

Sex Estimation from 3D Analysis of Paranasal Sinuses: A Multicenter Study Using Deep Learning and Machine Learning

Maria Fernanda A. F. Scarcela^{1,2}, Antonio Everton C. Teixeira²,
Diego S. de Mendonça⁵, Saulo A. F. de Oliveira³,
Bruno R. dos Santos Silva², Carlos Caminha^{1,2},
Fábio W. G. Costa⁴, Wellington Franco^{1,2}

¹ Kunumi Lab - Universidade Federal do Ceará (UFC)
Fortaleza – CE – Brasil

²Universidade Federal do Ceará - Campus Crateús

³Instituto Federal do Ceará (IFCE) - Campus Tauá

⁴Universidade Federal do Ceará, Centro de Ciências da Saúde
Departamento de Clínica Odontológica

⁵Centro de Ortodontia Paulo Picanço

fernanda08012004@gmail.com, caminha@ufc.br, wellington@crateus.ufc.br

Abstract. *The paranasal sinuses are used in forensic sex identification, typically through manual measurements on computed tomography (CT) images. This study proposes to automate this classification using 3D convolutional neural networks (ResNet-50, DenseNet-201) integrated with supervised algorithms (SVM, Random Forest, MLP, KNN). Eight combinations were tested, with ResNet-50 + RF achieving the best performance, with an accuracy of 92.60%, precision of 91.18%, recall of 94.55%, and F1-score of 92.64%. The results indicate high potential for the combined use of deep learning and supervised learning to determine sex from segmented masks of the paranasal sinuses, with relevant application in the forensic field.*

Resumo. *Os seios paranasais são usados na identificação forense de sexo, geralmente via medidas manuais em imagens de tomografia (CT). Esse estudo propõe automatizar essa classificação por meio de redes neurais convolucionais 3D (ResNet-50, DenseNet-201) integradas a algoritmos supervisionados (SVM, Random Forest, MLP, KNN). Foram testadas oito combinações, das quais a ResNet-50 + RF obteve melhor desempenho com acurácia de 92,60%, precisão de 91,18%, recall de 94,55% e F1-score de 92,64%. Os resultados indicam alto potencial do uso conjunto de deep learning e aprendizado supervisionado para determinar o sexo a partir de máscaras segmentadas dos seios paranasais, com aplicação relevante na área forense.*

1. Introdução

A determinação de sexo em contextos forenses é uma etapa crucial, não apenas por seu valor científico, mas também por sua relevância social [Nagare et al. 2018]. Nesse sentido, a

identificação de indivíduos é especialmente crítica em situações como desastres naturais, guerras e explosões, em que os corpos são encontrados com a estrutura bastante comprometida, em avançado estado de deterioração ou decomposição [Jasim and Al-Taei 2013]. Diante disso, o campo da análise forense dispõe de diversas técnicas para a identificação post mortem, muitas das quais envolvem a extração manual de métricas de regiões anatômicas associadas ao dimorfismo sexual, caracterizado por diferenças estruturais entre os sexos biológicos [Spradley 2016, Cohen et al. 2018].

Na literatura, diversos estudos analisam características estruturais ósseas, como a mandíbula [de Oliveira Gamba et al. 2016, de Oliveira Gamba et al. 2017], fêmur [Cavaignac et al. 2016, Moosa et al. 2021], dentes [Martins Filho et al. 2016, Zorba et al. 2011] e o forame magno [Jaitley et al. 2016, Pereira et al. 2021], com o intuito de extrair informações relevantes, como o sexo e a idade dos indivíduos. Estudos envolvendo medidas do crânio se destacam por sua resistência à deterioração e pela presença de diversos parâmetros associados ao dimorfismo sexual [Sidhu et al. 2014]. Entre as regiões cranianas de interesse, destacam-se os seios paranasais, estruturas com alto potencial discriminatório [de Barros et al. 2021, Cox et al. 2009].

Os seios paranasais são cavidades localizadas nos ossos ao redor das fossas nasais, incluindo o seio maxilar direito (SMD), seio maxilar esquerdo (SME), seio etmoidal (SE) e seio frontal (SF). Essas estruturas apresentam variações consideráveis entre indivíduos, como forma, dimensão e volume. Estudos como [de Mendonça et al. 2021] indicam que essas características podem ser úteis na estimativa de sexo em contextos forenses. Na radiologia oral e maxilofacial, essas medidas são geralmente obtidas por meio de exames de imagem, como Tomografia Computadorizada (TC), que permitem uma análise tridimensional detalhada das estruturas ósseas [Teke et al. 2007]. No entanto, a extração manual dessas medidas por especialistas é um processo dispendioso, sujeito a erros humanos e dependente da experiência do avaliador [da Silva et al. 2024].

Nos últimos anos, o uso de técnicas de Inteligência Artificial (IA) tem se expandido significativamente em diversas áreas, incluindo a medicina forense [Piraianu et al. 2023]. Entre os principais recursos da IA estão algoritmos capazes de reconhecer padrões complexos nos dados, possibilitando aplicações como classificação, regressão e segmentação [Liu et al. 2021]. No caso de imagens volumétricas, as CNNs tridimensionais (3D CNNs) representam uma abordagem promissora, pois preservam as informações espaciais, superando limitações das redes 2D [Tiwari et al. 2023]. Apesar dos desafios relacionados ao alto custo computacional e à necessidade de grandes volumes de dados anotados, as 3D CNNs têm demonstrado grande potencial na análise forense, contribuindo para tarefas como estimativa de sexo e análise anatômica detalhada [Thurzo et al. 2021, Cao et al. 2021].

