

ANÁLISE COMPARATIVA DE ARQUITETURAS CONVOLUCIONAIS PRÉ-TREINADAS PARA CLASSIFICAÇÃO DE TRAUMATISMOS DENTÁRIOS.

**Warley da Costa Silva¹, Marco Antônio de Souza Leite Cuadros²,
Renner Sartório Camargo³, Pamela Barbosa dos Santos⁴**

¹Instituto Federal do Espírito Santo - IFES, Serra, Brasil (Warleydc.silva@gmail.com)

²Instituto Federal do Espírito Santo - IFES, Serra, Brasil

³Instituto Federal do Espírito Santo - IFES, Serra, Brasil

⁴Universidade Estadual de Campinas - Unicamp, Campinas, Brasil

Resumo: Este trabalho compara as arquiteturas pré-treinadas ResNet-50, EfficientNetV2-S e ConvNeXt-Tiny na classificação de imagens de traumatismos dentários utilizando *Transfer Learning*. Foi desenvolvido um framework experimental padronizado, empregando o mesmo conjunto de dados e protocolo de treinamento para todas as redes. Os resultados permitiram comparar o desempenho e a eficiência computacional das arquiteturas, contribuindo para o apoio ao reconhecimento inicial dessas lesões.

Palavras-chave: Traumatismos dentários; Deep Learning; Transfer Learning; Redes Neurais Convolucionais; Visão Computacional.

INTRODUÇÃO

Os traumatismos dentários constituem um importante problema de saúde pública devido à elevada incidência e ao impacto funcional, estético e psicossocial decorrente dessas lesões. Além de comprometerem a saúde bucal e a qualidade de vida dos pacientes, essas injúrias frequentemente precisam de tratamentos complexos e acompanhamento clínico prolongado. O prognóstico está diretamente relacionado à rapidez e à adequação das condutas adotadas nos primeiros momentos após o trauma, tornando o atendimento inicial um fator determinante para a preservação dos elementos dentários e para a redução de complicações futuras (Bourguignon et al., 2020).

De acordo com as Diretrizes da International Association of Dental Traumatology (IADT), revisadas em 2020, os traumatismos dentários em dentes permanentes são classificados conforme o tipo de tecido afetado: tecidos duros, tecidos de sustentação e tecidos moles e pela extensão do comprometimento pulpar e periodontal.

Essa taxonomia, amplamente utilizada em odontologia traumatológica, permite padronizar condutas e facilitar a comunicação entre profissionais, além de subsidiar modelos computacionais de

diagnóstico (Levin et al., 2020; Bourguignon et al., 2020; Fouad et al., 2020).

A fratura de esmalte é considerada o trauma de menor gravidade, caracterizando-se por trincas ou perda superficial da camada externa do dente. Embora muitas vezes assintomática, pode causar sensibilidade ou afetar a estética. O tratamento envolve apenas o polimento das bordas ou restauração adesiva, dependendo da extensão (Levin et al., 2020).

No caso da fratura de esmalte e dentina, há perda mais profunda de estrutura dentária, envolvendo esmalte e dentina, mas sem exposição pulpar, os pacientes relatam dor à mastigação e sensibilidade térmica, e o exame clínico revela superfície irregular e bordas cortantes. A conduta recomendada é a proteção da dentina exposta com resina composta ou ionômero de vidro, seguida de acompanhamento clínico (Bourguignon et al., 2020).

Quando a fratura atinge a polpa dentária, ocorre exposição pulpar e sangramento, frequentemente acompanhado de dor intensa. O manejo clínico depende do tempo decorrido desde o trauma e da maturidade radicular, podendo incluir pulpotomia parcial (em dentes imaturos) ou tratamento endodôntico completo (em dentes maduros).

A IADT reforça que a proteção pulpar imediata é decisiva para a manutenção da vitalidade e prevenção de necrose (Fouad et al., 2020).

Figura 1 - Trauma em esmalte (círculo) e esmalte e dentina (retângulo); fratura coronária envolvendo esmalte, dentina e polpa; exposição pulpar.



Fonte: Adaptado de Sanabe, M. Cavalcante, L. Coldebella, C. et al. (2009).

A luxação extrusiva corresponde ao deslocamento parcial do dente para fora do alvéolo, mantendo ainda alguma inserção periodontal. Clinicamente, o dente apresenta-se alongado, com mobilidade aumentada e dor à percussão.

A literatura destaca que o reposicionamento manual imediato e a contenção flexível por duas semanas são medidas essenciais para o sucesso terapêutico (Levin et al., 2020).

Figura 2 - Luxação Extrusiva de 3 mm e Deslocamento palatal.



Fonte: Adaptado de Spinas. E, Zangani. A, Mallus. T. et al. (2024).

A intrusão é caracterizada pelo deslocamento do dente para o interior do alvéolo, em direção apical, resultando em encurtamento visível da coroa. É uma das lesões mais graves da dentição permanente, frequentemente associada a fratura alveolar e risco da anquilose. Segundo Bourguignon et al. (2020), quanto menor o tempo de impacto e mais leve o deslocamento (até 3 mm), maior a chance de re-erupção espontânea.

