

UNIVERSIDADE FEDERAL DE JUIZ DE FORA  
INSTITUTO DE CIÊNCIAS EXATAS  
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

# Modelos de aprendizado profundo para a estimativa de idade dentária através de imagens tomográficas da câmara pulpar

Alexandre Vieira Pereira Pacelli

JUIZ DE FORA  
DEZEMBRO, 2023

# Modelos de aprendizado profundo para a estimativa de idade dentária através de imagens tomográficas da câmara pulpar

ALEXANDRE VIEIRA PEREIRA PACELLI

Universidade Federal de Juiz de Fora  
Instituto de Ciências Exatas  
Departamento de Ciência da Computação  
Bacharelado em Ciência da Computação

Orientador: Heder Soares Bernardino  
Coorientador: Karina Lopes Devito

JUIZ DE FORA  
DEZEMBRO, 2023

MODELOS DE APRENDIZADO PROFUNDO PARA A ESTIMATIVA  
DE IDADE DENTÁRIA ATRAVÉS DE IMAGENS TOMOGRÁFICAS  
DA CÂMARA PULPAR

Alexandre Vieira Pereira Pacelli

MONOGRAFIA SUBMETIDA AO CORPO DOCENTE DO INSTITUTO DE CIÊNCIAS  
EXATAS DA UNIVERSIDADE FEDERAL DE JUIZ DE FORA, COMO PARTE INTE-  
GRANTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE  
BACHAREL EM CIÊNCIA DA COMPUTAÇÃO.

Aprovada por:

Heder Soares Bernardino  
Doutor em Modelagem Computacional

Karina Lopes Devito  
Doutora em Radiologia Odontológica

Saulo Moraes Villela  
Doutor em Engenharia de Sistemas e Computação

Marcelo Bernardes Vieira  
Doutor em Ciência da Computação

Leonardo Goliatt da Fonseca  
Doutor em Modelagem Computacional

JUIZ DE FORA  
20 DE DEZEMBRO, 2023

*Aos meus amigos e irmãos.*

*Aos pais, pelo apoio e sustento.*

## Resumo

A determinação da idade de um indivíduo é valiosa na ciência forense e em diversas áreas, auxiliando na identificação de vítimas em desastres e litígios civis. Para isso, existem diversos métodos para a estimativa da idade, destacando-se, na área odontológica, a idade dentária. Porém, esses métodos dependem da análise por especialistas na área, além de conter processos manuais demorados, cujos resultados são influenciados pela subjetividade do observador. Nos últimos anos, o aprendizado de máquina profundo tem sido fundamental na resolução de uma ampla variedade de problemas complexos. Por exemplo, estudos recentes que utilizam essa abordagem são capazes de prever a idade dentária com eficácia através de imagens radiográficas. Desse modo, neste trabalho é explorado o uso de modelos de aprendizado profundo, especificamente redes neurais convolucionais, para estimar a idade de indivíduos com base em imagens da câmara pulpar dos dentes obtidas através de imagens de Tomografia Computadorizada de Feixe Cônico. Os modelos desenvolvidos consideram cortes coronais e sagitais dos dentes incisivos centrais superiores, além de imagens que fazem uma junção dos dois cortes, utilizando uma base de dados da Faculdade de Odontologia da UFJF. Este estudo visa destacar o potencial desses modelos na estimativa de idade a partir de imagens tomográficas da câmara pulpar dentária, além de contribuir para a ciência forense e promover a colaboração interdisciplinar. Como resultados, o melhor modelo para os cortes juntos (concatenado) obteve acurácia de 0,527, enquanto o corte sagital obteve 0,516 e o coronal o valor de 0,429.

**Palavras-chave:** Redes Neurais Convolucionais, Estimativa da Idade Dentária, Aprendizado Profundo, Tomografia Computadorizada de Feixe Cônico, Câmara Pulpar.

# Abstract

The determination of an individual's age is valuable in forensic science and various fields, aiding in the identification of victims in disasters and civil litigation. For this purpose, there are various methods for age estimation, with dental age being prominent in the dental field. However, these methods rely on analysis by experts in the field and involve time-consuming manual processes, with results influenced by the subjectivity of the observer. In recent years, deep machine learning has been instrumental in solving a wide range of complex problems. For instance, recent studies using this approach have effectively predicted dental age through radiographic images. Thus, this work explores the use of deep learning models, specifically convolutional neural networks, to estimate the age of individuals based on images of the dental pulp chamber obtained through Cone Beam Computed Tomography images. The developed models consider coronal and sagittal sections of the upper central incisor teeth, as well as images that merge both sections, using a database from the UFJF School of Dentistry. This study aims to highlight the potential of these models in age estimation from dental pulp chamber tomographic images, contributing to forensic science and promoting interdisciplinary collaboration. As results, the best model for concatenated sections obtained an accuracy of 0.527, while the sagittal section achieved 0.516, and the coronal section achieved a value of 0.429.