Com o propósito de explorar essas técnicas na medicina forense, este trabalho propõe o uso de 3D CNNs para extrair automaticamente características de máscaras segmentadas dos seios paranasais, obtidas a partir de imagens de TCs. O objetivo central é estimar o sexo dos indivíduos com base nessas informações anatômicas. As representações extraídas serão utilizadas como entrada em algoritmos de aprendizado supervisionado, como *Support Vector Machine* (SVM), *Random Forest* (RF), *K-Nearest Neighbors* (KNN) e *Multilayer Perceptron* (MLP), com o intuito de realizar a classificação binária entre os sexos masculino e feminino.

A estrutura deste artigo está organizada da seguinte forma: a Seção 2 discute trabalhos relacionados, a Seção 3 descreve a abordagem proposta, a Seção 4 apresenta os resultados obtidos e a Seção 5, as principais limitações do trabalho. Por fim, as conclusões são apresentadas na Seção 6.

2. Trabalhos relacionados

Atualmente, técnicas de IA têm sido amplamente aplicadas à análise de imagens médicas, como Tomografia Computadorizada de Feixe Cônico (CBCT) e TCs, para tarefas como detecção, segmentação e classificação de estruturas anatômicas. Na área forense, abordagens baseadas em redes neurais têm demonstrado eficácia na identificação de dimorfismo sexual por meio da análise de estruturas craniofaciais [da Silva et al. 2024, Bewes et al. 2019].

Em [da Silva et al. 2024] é proposta uma rede neural com atenção guiada por anatomia em duas etapas (SDetNet) para a segmentação do seio frontal e classificação automática do sexo a partir de 310 imagens CBCTs obtidas na Coreia do Sul. O modelo utiliza uma entrada multicanal e incorpora um módulo de atenção anatômica, atingindo 97,9% na etapa de classificação. Embora a acurácia obtida tenha sido inferior em nosso trabalho, os resultados mostraram-se comparáveis, superando os 90% de acurácia. Ademais, o presente estudo buscou ampliar a variabilidade e representatividade da amostra ao incluir maior diversidade populacional, aspecto essencial na análise de seios paranasais, uma vez que estão sujeitos a influências regionais, étnicas e populacionais, as quais podem afetar a acurácia preditiva [Zheng et al. 2023].

[Hamidi et al. 2024] desenvolveu um modelo híbrido de aprendizado profundo composto por uma rede neural profunda (DNN) combinada com um algoritmo genérico (GA) para estimativa de idade e sexo a partir de doze características dos seios frontais e maxilares. Tais características, como volume e largura, foram obtidas a partir de 240 imagens CBCT por meio de um software. O GADNN demonstrou consistentemente maior acurácia em comparação a métodos como *Logistic Regression* (LR), *Random Forest* (RF) e *Multilayer Perceptron* (MLP), alcançando 86% de acurácia na tarefa de determinação de sexo. Por outro lado, a nossa abordagem se destacou ao alcançar 92,60% de acurácia, apesar de utilizar um conjunto de dados menor. Além disso, nosso modelo opera diretamente sobre as imagens sem depender da extração de características pré-definidas, embora exija previamente a segmentação manual das máscaras nas CTs.

O estudo de [Kondou et al. 2023] utilizou um modelo de atenção com múltiplas instâncias (MIL) baseado em DenseNet121 para determinar o sexo a partir de imagens 3D do crânio, alcançando 93% de acurácia em cadáveres do Leste Asiático. Diferentemente, este trabalho concentra-se exclusivamente nos seios paranasais, uma região mais sutil e morfologicamente variável, o que torna a tarefa mais complexa, no entanto atingindo resultados comparáveis. Esse foco permite a redução de dimensionalidade da imagem tomográfica, o que demanda menor consumo de memória da *Graphic Processing Unit* (GPU) e torna o processamento mais eficiente, com menor perda de informação.

O estudo de [Kim et al. 2021] utilizou redes convolucionais profundas (ResNet-152 e DenseNet-169), pré-treinadas no ImageNet, para prever sexo e idade a partir de radiografias dos seios paranasais, obtendo acurácia de até 98% para sexo e 77,6% para grupos etários, em um extenso conjunto de 4.160 imagens. Além disso, empregaram ma-

pas de ativação (CAM) para identificar regiões anatômicas determinantes, como os seios maxilares e etmoidais. Apesar do excelente desempenho, o uso exclusivo de imagens 2D limita a profundidade anatômica capturada. Em contraste, o presente trabalho utiliza dados volumétricos (CTs) segmentados manualmente e explora o potencial de arquiteturas modernas 3D (ResNet50 e DenseNet201), com objetivo de analisar de forma mais contextual e detalhada as características morfológicas dos seios paranasais. Embora o conjunto de dados seja menor (108 exames), a metodologia adotada busca mitigar essa limitação com técnicas robustas de aumento de dados, validação cruzada estratificada e um pipeline modular que separa a extração de embeddings da classificação supervisionada.