Entretanto, em situações de urgência, o acesso imediato ao cirurgião-dentista nem sempre é possível, fazendo com que pacientes e familiares necessitem tomar decisões sem orientação especializada.

Nesse contexto, soluções computacionais capazes de auxiliar o reconhecimento inicial dos traumatismos

dentários tornam-se ferramentas promissoras para apoiar a tomada de decisão até a realização do atendimento profissional.

Figura 3 - Fotografia Clínica pré-operatória de um dente incisivo central intruído.



Fonte: Adaptado de Sule, K., Malik, A., Kothari, A., Rao, V. et al. (2024).

Na luxação lateral, o dente é deslocado horizontalmente, geralmente em direção vestibular ou palatina, acompanhada por ruptura parcial do osso alveolar.

A IADT recomenda reposicionamento manual sob anestesia e contenção flexível por até quatro semanas (Fouad et al., 2020).

Figura 4 - Luxação lateral do dente.



Fonte: Adaptado de DentaGama (2023).

A avulsão representa a forma mais severa de traumatismo dentário. Nesses casos, o sucesso do tratamento depende essencialmente do tempo seco extraoral e do meio de armazenamento.

Fouad et al. (2020) e Levin et al. (2020) ressaltam que o reimplante imediato, preferencialmente dentro de 15 minutos, é o principal fator prognóstico de sobrevida periodontal. Quando o reimplante é tardio, ocorrem frequentemente reabsorções inflamatórias e substitutivas.

Figura 5 – Vista intraoral evidenciando a avulsão dos dentes 11, 21 e 22.



Fonte: Adaptado de El Khatib et al. (2022).

A diversidade clínica dessas lesões evidencia a complexidade do reconhecimento inicial, justificando a investigação de métodos computacionais capazes de auxiliar sua classificação automática.

Estudos recentes demonstram que técnicas baseadas em Inteligência Artificial vêm apresentando elevado potencial para auxiliar a interpretação de imagens odontológicas e apoiar a identificação de diferentes tipos de traumatismos dentários (Çege et al., 2025; Kuru, Aşık e Demir, 2025).

O sucesso do tratamento desses traumatismos depende, em grande parte, da conduta adotada nos primeiros minutos após o acidente. Entretanto, evidências recentes demonstram que mesmo indivíduos frequentemente expostos ao risco, como atletas, apresentam conhecimento insuficiente sobre o manejo inicial dessas lesões. Entre os principais problemas observados estão o armazenamento inadequado do dente avulsionado, a incerteza quanto ao reimplante imediato e o desconhecimento das medidas corretas de primeiros socorros, reforçando a necessidade de ferramentas capazes de fornecer suporte rápido e baseado em evidências (Santinoni et al., 2024).

Paralelamente, os avanços da Inteligência Artificial, especialmente da Visão Computacional e do Aprendizado Profundo (*Deep Learning*), ampliaram significativamente o desenvolvimento de sistemas capazes de analisar imagens odontológicas com elevado desempenho. Entre as técnicas mais empregadas destacam-se as Redes Neurais Convolucionais (Convolutional Neural Networks – CNNs), que vêm sendo amplamente utilizadas em tarefas de classificação de imagens médicas e odontológicas devido à sua capacidade de extrair automaticamente características relevantes e identificar padrões complexos presentes nas imagens (Villena et al., 2024).

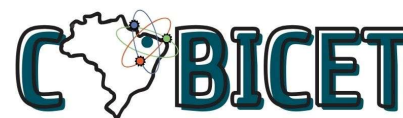
Uma estratégia amplamente adotada nesse contexto é o *Transfer Learning*, que permite reutilizar modelos previamente treinados em grandes bases de imagens, reduzindo o tempo de treinamento e aumentando a capacidade de generalização quando aplicado a conjuntos de dados específicos. Entre as arquiteturas mais utilizadas destacam-se a ResNet-50, a EfficientNetV2-S e a ConvNeXt-Tiny, que representam diferentes gerações de redes convolucionais e apresentam características distintas em termos de desempenho e eficiência computacional.

Embora estudos recentes demonstrem resultados promissores da Inteligência Artificial aplicada à Odontologia, observa-se que ainda são escassos trabalhos que realizem comparações sistemáticas entre diferentes arquiteturas pré-treinadas utilizando um mesmo protocolo experimental para classificação de traumatismos dentários. A utilização de conjuntos de dados distintos, diferentes estratégias de treinamento e métricas variadas dificulta a comparação direta entre os modelos e a identificação da arquitetura mais adequada para essa aplicação.

Diante desse contexto, este trabalho apresenta uma análise comparativa das arquiteturas pré-treinadas ResNet-50, EfficientNetV2-S e ConvNeXt-Tiny para classificação de imagens de traumatismos dentários utilizando a estratégia de *Transfer Learning*. Para garantir uma comparação justa entre os modelos, foi desenvolvido um framework experimental padronizado, no qual todas as arquiteturas foram treinadas e avaliadas utilizando o mesmo conjunto de dados, os mesmos hiperparâmetros e o mesmo protocolo experimental. O objetivo é identificar a arquitetura que apresenta o melhor desempenho para apoiar o reconhecimento inicial de traumatismos dentários, contribuindo para o desenvolvimento de ferramentas baseadas em Inteligência Artificial aplicadas à Odontologia.

