic odonto-stomatology*, v. 42, n. 1, p. 22–29, 2024.

MILANI, Omid Halimi et al. Automated classification of midpalatal suture maturation stages from CBCTs using an end-to-end deep learning framework. *Scientific Reports*, v. 15, n. 1, p. 18783, 2025.

MINNEMA, Jordi et al. Segmentation of dental cone-beam CT scans affected by metal artifacts using a mixed-scale dense convolutional neural network. *Medical Physics*, v. 46, n. 11, p. 5027–5035, 2019.

MIRANDA, Felicia et al. Interpretable artificial intelligence for classification of alveolar bone defect in patients with cleft lip and palate. *Scientific Reports*, v. 13, n. 1, p. 15861, 2023.

MORGAN, Nermin et al. Convolutional neural network for automatic maxillary sinus segmentation on cone-beam computed tomographic images. *Scientific Reports*, v. 12, n. 1, p. 7523, 2022.

MOUFTI, Mohammad Adel et al. Developing an artificial intelligence solution to autosegment the edentulous mandibular bone for implant planning. *European Journal of Dentistry*, v. 17, n. 4, p. 1330–1337, 2023.

NAMATEVS, Ivars et al. Modular neural networks for osteoporosis detection in mandibular cone-beam computed tomography scans. *Tomography*, v. 9, n. 5, p. 1772–1786, 2023.

NOGUEIRA-REIS, Fernanda et al. Three-dimensional maxillary virtual patient creation by convolutional neural network-based segmentation on cone-beam computed tomography images. *Clinical Oral Investigations*, v. 27, n. 3, p. 1133–1141, 2023.

NOGUEIRA-REIS, Fernanda et al. Full virtual patient generated by artificial intelligence-driven integrated segmentation of craniomaxillofacial structures from CBCT images. *Journal of Dentistry*, v. 141, n. 104829, p. 104829, 2024.

OLIVEIRA, Matheus L. et al. Development and evaluation of a deep learning model to reduce exomass-related metal artefacts in cone-beam CT: an ex vivo study using porcine mandibles. *Dentomaxillofacial Radiology*, v. 54, n. 2, p. 109–117, 2025.

ON, Sungchul et al. Improving accuracy for inferior alveolar nerve segmentation with multi-label of anatomical adjacent structures using active learning in cone-beam computed tomography. *Scientific Reports*, v. 15, n. 1, p. 7441, 2025.

ORHAN, Kaan et al. AI-based automatic segmentation of craniomaxillofacial anatomy from CBCT scans for automatic detection of pharyngeal airway evaluations in OSA patients. *Scientific Reports*, v. 12, n. 1, p. 11863, 2022.

ORHAN, Kaan et al. Assessing the reliability of CBCT-based AI-generated STL files in diagnosing osseous changes of the mandibular condyle: a comparative study with ground truth diagnosis. *Dentomaxillofacial Radiology*, v. 52, n. 7, p. 20230141, 2023.

OUSEHAL, Lahcen; MAJID, Amine; ELBENNA, S. The factors driving Moroccan orthodontists to adopt cone-beam imaging in clinical practice. *OAlib*, v. 11, n. 07, p. 1–15, 2024.

OZUDOGRU, S. et al. Artificial intelligence system for automatic tooth detection and numbering in the mixed dentition in CBCT. *European Journal of Paediatric Dentistry*, v. 26, n. 2, p. 140–146, 2025.

PALKOVICS, Daniel et al. Assessment of hard tissue changes after horizontal guided bone regeneration with the aid of deep learning CBCT segmentation. *Clinical Oral Investigations*, v. 29, n. 1, p. 59, 2025.

PAN, Xiao et al. Development and verification of a convolutional neural network-based model for automatic mandibular canal localization on multicenter CBCT images. *BMC Oral Health*, v. 25, n. 1, p. 1352, 2025.

PARK, Jungeun et al. Clinical validity and precision of deep learning-based cone-beam computed tomography automatic landmarking algorithm. *Imaging Science in Dentistry*, v. 54, n. 3, p. 240–250, 2024.

PIANYKH, Oleg S. et al. Continuous learning AI in radiology: implementation principles and early applications. *Radiology*, v. 297, n. 1, p. 6–14, 2020.

PICOLI, Fernando Fortes et al. Risk assessment of inferior alveolar nerve injury after wisdom tooth removal using 3D AI-driven models: a within-patient study. *Journal of Dentistry*, v. 139, n. 104765, p. 104765, 2023.

PIRAYESH, Zeynab et al. A hierarchical deep learning approach for diagnosing impacted canine-induced root resorption via cone-beam computed tomography. *BMC Oral Health*, v. 24, n. 1, p. 982, 2024.

REDUWAN, Nor Hidayah et al. Application of deep learning and feature selection technique on external root resorption identification on CBCT images. *BMC Oral Health*, v. 24, n. 1, p. 252, 2024.

ROONGRUANGSILP, Pathompong; NARKBUAKAEW, Walita; KHONGKHUNTHIAN, Pathawee. Performance of two different artificial intelligence models in dental implant

planning among four different implant planning software: a comparative study. *BMC Oral Health*, v. 25, n. 1, p. 984, 2025.

SAKAI, Takahiko et al. Development of artificial intelligence model for supporting implant drilling protocol decision making. *Journal of Prosthodontic Research*, v. 67, n. 3, p. 360–365, 2023.