No trabalho conduzido por [Senol et al. 2023], o objetivo foi estimar sexo e idade por meio de algoritmos de aprendizado de máquina aplicados a parâmetros métricos extraídos dos dentes caninos e molares superiores obtidos via CBCTs. Utilizando 240 imagens (120 de cada sexo), foram realizadas 14 medições específicas em estruturas dentárias, como ângulos e distâncias entre dentes e pontos anatômicos de referência. Os autores empregaram diversos algoritmos tradicionais de aprendizado supervisionado, como *AdaBoost*, *Random Forest* e *Logistic Regression*, alcançando acurácia de até 81% na classificação de sexo. Apesar dos bons resultados, o estudo dependeu de múltiplas medições manuais, o que pode introduzir variabilidade interobservador e limita sua escalabilidade. Em contraste, o presente trabalho propõe uma abordagem baseada em aprendizado profundo com 3D CNNs, aplicadas diretamente aos volumes completos da região dos seios paranasais, sem a necessidade de extração manual de atributos.

3. Metodologia

Esta seção descreve a metodologia adotada para a determinação de sexo a partir de máscaras de segmentação dos seios paranasais extraídas de tomografias computadorizadas (CTs). A abordagem adotada inclui o uso de 3D CNNs para a extração de representações (*embeddings*), que são posteriormente utilizadas por classificadores supervisionados para concluir a predição. A Figura 1 demonstra o *pipeline* da proposta.

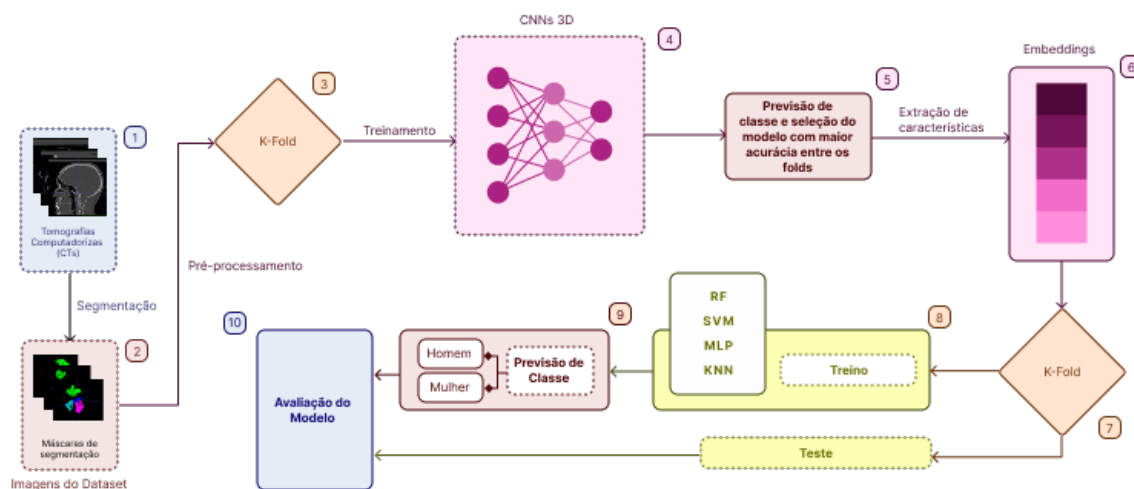


Figura 1. Fluxograma da proposta.

Inicialmente ocorre a aquisição dos dados, em que as tomografias computadorizadas da região craniana foram coletadas e posteriormente segmentadas por um profissional especialista. Posteriormente, CNNs foram utilizadas para a extração de características dos exames, visando classificação entre sexos masculino e feminino. Tais características são extraídas para a aplicação de algoritmos de aprendizado de máquina tradicional, tais como *RF*, *SVM*, *MLP* e *KNN*. Por fim, na etapa de treinamento/testes, a técnica de validação cruzada *K-Fold* foi utilizada, de forma a garantir a isonomia dos algoritmos, evitando assim a contaminação do conjunto de testes na etapa de treinamento.

3.1. Aquisição dos dados

O conjunto de dados utilizado neste estudo é composto por 108 tomografias computadorizadas da região craniana de indivíduos com idades entre 18 e 49 anos, sendo 54 do sexo masculino e 54 do sexo feminino. As imagens foram coletadas nos estados do Pará, Ceará e Bahia, provenientes de indivíduos naturais de cada respectivo estado. A distribuição das classes por estado está apresentada na Tabela 1.

Foram utilizados dois modelos de tomógrafos: *Somatom Emotion 6* (Siemens, Alemanha) e *Light Speed VCT* (GE Healthcare, EUA). As aquisições seguiram os seguintes parâmetros: incremento de mesa de 1 mm, 130 kVp, corrente entre 100–200 mA, espessura de até 2,0 mm, campo de visão (FOV) de 180 mm e tempo de rotação de 0,6 s. Como critérios de exclusão, foram consideradas imagens de crânios que apresentavam fraturas, patologias ou alterações craniofaciais, visando garantir a homogeneidade da amostra e minimizar possíveis vieses.