MATERIAL E MÉTODOS

O presente estudo foi desenvolvido com o objetivo de comparar o desempenho de diferentes arquiteturas de Redes Neurais Convolucionais (CNNs) pré-treinadas na classificação automática de traumatismos dentários a partir de imagens clínicas. Para isso, todo o fluxo experimental foi estruturado de forma padronizada e reproduzível, garantindo que as diferenças de desempenho observadas entre os modelos fossem decorrentes exclusivamente de suas características arquitetônicas e não de variações associadas ao pré-processamento, divisão dos dados ou estratégia de treinamento. A escolha desse desenho metodológico segue recomendações recentes da literatura, que destacam a importância de protocolos experimentais padronizados para a avaliação de sistemas de



Inteligência Artificial aplicados à Odontologia (Alshahri et al., 2023; Tanaka et al., 2024; Parveen e Dar, 2023).

O conjunto de imagens utilizado foi composto por 1.440 fotografias clínicas de traumatismos dentários, distribuídas em 7 categorias clínicas, de acordo com a taxonomia adotada pela *International Association of Dental Traumatology* (IADT), além de uma oitava classe, denominada "Várias Fraturas", destinada às imagens que apresentavam simultaneamente mais de um tipo de traumatismo dentário (Bourguignon et al., 2020).

O conjunto de dados apresenta elevada heterogeneidade visual, incluindo variações de iluminação, distância da câmera, profundidade de campo, presença de sangue e tecidos adjacentes, além de diferenças anatômicas entre os pacientes. Essa variabilidade foi mantida intencionalmente, uma vez que estudos recentes demonstram que conjuntos de dados excessivamente padronizados tendem a superestimar o desempenho dos modelos e reduzir sua capacidade de generalização em aplicações clínicas (Uddin et al., 2024; Zhang et al., 2024; Foster e Del Fiore, 2024).

O pré-processamento das imagens constituiu uma etapa essencial para garantir a padronização visual do conjunto de dados. Inicialmente, as imagens foram organizadas na plataforma Roboflow para gerenciamento do banco de dados. Em seguida, foram convertidas para o formato RGB, redimensionadas para 224×224 pixels, dimensão compatível com as arquiteturas convolucionais avaliadas, e normalizadas para o intervalo $[0,1]$, favorecendo a estabilidade do processo de treinamento. Posteriormente, o conjunto de dados foi dividido em três subconjuntos mutuamente exclusivos, correspondentes a 70% para treinamento, 20% para validação e 10% para teste, preservando a distribuição das oito classes e evitando o vazamento de dados entre as etapas de treinamento e avaliação.

Foram avaliadas três arquiteturas convolucionais pré-treinadas amplamente empregadas em tarefas de classificação de imagens médicas: ResNet-50, EfficientNetV2-S e ConvNeXt-Tiny. Todas as arquiteturas foram inicializadas com pesos previamente treinados na base ImageNet e posteriormente adaptadas para a classificação das oito categorias de traumatismos dentários por meio da estratégia de *Transfer Learning*. A seleção desses modelos buscou representar diferentes gerações de arquiteturas convolucionais, permitindo analisar o impacto das suas características estruturais no reconhecimento automático das lesões (Kim et al., 2023; Chen et al., 2023; Zhang et al., 2024; Santos e Ribeiro, 2024; Steiner, Müller e Köhler, 2024).

Para garantir equidade experimental, todas as arquiteturas foram submetidas ao mesmo protocolo de treinamento supervisionado, utilizando idênticos critérios de otimização, divisão dos dados e métricas de avaliação. Os hiperparâmetros foram definidos de forma a assegurar estabilidade durante o treinamento e minimizar problemas de sobreajuste ou subtreinamento. O treinamento foi realizado na plataforma Google Colaboratory utilizando aceleração por GPU NVIDIA Tesla T4, empregando imagens de entrada com resolução de 224×224 pixels, taxa de aprendizagem constante, tamanho de lote (*batch size*) igual a 16 e número fixo de épocas para todas as arquiteturas. O modelo com melhor desempenho no conjunto de validação foi automaticamente selecionado para avaliação final.

A avaliação foi realizada exclusivamente sobre o conjunto de teste, utilizando métricas amplamente empregadas em problemas de classificação multiclasse, incluindo Accuracy, Precision, Recall e F1-score. Além das métricas de desempenho, foram analisados aspectos relacionados à eficiência computacional, como tempo de treinamento, número de parâmetros e velocidade de inferência, permitindo uma comparação abrangente entre as arquiteturas investigadas. Esse conjunto de análises segue recomendações recentes da literatura, que ressaltam a importância de avaliar não apenas o desempenho preditivo, mas também a viabilidade prática da aplicação de modelos de Inteligência Artificial em ambientes clínicos (Parveen e Dar, 2023; Steiner, Müller e Köhler, 2024; Barrett e Cunningham, 2024; De Luca, 2024).

Todo o fluxo metodológico foi desenvolvido de forma padronizada e reprodutível, permitindo que o protocolo experimental possa ser replicado em futuros estudos e aplicado a conjuntos de dados mais amplos.