SANTOS-JUNIOR, Airtton Oliveira et al. A novel artificial intelligence-powered tool for automated root canal segmentation in single-rooted teeth on cone-beam computed tomography. *International Endodontic Journal*, v. 58, n. 4, p. 658–671, 2025.

ŞENOL, D. et al. Gender estimation from morphometric measurements of mandibular lingula by using machine learning algorithms and artificial neural networks. *Nigerian Journal of Clinical Practice*, v. 27, n. 6, p. 732–738, 2024.

SERINDERE, Gozde et al. Evaluation of maxillary sinusitis from panoramic radiographs and cone-beam computed tomographic images using a convolutional neural network. *Imaging Science in Dentistry*, v. 52, n. 2, p. 187–195, 2022.

SHETTY, Shishir et al. Application of artificial intelligence-based detection of furcation involvement in mandibular first molar using cone beam tomography images: a preliminary study. *BMC Oral Health*, v. 24, n. 1, p. 1476, 2024.

SHETTY, Shishir et al. Accuracy of deep learning models in the detection of accessory ostium in coronal cone beam computed tomographic images. *Scientific Reports*, v. 15, n. 1, p. 8324, 2025a.

SHETTY, Shishir et al. Machine learning models in the detection of MB2 canal orifice in CBCT images. *International Dental Journal*, v. 75, n. 3, p. 1640–1648, 2025b.

SIN, Çağla et al. A deep learning algorithm proposal to automatic pharyngeal airway detection and segmentation on CBCT images. *Orthodontics & Craniofacial Research*, v. 24, suppl. 2, n. S2, p. 117–123, 2021.

SLIM, Marie Louise et al. AI-driven segmentation of the pulp cavity system in mandibular molars on CBCT images using convolutional neural networks. *Clinical Oral Investigations*, v. 28, n. 12, p. 650, 2024.

SWAITY, Abdullah et al. Deep learning driven segmentation of maxillary impacted canine on cone beam computed tomography images. *Scientific Reports*, v. 14, n. 1, p. 369, 2024.

TALAAT, Wael M. et al. An artificial intelligence model for the radiographic diagnosis of osteoarthritis of the temporomandibular joint. *Scientific Reports*, v. 13, n. 1, p. 15972, 2023.

TANG, Haomin et al. Prediction of midpalatal suture maturation stage based on transfer learning and enhanced vision transformer. *BMC Medical Informatics and Decision Making*, v. 24, n. 1, p. 232, 2024.

TANG, Haomin et al. Automatic segmentation and landmark detection of 3D CBCT images using semi-supervised learning for assisting orthognathic surgery planning. *Scientific Reports*, v. 15, n. 1, p. 8814, 2025.

TANIKAWA, Chihiro et al. Development of artificial intelligence-supported automatic three-dimensional surface cephalometry. *Orthodontics & Craniofacial Research*, v. 28, n. 4, p. 636–646, 2025.

USMAN, Muhammad et al. Dual-stage deeply supervised attention-based convolutional neural networks for mandibular canal segmentation in CBCT scans. *Sensors*, v. 22, n. 24, p. 9877, 2022.

VASEY, B. et al. DECIDE-AI: a new reporting guideline and its relevance to artificial intelligence studies in radiology. *Clinical Radiology*, v. 78, n. 2, p. 130–136, 2023.

VINAYAHALINGAM, Shankeeth et al. Deep learning for automated segmentation of the temporomandibular joint. *Journal of Dentistry*, v. 132, n. 104475, p. 104475, 2023.

WANG, H. et al. Multiclass CBCT image segmentation for orthodontics with deep learning. *Journal of Dental Research*, v. 100, n. 9, p. 943–949, 2021a.

WANG, Xiaoyu et al. 3D morphometric quantification of maxillae and defects for patients with unilateral cleft palate via deep learning-based CBCT image auto-segmentation. *Orthodontics & Craniofacial Research*, v. 24, suppl. 2, n. S2, p. 108–116, 2021b.

XIAO, Yanjun et al. Construction of a new automatic grading system for jaw bone mineral density level based on deep learning using cone beam computed tomography. *Scientific Reports*, v. 12, n. 1, p. 12841, 2022.

XU, Shuxi et al. Automatic detection of orthodontically induced external root resorption based on deep convolutional neural networks using CBCT images. *Scientific Reports*, v. 15, n. 1, p. 22984, 2025.

YANG, Pan et al. Detection of vertical root fractures by cone-beam computed tomography based on deep learning. *Dentomaxillofacial Radiology*, v. 52, n. 3, p. 20220345, 2023.

YANG, Su et al. Comparison of 2D, 2.5D, and 3D segmentation networks for mandibular canals in CBCT images: a study on public and external datasets. *BMC Oral Health*, v. 25, n. 1, p. 1126, 2025.

YASIN, Elham Tahsin et al. Automatic mandibular third molar and mandibular canal relationship determination based on deep learning models for preoperative risk reduction. *Clinical Oral Investigations*, v. 29, n. 4, p. 203, 2025.

YONG, Tae-Hoon et al. QCBCT-NET for direct measurement of bone mineral density from quantitative cone-beam CT: a human skull phantom study. *Scientific Reports*, v. 11, n. 1, p. 15083, 2021.

YOO, Yeon-Sun et al. Comparison of 2D, 2.5D, and 3D segmentation networks for maxillary sinuses and lesions in CBCT images. *BMC Oral Health*, v. 23, n. 1, p. 866, 2023.