Tabela 1. Distribuição das imagens por estado

Estado	Homens	Mulheres	Total
Pará	34	34	68
Ceará	8	8	16
Bahia	12	12	24
Total	54	54	108

3.2. Pré-processamento

Após a coleta dos exames, as máscaras dos seios paranasais foram segmentadas por um cirurgião-dentista especialista utilizando os softwares ITK-SNAP (v3.8.0) e 3D Slicer (v4.10.2), como apresentadas na Figura 2. Os arquivos foram salvos no formato *.nii*, com apenas um canal de cor e dimensão padrão de $512 \times 512 \times S$ *voxels*, em que *S* é o número de cortes (*slices*), o qual varia entre os exames.

Cada seio paranasal é representado por um valor distinto nos voxels da máscara. Para garantir a coerência entre as amostras, realizou-se uma padronização manual dos rótulos, atribuindo os valores 0,0 para o fundo, 2,0 para o seio frontal, 3,0 para o esfenoidal, 4,0 para o maxilar direito e 5,0 para o maxilar esquerdo. Além da abordagem multivalorada para a rotulagem realizada, também foi considerada uma versão binarizada das máscaras, com o valor 1,0 para os seios marcados e 0,0 para o fundo. Ambas as abordagens foram avaliadas, com o objetivo de investigar se a diferenciação anatômica entre os seios impacta no desempenho da rede, em comparação à segmentação como uma única estrutura.

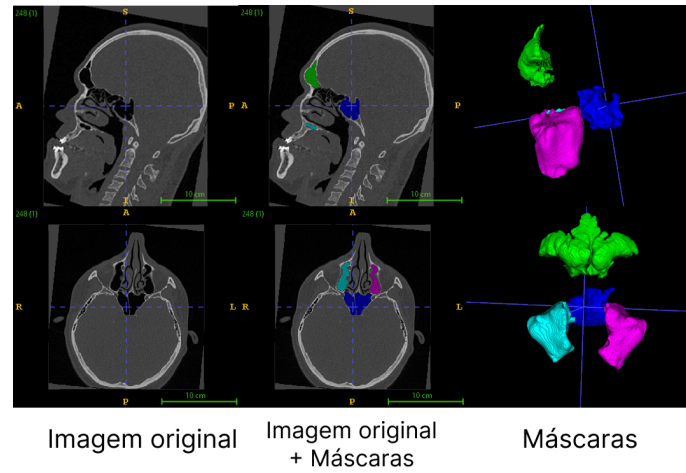


Figura 2. Exemplo de imagens CTs e suas respectivas máscaras de segmentação. Imagem baseada na ilustrada no artigo [da Silva et al. 2024]

O pré-processamento foi realizado com auxílio das bibliotecas PyTorch (v2.6.0)¹ e MONAI (v1.4.0)². Devido ao tamanho variável e à alta resolução das imagens, os volumes foram recortados nas três dimensões (coronal, axial e sagital) para isolar os seios paranasais. Em seguida, aplicou-se o preenchimento (*SpatialPad*) para uniformizar as dimensões e, por fim, os volumes foram redimensionados com o método interpolação (modo *nearest*) para 256×256×128 voxels. A Figura 3 apresenta essa etapa de pré-processamento. Tal abordagem permite reduzir o uso de memória, enquanto minimiza a perda de informação oriunda da segmentação tridimensional.

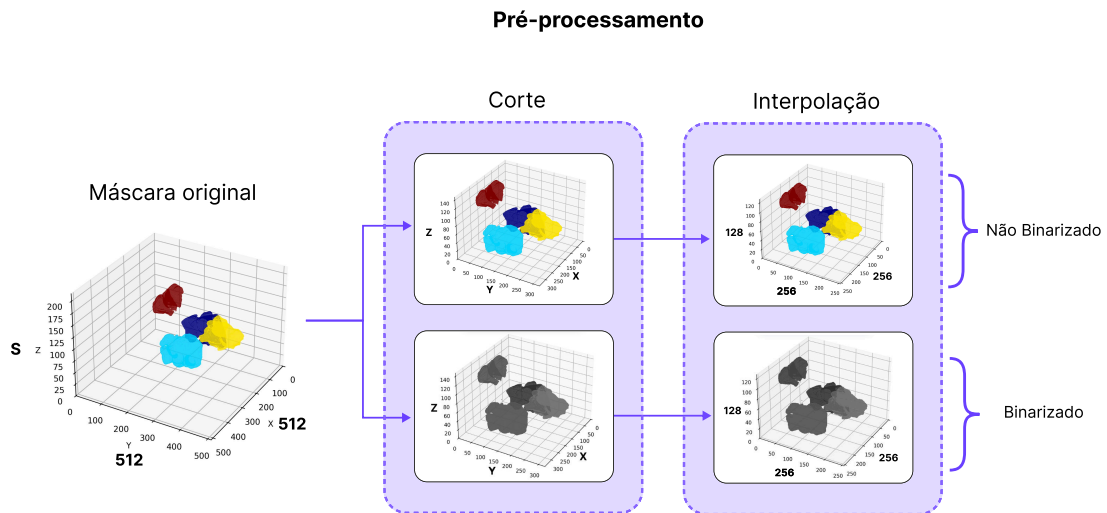


Figura 3. Ilustração do pré-processamento realizado nas máscaras de segmentação.