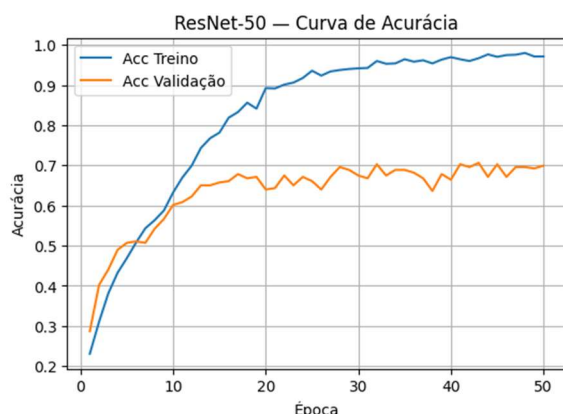
O presente trabalho contribui para a avaliação comparativa de arquiteturas convolucionais pré-treinadas voltadas ao reconhecimento inicial de traumatismos dentários, fornecendo evidências sobre o desempenho de diferentes modelos em um mesmo cenário experimental.

RESULTADOS E DISCUSSÃO

O comportamento do ResNet-50 durante o treinamento mostrou uma trajetória de aprendizado progressiva e relativamente estável, conforme documentado em aplicações médicas envolvendo redes residuais profundas (Kim et al., 2023; Chen et al., 2023). A curva de precisão mostra que o modelo rapidamente abandona o regime inicial de aleatoriedade, alcançando aproximadamente 60% de acurácia no conjunto de validação, nas primeiras 10 épocas, enquanto a precisão do treinamento se aproxima de 80%.

A partir desse ponto, as duas curvas mantêm tendências ascendentes, embora em velocidades diferentes: enquanto a precisão do treinamento evolui continuamente até atingir valores próximos de 97%, a precisão da validação se estabiliza na faixa de 68% a 70%.

Figura 6 – Curva de Acurácia – ResNet50.

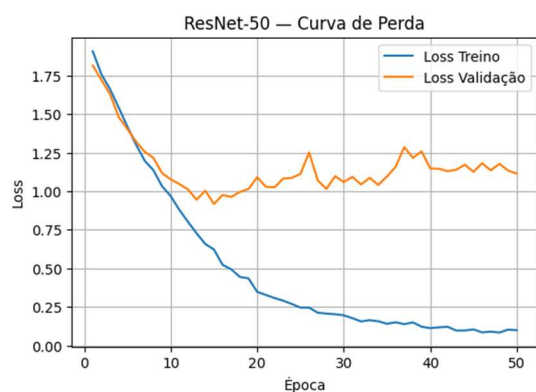


Fonte: Elaboração própria (2026).

Essa separação gradual revela um padrão típico de sobreajuste leve, no qual o modelo se adapta cada vez mais a exemplos de treinamento, sem produzir ganhos proporcionais em generalizabilidade. A análise da curva de perda reforça essa interpretação. A perda de treinamento diminui robustamente, atingindo valores muito baixos ao final das temporadas, enquanto a perda de validação atinge um platô intermediário em torno de 1,0 e permanece oscilante.

Essa divergência entre as perdas, confirma que o modelo captura representações significativas, mas começa a investir em otimização adicional no refinamento das características específicas do conjunto de treinamento (Parveen e Dar, 2023; Barrett e Cunningham, 2024).

Figura 7 – Curva de Perda – ResNet50.

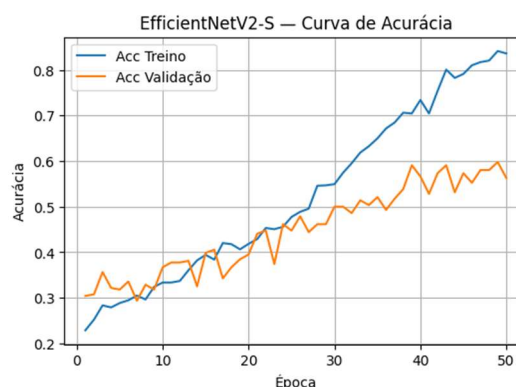


Fonte: Elaboração própria (2026).

Apesar disso, o treinamento é claramente bem-sucedido e mantém a capacidade de generalizar em um nível adequado (Chen et al., 2023; Zhang et al., 2024). O melhor desempenho de validação é registrado por volta da época 43, estabelecendo 70,63% como a precisão máxima validada do ResNet-50.

O EfficientNetV2-S apresentou o comportamento menos estável entre as três CNNs. A curva de precisão do treinamento segue uma trajetória ascendente consistente, atingindo valores acima de 80% ao final das temporadas, mas a precisão da validação progride de forma mais lenta e irregular, permanecendo predominantemente na faixa de 50% a 60%.

Figura 8 – Curva de Acurácia – EfficientNetV2-S.



Fonte : Elaboração própria (2026).

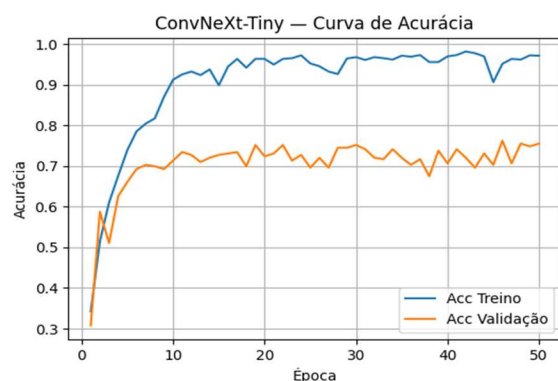
O treinamento do ConvNeXt-Tiny mostrou o comportamento mais eficiente e equilibrado entre as três arquiteturas.

A curva de precisão do treinamento demonstra aprendizado extremamente rápido: já nas primeiras 10 temporadas, o modelo atinge cerca de 85% das respostas corretas e continua evoluindo até estabilizar em valores próximos de 95%.

A precisão da validação, por outro lado, acompanha a tendência de crescimento, atingindo rapidamente a faixa de 70% nas primeiras temporadas e mantendo, com pequenas flutuações, entre 70% e 76% durante o restante do treinamento.

A distância entre as curvas é esperada, já que redes profundas inevitavelmente se ajustam de forma mais intensa aos dados de treinamento, mas a separação não se amplia com o tempo, o que revela estabilidade e capacidade de generalização (Zhang et al., 2024; Steiner, Müller e Köhler, 2024; Andrade e Gomes, 2024).

Figura 9 – Curva de Acurácia – ConvNext-Tiny.

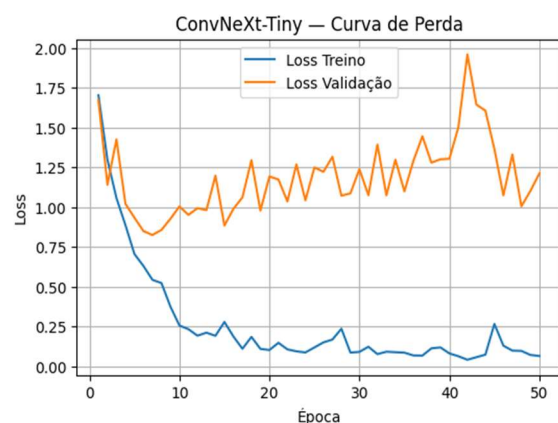


Fonte: Elaboração própria (2026).

Essa leitura é apoiada pela curva de perda. A perda de treinamento mostra uma queda acentuada, aproximando-se de valores próximos a zero nas últimas temporadas, enquanto a perda de validação, embora não siga uma trajetória monotonamente decrescente, permanece dentro de margens relativamente estáveis e próxima ao seu ponto mínimo (Zhang et al., 2024; Steiner, Müller e Köhler, 2024).

A ausência de uma tendência de alta acentuada que caracterizaria um sobreajuste severo indica que o modelo alcança um equilíbrio adequado entre complexidade e regularização (Zhang et al., 2024; Andrade e Gomes, 2024). O comportamento conjunto das curvas torna evidente por que o ConvNeXt-Tiny alcançou o melhor desempenho classificatório: a arquitetura foi capaz de aprender representações discriminativas rapidamente, sem perder a generalização durante todo o processo de otimização (Zhang et al., 2024; Steiner, Müller e Köhler, 2024; Andrade e Gomes, 2024).

Figura 10 – Curva de Perda – ConvNext-Tiny.



Fonte: Elaboração própria (2026).

A análise individualizada dos comportamentos de treinamento revela que os três modelos, apesar de terem sido submetidos ao mesmo protocolo de ajuste fino, respondem de forma significativamente diferente às condições impostas pelo banco de dados.

O ResNet-50 demonstra aprendizado sólido, embora com overfit moderado; EfficientNetV2-S demonstra instabilidade e dificuldade em generalizar; e o ConvNeXt-Tiny combina velocidade de convergência com estabilidade das métricas de validação, consolidando-se como a arquitetura de melhor desempenho (Zhang et al., 2024; Steiner, Müller e Köhler, 2024; Andrade e Gomes, 2024).

O comportamento do ConvNeXt-Tiny se destaca pela rapidez de convergência e estabilidade ao longo das estações. A curva de precisão do treinamento cresceu rapidamente desde as primeiras iterações, ultrapassando a marca de 80% nas primeiras temporadas e aproximando-se de valores próximos a 95% conforme o treinamento avança.

A precisão da validação segue essa tendência, embora em um nível mais baixo, estabilizando-se em torno de 70% a 76% durante o treinamento.

A distância entre as curvas de treinamento e validação é esperada, dada a profundidade da arquitetura e o tamanho moderado do conjunto de dados, mas não assume proporções que comprometam a generalização.

Isso também se reflete na curva de perda: enquanto a perda de treinamento cai rapidamente para valores próximos de zero, a perda de validação atinge uma faixa intermediária e oscila em torno de um platô, sem uma tendência clara de divergência.

Esse comportamento indica um modelo capaz de capturar padrões discriminativos específicos sem colapsar em excesso extremo, o que justifica o fato de que o ConvNeXt-Tiny alcançou a melhor precisão validada entre as três CNNs (Zhang et al., 2024; Steiner, Müller e Köhler, 2024; Andrade e Gomes, 2024).

O ResNet-50 apresentou um perfil de aprendizado diferente, mas igualmente revelador. As curvas de precisão mostram crescimento progressivo tanto no treinamento quanto na validação, com a precisão do treinamento ultrapassando 90% e se aproximando de 97% nas últimas temporadas, enquanto a precisão da validação permanece na faixa de 65% a 70%.