¹<https://pytorch.org/>

²<https://monai.io/>

3.3. Extração de características

Os *embeddings* foram extraídos a partir de 3D CNNs disponibilizadas pela biblioteca *timm3d* (*PyTorch Volumes Models for 3D data*, v1.0.1)³, uma extensão da biblioteca *timm* (*PyTorch Image Models*). A escolha dessa biblioteca fundamentou-se na facilidade de integração com o PyTorch e na disponibilidade de versões 3D.

As arquiteturas exploradas neste estudo incluíram tridimensionais da ResNet-50 e da DenseNet201. A ResNet-50 pertence à família das *Residual Networks*, caracterizada pelo uso de conexões residuais que facilitam o treinamento de redes profundas, atenuando problemas relacionados ao desaparecimento do gradiente. Já a DenseNet-201 adota uma estrutura densa de conexões entre camadas, visando à reutilização das características extraídas e contribuindo para maior desempenho com menor risco de *overfitting*.

Inicialmente, os experimentos foram conduzidos com as camadas convolucionais congeladas. No entanto, os resultados obtidos nessa configuração não foram satisfatórios. Logo, optou-se pelo treinamento completo das arquiteturas, mantendo a estrutura base das CNNs e modificando a camada totalmente conectada final por uma camada com um único neurônio e ativação sigmoide. Como função de perda, foi utilizada a *Binary Cross Entropy* (BCELoss), e o otimizador adotado foi o Adam. A métrica principal de avaliação foi a acurácia binária (*binary accuracy*), a qual representa uma avaliação média do desempenho dos algoritmos na identificação do gênero.

O treinamento foi realizado com validação cruzada estratificada em 5 *folds* (*Stratified K-Fold*), técnica que divide o conjunto de dados em cinco partes, mantendo o balanceamento entre as classes. Em cada iteração, quatro *folds* foram utilizados para treinamento e um para teste. A porção destinada ao treino foi novamente subdividida por meio da função *train_test_split* da biblioteca Scikit⁴, reservando-se 10% dos dados para validação. Essa divisão permitiu a aplicação da técnica de *early stopping*, com paciência de 10 épocas e limiar mínimo de melhoria de 0,1 na perda de validação, com o objetivo de prevenir *overfitting* e preservar os melhores pesos ao longo do treinamento. O treinamento foi conduzido com *batch size* igual a 2 devido às limitações da GPU. Dessa forma, totalizaram-se 77 amostras para treinamento, 9 para validação e 22 para teste por *fold*. Durante o treinamento, foram aplicadas técnicas de *data augmentation* com o objetivo de aumentar a variabilidade das amostras e promover maior robustez do modelo ao lidar com variações de posicionamento, orientação e escala dos volumes. As transformações incluíram rotações aleatórias de até $\pm 10^\circ$, espelhamento horizontal, zoom aleatório entre 90% e 110% e transformações afins com leve cisalhamento.

3.4. Classificação

Nesta etapa, foram utilizados apenas os *embeddings* extraídos pela CNN que obteve o melhor desempenho entre os modelos treinados durante o *K-Fold*. Esses mapas foram inicialmente salvos em formato de tensor e, posteriormente, convertidos para numpy arrays, a fim de possibilitar sua integração com os classificadores da biblioteca Scikit-learn (versão 1.6.1). Os algoritmos de aprendizado supervisionado aplicados para distinguir entre amostras do sexo masculino e feminino, com base nos vetores de características

³<https://github.com/ZFTurbo/timm3d>

⁴<https://scikit-learn.org/stable/>

extraídos, foram: K-Nearest Neighbors (KNN), Multi-Layer Perceptron (MLP), Support Vector Machine (SVM) e Random Forest (RF).

Na avaliação dos classificadores, foram utilizados os mesmos índices de validação cruzada estratificada (*K-Fold*) da etapa anterior, assegurando que não houvesse sobreposição entre os conjuntos de treino e teste, mantendo assim uma avaliação isonômica entre os conjuntos amostrados. Nessa fase, não foi realizada uma divisão adicional para validação, portanto as partições corresponderam a **86** amostras para treino e **22** para teste em cada *fold*. Quanto à configuração dos hiperparâmetros, o KNN foi ajustado com *5 neighbors*; a SVM utilizou *kernel* do tipo *Radial Basis Function* (RBF); o MLP foi treinado por até 300 *épocas*, com semente aleatória fixada para garantir reprodutibilidade; e o RF empregou 100 árvores de decisão. Os demais parâmetros foram mantidos com seus valores padrão definidos pela biblioteca Scikit-learn.

Para avaliar o desempenho dos modelos, foram utilizadas as seguintes métricas, a saber, Acurácia (ACC), Precisão (*Precision*), *Recall*, *F1-score* (F1). Para problemas de classificação binária com classes balanceadas, essas métricas fornecem uma visão abrangente do desempenho dos modelos, especialmente em termos de sua capacidade de generalização.

4. Resultados e Discussão

O presente trabalho propõe o uso de duas redes neurais convolucionais 3D (ResNet-50 e DenseNet-201) para extração de características a partir de máscaras dos seios paranasais obtidas a partir de 108 tomografias computadorizadas. Essas características foram utilizadas como entrada para quatro classificadores supervisionados tradicionais: KNN, SVM, MLP e RF, totalizando 8 (oito) combinações de modelos distintas.