As curvas de perda reforçam essa leitura: a perda de treinamento cai acentuadamente para valores muito baixos, enquanto a perda de validação atinge um nível

mínimo em torno de 1,0 e permanece oscilando sem redução adicional significativa.

Esse descompasso entre desempenho em treinamento e validação caracteriza um leve overfit, no qual o modelo se torna muito especializado em exemplos de treinamento, com ganhos marginais na generalização (Parveen e Dar, 2023; Steiner, Müller e Köhler, 2024; Barrett e Cunningham, 2024).

Ainda assim, o modelo alcança melhor precisão validada por volta da época 43, com um valor em torno de 70,63%, o que o posiciona como uma alternativa robusta, embora inferior ao ConvNeXt-Tiny em termos de equilíbrio entre capacidade de aprendizagem e generalização (Zhang et al., 2024; Steiner, Müller e Köhler, 2024).

O EfficientNetV2-S, por sua vez, apresentou o comportamento mais instável entre as arquiteturas de classificação.

As curvas de precisão revelam um aumento gradual no treinamento, com valores que chegam a cerca de 80% ao final das temporadas, mas a precisão da validação permanece sistematicamente abaixo dos outros modelos, oscilando em torno de 50% a 60%.

A análise das curvas de perda expõe essa dificuldade ainda mais claramente: enquanto a perda de treinamento diminui de forma previsível, a perda de validação apresenta amplas flutuações, com picos e vales sucessivos, sem convergência bem definida.

Essa discrepância indica que o modelo não conseguiu encontrar um ponto estável de generalização para o conjunto de validação, possivelmente devido à combinação da ausência de pesos pré-treinados adequados, à sensibilidade da arquitetura às condições de inicialização e ao tamanho limitado do banco de dados.

Como consequência, o EfficientNetV2-S alcançou a menor precisão final entre as três CNNs, confirmando sua posição de desempenho inferior na comparação (Zhang et al., 2024; Santos e Ribeiro, 2024; Steiner, Müller e Köhler, 2024).

O desempenho do ResNet-50 demonstrou comportamento sólido em classes bem representadas do conjunto, refletido pelo *score f1* de 0,932 em Avulsão e 0,739 em Fratura do Esmalte.

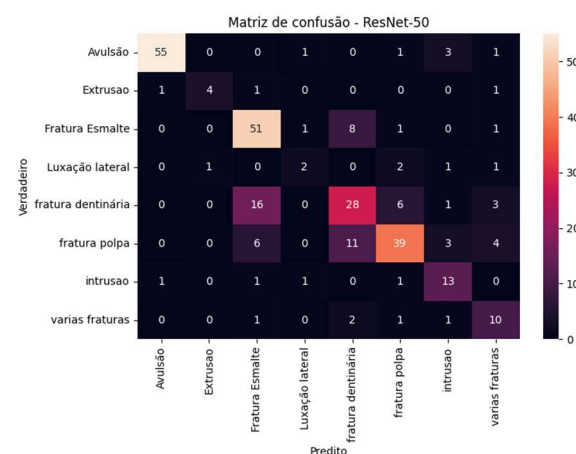
Esses valores mostram que a rede capturou com sucesso padrões radiográficos que apresentam maior contraste estrutural e que tendem a ocorrer de forma marcada na imagem.

No entanto, houve uma diminuição substancial nas categorias de frequência mais baixa, principalmente em extrusão e discordância lateral.

Esse fenômeno está associado ao fato de que o ResNet-50 realiza a classificação sobre a imagem completa, o que expõe o modelo a ruídos anatômicos externos à região de interesse e intensifica o impacto do desequilíbrio do conjunto de dados.

Esse padrão de comportamento sugere que, em cenários radiográficos em que a lesão ocupa uma região espacial limitada e com baixo contraste, o ResNet-50 tende a ser baseado em padrões globais em vez de locais, dificultando o reconhecimento de traumas sutis (Parveen e Dar, 2023; Steiner, Müller e Köhler, 2024; Barrett e Cunningham, 2024).

Figura 11 – Matriz de Confusão – ResNet 50.



Fonte: Elaboração própria (2026).

O EfficientNetV2-S apresentou o desempenho mais instável entre os classificadores, com uma precisão de aproximadamente 59,8%.

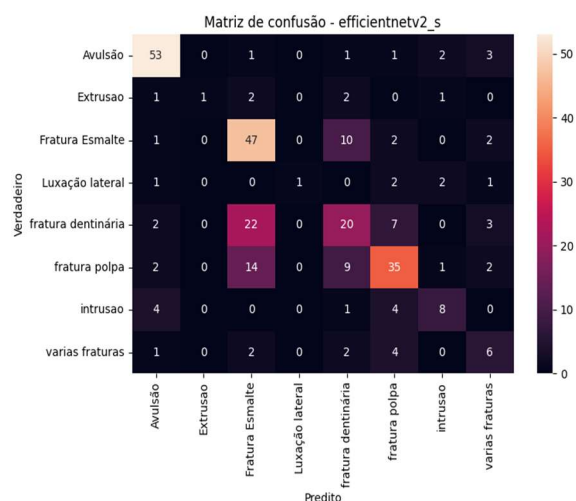
Apesar do excelente desempenho na Avulsão, houve deterioração marcada nas regiões anatômicas de baixa incidência.

A análise dos resultados indica que o EfficientNetV2-S, embora eficiente do ponto de vista computacional, demonstrou vulnerabilidade à sobreposição morfológica entre classes que apresentam similaridades estruturais, e ainda maior sensibilidade ao desequilíbrio da amostra.

O fato de classes como Extrusão e Fratura da Dentina apresentarem valores reduzidos de *recordação* revela que a rede priorizou padrões de distribuição predominantes, em detrimento da diferenciação das classes minoritárias.

Esses achados demonstram que a otimização paramétrica da espinha dorsal não é, isoladamente, uma garantia de desempenho robusto em cenários clínicos com variabilidade visual e baixa homogeneidade entre categorias (Zhang et al., 2024; Santos e Ribeiro, 2024; Steiner, Müller e Köhler, 2024).

Figura 12 – Matriz de Confusão - EfficientNetV2-S.



Fonte: Elaboração própria (2026).

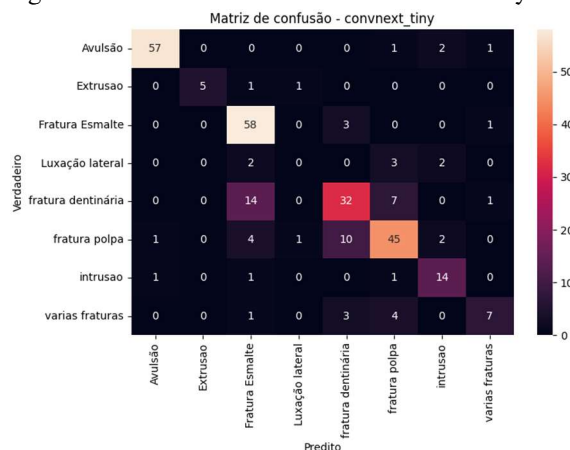
A ConvNeXt-Tiny foi a arquitetura de classificação com melhor desempenho entre as CNNs analisadas, alcançando uma precisão de 76,2% e mostrando estabilidade em praticamente todas as classes com maior densidade amostral.

A rede obteve um *escore f1* de 0,950 em Avulsão, 0,811 em Fratura de Esmalte, 0,726 em Fratura Pulpal e 0,757 em Intrusão, destacando maior capacidade de generalização em cenários clínicos heterogêneos (Zhang et al., 2024; Steiner, Müller e Köhler, 2024; Andrade e Gomes, 2024).

Do ponto de vista comportamental, o ConvNeXt-Tiny era menos suscetível à interferência do conteúdo global da imagem e demonstrava maior competência na extração de padrões discriminativos locais relevantes, o que explica seu desempenho superior em lesões estruturais (Zhang et al., 2024; Steiner, Müller e Köhler, 2024; Andrade e Gomes, 2024).

Mesmo assim, a classe de Luxação lateral manteve desempenho crítico, indicando que o aumento da profundidade arquitetônica e dos mecanismos de atenção espacial não foi suficiente para compensar a falta de exemplos dessa categoria no treinamento (Tanaka et al., 2024; Foster e Del Fiore, 2024).

Figura 13 – Matriz de Confusão – Convnext-tiny.



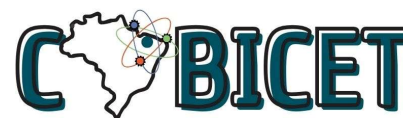
Fonte: Elaboração própria (2026).

CONCLUSÃO

Os resultados obtidos demonstram que arquiteturas convolucionais pré-treinadas representam uma abordagem eficaz para a classificação automática de traumatismos dentários a partir de fotografias clínicas. A utilização de um framework experimental padronizado permitiu comparar, de forma imparcial, o desempenho das arquiteturas ResNet-50, EfficientNetV2-S e ConvNeXt-Tiny, garantindo que as diferenças observadas fossem atribuídas exclusivamente às características de cada modelo.

Entre as arquiteturas avaliadas, a ConvNeXt-Tiny apresentou o melhor equilíbrio entre capacidade de aprendizagem, desempenho de classificação e eficiência computacional, demonstrando maior robustez diante da heterogeneidade visual das imagens e da complexidade inerente às sete categorias de traumatismos dentários investigadas, além da classe composta por casos de várias fraturas. Os resultados reforçam o potencial das arquiteturas convolucionais modernas para o reconhecimento automático de padrões clínicos sutis, contribuindo para aplicações de Inteligência Artificial voltadas à Odontologia.

Por fim, este estudo estabelece uma linha de base comparativa para futuras pesquisas envolvendo classificação automática de traumatismos dentários por meio de Redes Neurais Convolucionais. Além de identificar a arquitetura com melhor desempenho no cenário investigado, o trabalho fornece um protocolo experimental reprodutível que poderá ser utilizado em estudos futuros envolvendo conjuntos de dados ampliados, novas arquiteturas de aprendizado profundo e sistemas computacionais de apoio ao reconhecimento inicial de traumatismos dentários.