Tabela 2. Resultados de acurácia, precisão, recall e f1-score (%) de classificadores supervisionados aplicados a embeddings extraídos por CNNs 3D (ResNet-50 e DenseNet-201) com máscaras de segmentação binarizadas.

Modelo	Acc (%)	Precision (%)	Recall (%)	F1-Score (%)
Resnet50 only	67.79 ± 11.49	66.27 ± 10.80	76.36 ± 13.61	70.32 ± 9.77
Resnet50 + KNN	90.69 ± 6.65	91.18 ± 9.15	90.73 ± 5.76	90.74 ± 6.49
Resnet50 + MLP	92.55 ± 4.71	94.18 ± 4.76	90.73 ± 5.76	92.37 ± 4.87
Resnet50 + RF	92.60 ± 4.65	91.18 ± 5.31	94.55 ± 7.27	92.64 ± 4.84
Resnet50 + SVM	92.55 ± 5.59	94.36 ± 7.37	90.73 ± 5.76	92.38 ± 5.71
Densenet201 only	54,63 ± 16,72	55,31 ± 16,23	64,82 ± 14,58	59,22 ± 14,60
Densenet201 + KNN	75,00 ± 6,00	74,23 ± 4,20	75,93 ± 10,48	74,91 ± 7,27
Densenet201 + MLP	70,37 ± 5,71	71,01 ± 4,51	68,52 ± 10,48	69,49 ± 6,87
Densenet201 + RF	72,22 ± 6,80	72,88 ± 6,00	70,37 ± 9,44	71,51 ± 7,47
Densenet201 + SVM	71,30 ± 4,72	74,18 ± 2,66	64,81 ± 9,44	68,98 ± 6,45

Tabela 3. Resultados de acurácia, precisão, recall e f1-score (%) de classificadores supervisionados aplicados a embeddings extraídos por CNNs 3D (ResNet-50 e DenseNet-201) com máscaras de segmentação não binarizadas.

Modelo	Acc (%)	Precision (%)	Recall (%)	F1-Score (%)
Resnet50 only	62.03 ± 11.76	62.32 ± 12.03	59.45 ± 19.26	59.97 ± 14.76
Resnet50 + KNN	89.87 ± 4.39	89.62 ± 6.29	90.73 ± 8.14	89.83 ± 4.51
Resnet50 + MLP	89.00 ± 8.89	89.40 ± 9.01	89.09 ± 14.55	88.67 ± 9.97
Resnet50 + RF	87.19 ± 10.49	85.85 ± 9.71	89.09 ± 14.55	87.14 ± 11.20
Resnet50 + SVM	90.87 ± 7.58	90.26 ± 9.55	92.73 ± 8.91	91.12 ± 7.30
Densenet201 only	52.73 ± 7.08	53.48 ± 9.43	44.18 ± 28.11	44.94 ± 13.97
Densenet201 + KNN	42.64 ± 5.76	42.88 ± 7.06	40.73 ± 4.08	41.59 ± 4.82
Densenet201 + MLP	47.32 ± 8.52	36.51 ± 19.80	49.82 ± 33.08	41.56 ± 24.45
Densenet201 + RF	49.00 ± 5.57	49.13 ± 5.43	46.55 ± 12.52	47.02 ± 7.72
Densenet201 + SVM	48.05 ± 4.55	44.24 ± 9.06	24.00 ± 16.76	28.90 ± 14.90

As métricas apresentadas nas Tabelas 2 e 3 demonstram desempenho competitivo entre as abordagens utilizando máscaras binarizadas e não binarizadas, quando empregada a arquitetura ResNet50 como extrator de características. No cenário com máscaras binarizadas, a configuração **ResNet50 + RF** apresentou o melhor desempenho geral, alcançando acurácia de **92,60%**, recall de **94,55%**, F1-score de **92,64%** e precisão de **91,18%**. No caso das máscaras não binarizadas, destaca-se a combinação **Resnet50 + SVM** que atingiu acurácia de **90,87%**, recall de **92,73%**, F1-score de **91,12%** e precisão de **90,26%**. Esses resultados não apenas refletem um alto poder preditivo, mas também revelam consistência entre as partições de validação. O alto F1-score indica que os modelos conseguiram manter o equilíbrio entre sensibilidade e precisão, aspecto favorecido pelo uso de um conjunto de dados balanceado, o que possibilitou diferenciar bem as classes mesmo diante da heterogeneidade morfológica. Ao comparar segmentações anatômicas (máscaras não binarizadas) com volumes binarizados, observou-se que a versão binarizada apresentou desempenho superior, sugerindo que a simplificação estrutural pode favorecer a robustez e a generalização do modelo. Apesar de os modelos treinados com máscaras não binarizadas apresentarem acurácia competitiva em alguns casos, eles demonstraram maior desequilíbrio entre classes, como evidenciado por F1-scores mais instáveis (maior desvio padrão). Em contraste, os modelos que utilizaram segmentações anatômicas rotuladas apresentaram representações mais discriminativas e desempenho mais equilibrado entre as classes. Redes mais profundas, como a ResNet-50, demonstraram maior capacidade para capturar traços morfológicos complexos, mas necessitaram do uso de classificadores externos para uma generalização eficaz, pois não obtiveram resultados satisfatórios isoladamente. Em contrapartida, a DenseNet-201 teve desempenho muito inferior em ambos os cenários, especialmente com máscaras não binarizadas, provavelmente devido à incompatibilidade arquitetural ou limitações no treinamento com dados volumétricos.