AGRADECIMENTOS

Os autores agradecem ao Programa de Pós-Graduação em Engenharia de Controle e Automação (ProPECAut) do Instituto Federal do Espírito Santo (Ifes) e à Universidade Aberta Capixaba (UnAC) pelo apoio institucional e pela oportunidade de desenvolvimento desta pesquisa.

Os autores reconhecem, ainda, o significativo esforço multidisciplinar empregado na construção do conjunto de dados utilizado neste estudo. A etapa de anotação e validação das imagens clínicas, conduzida com o apoio de conhecimentos especializados da área odontológica, demandou um criterioso processo de identificação e classificação. Esse trabalho foi essencial para garantir a qualidade do conjunto de dados e, consequentemente, a confiabilidade do treinamento.

REFERÊNCIAS

ALSHAHRI, H. et al. **Artificial intelligence for dental trauma analysis: a systematic review.** *Diagnostics*, 2023.

BARRETT, R.; CUNNINGHAM, S. **Concerns about bias and fairness in AI-assisted diagnostic dentistry.** *Artificial Intelligence in Healthcare*, v. 3, p. 42–60, 2024.

BOURGUIGNON, C. et al. **International Association of Dental Traumatology guidelines for the management of traumatic dental injuries: 1. Fractures and luxations.** *Dental Traumatology*, v. 36, n. 4, p. 314–330, 2020.

CHEN, Y. et al. **Automated diagnosis of root fractures using deep learning.** *IEEE Access*, 2023.

ÇEGE, A. et al. **Artificial intelligence applications in dental traumatology.** 2025.

DE LUCA, A. **Clinical applicability of deep learning in dental traumatology: a realist review.** *Journal of Dental Research*, v. 103, p. 498–515, 2024.

DENTAGAMA. **Lateral Luxation of Tooth – Prognosis, Management and Treatment.** 2023. Disponível em: <https://dentagama.com/news/lateral-luxation-tooth>. Acesso em: 21 jul. 2026.

EL KHATIB, K. et al. **Management of 3 Avulsed Permanent Teeth: Case Report of a 3-Year Follow-up.** *Case Reports in Dentistry*, 2022.

FOSTER, G. M.; DEL FIORE, A. **Lack of benchmarking in dental trauma AI: an open**

challenge for reproducibility. *Diagnostics*, v. 14, 2024.

FOUAD, A. F. et al. **International Association of Dental Traumatology guidelines for the management of traumatic dental injuries: 3. Avulsion of permanent teeth.** *Dental Traumatology*, v. 36, n. 4, p. 331–342, 2020.

KIM, J. et al. **Deep learning-based proximal caries detection from dental photographs.** *Scientific Reports*, 2023.

KURU, A.; AŞIK, M.; DEMIR, E. **Artificial intelligence in traumatic dental injury diagnosis.** 2025.

LEVIN, L. et al. **International Association of Dental Traumatology guidelines for the management of traumatic dental injuries: general introduction.** *Dental Traumatology*, v. 36, n. 4, p. 309–313, 2020.

PARVEEN, S.; DAR, M. U. **Reproducibility barriers in medical deep learning: a critical methodological perspective.** *Nature Biomedical Engineering*, v. 8, p. 807–823, 2023.

SANABE, M. E.; CAVALCANTE, L. B.; COLDEBELLA, C. R.; ABREU-E-LIMA, F. C. B. **Urgências em traumatismos dentários: classificação, características e procedimentos.** *Revista Paulista de Pediatria*, v. 27, n. 4, p. 447–451, 2009.

SANTOS, E.; RIBEIRO, V. **EfficientNet performance on medical datasets with high anatomical variability.** *Journal of Imaging*, v. 10, p. 1–15, 2024.

SANTINONI, C. et al. **Knowledge of athletes regarding emergency management of traumatic dental injuries.** 2024.

SPINAS, E.; ZANGANI, A.; MALLUS, T.; MAROGNA, G.; CARBONI, L. **Extrusive and Lateral Luxation Injuries in Permanent Teeth: Literature Review and Treatment Complications Update.** *European Journal of Paediatric Dentistry*, v. 25, n. 2, p. 163–167, 2024.

SULE, K.; MALIK, A.; KOTHARI, A.; RAO, V. et al. **Management of Luxation Injuries in Line with International Association of Dental Traumatology Guidelines: Two Case Reports.** *Journal of Conservative Dentistry and Endodontics*, 2024.

STEINER, G.; MÜLLER, R.; KÖHLER, S. **Transformer-based architectures outperform**



CNNs in small-scale clinical imaging scenarios.
Computer Methods and Programs in Biomedicine, v. 242, 2024.

TANAKA, T. et al. **Advances and limitations of AI-based dental trauma detection: a 2024 perspective.**
Journal of Dentistry, 2024.

UDDIN, S. et al. **AI-assisted periodontal assessment using convolutional neural networks.** *Computers in Biology and Medicine*, 2024.

VILLENA, F. et al. **Artificial intelligence and deep learning in dentistry: current applications and future perspectives.** 2024.

ZHANG, X. et al. **Comparative evaluation of CNNs and Vision Transformers for medical image analysis.** *Medical Image Analysis*, 2024.