5. Ameaças à validação

Podemos citar como ameaças a validade, que todas as segmentações foram realizadas manualmente por um único especialista, o que pode introduzir viés do observador e inconsistências na definição anatômica. Além disso, as limitações computacionais também

impactaram o processo de treinamento. Devido à alta demanda de memória dos dados volumétricos, foi necessário utilizar tamanhos de lote reduzidos (*batch size* pequeno), o que é conhecido por gerar estimativas de gradiente mais ruidosas e reduzir a eficácia da normalização em lote (*batch normalization*). Essa limitação afeta especialmente os modelos mais profundos e pode ter influenciado a estabilidade do treinamento.

Por fim, embora o conjunto de dados inclua indivíduos de três estados brasileiros, ele não representa adequadamente a diversidade étnico-racial do país. A população brasileira é altamente miscigenada, o que resulta em grande variabilidade anatômica, especialmente em estruturas craniofaciais como os seios paranasais. Para que os modelos possam ser aplicados em contextos forenses de forma mais ampla, é fundamental garantir uma representatividade demográfica mais abrangente.

6. Conclusão

Nosso trabalho traz a combinação de 3D CNNs com classificadores tradicionais para prever o sexo com base na morfologia dos seios paranasais. O melhor desempenho foi encontrado em ResNet-50 + RF com a abordagem binarizada, que se destacou pela alta acurácia, estabilidade e capacidade de lidar com o equilíbrio entre as classes — um aspecto crítico em conjuntos de dados forenses reais. Os resultados indicam que a simplificação anatômica via segmentações binarizadas gerou desempenho superior, mostrando que detalhes rotulados nem sempre contribuem para melhor predição.

Por fim, a combinação de 3D CNNs com classificadores externos representa uma abordagem viável e eficaz para a identificação forense de sexo, especialmente ao lidar com imagens médicas volumétricas de populações anatomicamente diversas, como a brasileira. Desta forma, os modelos desenvolvidos podem ser empregados como assistentes ao especialista em cenários forenses, auxiliando na tomada de decisão. Além de reduzir o tempo de inferência e melhorar a precisão na classificação de sexo, o sistema também é robusto a variações na resolução e qualidade das imagens, reforçando seu potencial para aplicações práticas. Como trabalhos futuros podemos citar a ampliação e diversificação geográfica do conjunto de dados, por meio de parcerias com novas instituições. Além disso, é fundamental automatizar o processo de segmentação e aprimorar as arquiteturas para construir sistemas mais robustos e generalizáveis.

Referências

- Bewes, J., Low, A., Morphett, A., Pate, F. D., and Henneberg, M. (2019). Artificial intelligence for sex determination of skeletal remains: Application of a deep learning artificial neural network to human skulls. *Journal of forensic and legal medicine*, 62:40–43.
- Cao, Y., Ma, Y., Vieira, D. N., Guo, Y., Wang, Y., Deng, K., Chen, Y., Zhang, J., Qin, Z., Chen, F., et al. (2021). A potential method for sex estimation of human skeletons using deep learning and three-dimensional surface scanning. *International Journal of Legal Medicine*, 135(6):2409–2421.
- Cavaignac, E., Savall, F., Faruch, M., Reina, N., Chiron, P., and Telmon, N. (2016). Geometric morphometric analysis reveals sexual dimorphism in the distal femur. *Forensic science international*, 259:246–e1.
- Cohen, O., Warman, M., Fried, M., Shoffel-Havakuk, H., Adi, M., Halperin, D., and Lahav, Y. (2018). Volumetric analysis of the maxillary, sphenoid and frontal sinuses:

- A comparative computerized tomography based study. *Auris Nasus Larynx*, 45(1):96–102.
- Cox, M., Malcolm, M., and Fairgrieve, S. I. (2009). A new digital method for the objective comparison of frontal sinuses for identification. *Journal of forensic sciences*, 54(4):761–772.
- da Silva, R. L. B., Yang, S., Kim, D., Kim, J. H., Lim, S.-H., Han, J., Kim, J.-M., Kim, J.-E., Huh, K.-H., Lee, S.-S., et al. (2024). Automatic segmentation and classification of frontal sinuses for sex determination from cbct scans using a two-stage anatomy-guided attention network. *Scientific Reports*, 14(1):11750.
- de Barros, F., da Silva Fernandes, C. M., Kuhnen, B., Scarso Filho, J., Gonçalves, M., and da Costa Serra, M. (2021). Paranasal sinuses and human identification. *Research, Society and Development*, 10(9):e48710918161–e48710918161.
- de Mendonça, D. S., Kurita, L. M., Carvalho, F. S. R., Tuji, F. M., Silva, P. G. d. B., Bezerra, T. P., de Aguiar, A. S. W., and Gurgel Costa, F. W. (2021). Development and validation of a new formula for sex estimation based on multislice computed tomographic measurements of maxillary and frontal sinuses among brazilian adults. *Dentomaxillofacial Radiology*, 50(6):20200490.
- de Oliveira Gamba, T., Alves, M. C., and Haiter-Neto, F. (2016). Mandibular sexual dimorphism analysis in cbct scans. *Journal of forensic and legal medicine*, 38:106–110.
- de Oliveira Gamba, T., Yamasaki, M. C., Groppo, F. C., da Silveira, H. L. D., de Almeida Boscolo, S. M., Sanderink, G. C., and Berkhout, W. E. R. (2017). Validation study of a new method for sexual prediction based on cbct analysis of maxillary sinus and mandibular canal. *Archives of oral biology*, 83:118–123.
- Hamidi, O., Afrasiabi, M., and Namaki, M. (2024). 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*, 24(1):50.
- Jaitley, M., Phulambrikar, T., Kode, M., Gupta, A., and Singh, S. K. (2016). Foramen magnum as a tool for sexual dimorphism: A cone beam computed tomography study. *Indian Journal of Dental Research*, 27(5):458–462.
- Jasim, H. H. and Al-Taei, J. A. (2013). Computed tomographic measurement of maxillary sinus volume and dimension in correlation to the age and gender (comparative study among individuals with dentate and edentulous maxilla). *Journal of Baghdad College of Dentistry*. Accessed: 2025-06-30.
- Kim, D.-K., Cho, B.-J., Lee, M.-J., and Kim, J. H. (2021). Prediction of age and sex from paranasal sinus images using a deep learning network. *Medicine*, 100(7):e24756.
- Kondou, H., Morohashi, R., Kimura, S., Idota, N., Matsunari, R., Ichioka, H., Bandou, R., Kawamoto, M., Ting, D., and Ikegaya, H. (2023). Artificial intelligence-based forensic sex determination of east asian cadavers from skull morphology. *Scientific Reports*, 13(1):21026.

- Liu, X., Song, L., Liu, S., and Zhang, Y. (2021). A review of deep-learning-based medical image segmentation methods. *Sustainability*, 13(3):1224.
- Martins Filho, I. E., Lopez-Capp, T. T., Biazevic, M. G. H., and Michel-Crosato, E. (2016). Sexual dimorphism using odontometric indexes: Analysis of three statistical techniques. *Journal of Forensic and Legal Medicine*, 44:37–42.
- Moosa, S. S., Shaikh, M. H. R., Khwaja, M., Shaikh, S. A. H., Siddiqui, F. B., Daimi, S. R. H., Hiware, S. D., Ismail, E. E., and Begum, Y. (2021). Sexual dimorphic parameters of femur: a clinical guide in orthopedics and forensic studies. *Journal of Medicine and Life*, 14(6):762.
- Nagare, S. P., Chaudhari, R. S., Birangane, R. S., and Parkarwar, P. C. (2018). Sex determination in forensic identification, a review. *Journal of forensic dental sciences*, 10(2):61–66.
- Pereira, J. G. D., Santos, J. B. S., Sousa, S. P. d., Franco, A., and Silva, R. H. A. (2021). Frontal sinuses as tools for human identification: a systematic review of imaging methods. *Dentomaxillofacial Radiology*, 50(5):20200599.
- Piraiianu, A.-I., Fulga, A., Musat, C. L., Ciobotaru, O.-R., Poalelungi, D. G., Stamate, E., Ciobotaru, O., and Fulga, I. (2023). Enhancing the evidence with algorithms: how artificial intelligence is transforming forensic medicine. *Diagnostics*, 13(18):2992.
- Senol, D., Secgin, Y., Duman, B. S., Toy, S., and Oner, Z. (2023). Sex and age estimation with machine learning algorithms with parameters obtained from cone beam computed tomography images of maxillary first molar and canine teeth. *Egyptian Journal of Forensic Sciences*, 13(1):27.
- Sidhu, R., Chandra, S., Devi, P., Taneja, N., Sah, K., and Kaur, N. (2014). Forensic importance of maxillary sinus in gender determination: A morphometric analysis from western uttar pradesh, india. *European Journal of General Dentistry*, 3(01):53–56.
- Spradley, M. K. (2016). Metric methods for the biological profile in forensic anthropology: sex, ancestry, and stature. *Academic forensic pathology*, 6(3):391–399.
- Teke, H. Y., Duran, S., Canturk, N., and Canturk, G. (2007). Determination of gender by measuring the size of the maxillary sinuses in computerized tomography scans. *Surgical and radiologic anatomy*, 29:9–13.
- Thurzo, A., Kosnáčová, H. S., Kurilová, V., Kosmel', S., Beňuš, R., Moravanský, N., Kováč, P., Kuracinová, K. M., Palkovič, M., and Varga, I. (2021). Use of advanced artificial intelligence in forensic medicine, forensic anthropology and clinical anatomy. In *Healthcare*, volume 9, page 1545. MDPI.
- Tiwari, S., Jain, G., Shetty, D. K., Sudhi, M., Balakrishnan, J. M., and Bhatta, S. R. (2023). A comprehensive review on the application of 3d convolutional neural networks in medical imaging. *Engineering Proceedings*, 59(1):3.
- Zheng, B., Zhong, Y., Al-Worafi, N. A., and Liu, Y. (2023). The dimensional and morphological assessment of the frontal sinus in sex estimation among different populations. *Head & Face Medicine*, 19(1):8.
- Zorba, E., Moraitis, K., and Manolis, S. K. (2011). Sexual dimorphism in permanent teeth of modern greeks. *Forensic science international*, 210(1-3):74–81.