**Keywords:** Convolutional Neural Networks, Dental Age Estimation, Deep Learning, Cone Beam Computed Tomography, Pulp Chamber.

## Agradecimentos

Primeiramente, gostaria de agradecer a Deus por estar ao meu lado e me ajudar a superar os desafios.

Aos meus pais, por todo o apoio, carinho e incentivo que me fornecem sempre.

Ao professor Heder Soares e professora Karina Lopes pela orientação e oportunidade de participar deste projeto, sem a qual este trabalho não se realizaria, além de todas as correções e suporte. Ao aluno William Lopes pelo apoio durante o projeto e coleta das imagens tomográficas.

Aos meus amigos pelo companheirismo e amizade ao longo desse processo.

Aos membros da banca de avaliação final pela presença e por aceitarem debater sobre a pesquisa.

Aos professores do Departamento de Ciência da Computação pelos seus ensinamentos e aos funcionários do curso, que durante esses anos, contribuíram de algum modo para o nosso enriquecimento pessoal e profissional.

*“It seems like the pressure that could have crushed us made us into diamonds instead. And what didn’t kill us actually did make us stronger”.*

*Taylor Swift*

# Conteúdo

<b>Lista de Figuras</b>	<b>8</b>
<b>Lista de Tabelas</b>	<b>11</b>
<b>Lista de Abreviações</b>	<b>12</b>
<b>1 Introdução</b>	<b>13</b>
1.1 Apresentação do tema . . . . .	13
1.2 Contextualização . . . . .	13
1.3 Descrição do Problema . . . . .	14
1.4 Justificativa / Motivação . . . . .	16
1.5 Objetivos . . . . .	17
<b>2 Fundamentação Teórica</b>	<b>18</b>
2.1 Morfologia dentária e sua relação com a idade . . . . .	18
2.2 Tomografia Computadorizada de Feixe Cônico . . . . .	20
2.3 Redes Neurais . . . . .	22
2.4 Redes Neurais Convolucionais . . . . .	24
2.5 Considerações Finais . . . . .	27
<b>3 Trabalhos Relacionados</b>	<b>28</b>
3.1 Precisão na estimativa de idade por meio de aprendizado profundo: Interferência humana ou autonomia? . . . . .	28
3.2 Determinação do grupo etário utilizando inteligência artificial em imagens do primeiro molar . . . . .	31
3.3 Estimativa de idade com base na segmentação da câmara pulpar em 3D . . . . .	34
3.4 Redes neurais profundas para estimativa da idade a partir de radiografia panorâmica . . . . .	36
3.5 Classificação de idade baseada em radiografia panorâmica através de método manual e CNN . . . . .	38
3.6 Avaliação de duas redes neurais convolucionais na estimativa de idade dentária . . . . .	40
3.7 Considerações Finais . . . . .	42
<b>4 Método Proposto</b>	<b>45</b>
4.1 Descrição dos Conjuntos de Dados . . . . .	46
4.2 Arquitetura dos Modelos . . . . .	47
4.3 Definição dos Hiperparâmetros . . . . .	49
4.4 Critérios de Avaliação . . . . .	50
<b>5 Resultados</b>	<b>53</b>
5.1 Ambiente de Desenvolvimento . . . . .	54
5.2 Bibliotecas Utilizadas . . . . .	55
5.3 Conjunto de Dados 1 - Corte Coronal . . . . .	55
5.3.1 Modelo <i>Nano</i> . . . . .	55

5.3.2	Modelo <i>Small</i> . . . . .	57
5.3.3	Modelo <i>Medium</i> . . . . .	59
5.3.4	Modelo <i>Large</i> . . . . .	61
5.3.5	Modelo <i>Extra Large</i> . . . . .	63
5.4	Conjunto de Dados 2 - Corte Sagital . . . . .	65
5.4.1	Modelo <i>Nano</i> . . . . .	65
5.4.2	Modelo <i>Small</i> . . . . .	67
5.4.3	Modelo <i>Medium</i> . . . . .	69
5.4.4	Modelo <i>Large</i> . . . . .	71
5.4.5	Modelo <i>Extra Large</i> . . . . .	73
5.5	Conjunto de Dados 3 - Corte Concatenado . . . . .	75
5.5.1	Modelo <i>Nano</i> . . . . .	75
5.5.2	Modelo <i>Small</i> . . . . .	77
5.5.3	Modelo <i>Medium</i> . . . . .	79
5.5.4	Modelo <i>Large</i> . . . . .	81
5.5.5	Modelo <i>Extra Large</i> . . . . .	83
5.6	Comparação dos Modelos . . . . .	85
5.6.1	Corte Coronal . . . . .	85
5.6.2	Corte Sagital . . . . .	90
5.6.3	Corte Concatenado . . . . .	94
5.6.4	Melhores Modelos de cada Corte . . . . .	98
5.7	Análise Visual . . . . .	99
5.7.1	Corte Coronal . . . . .	99
5.7.2	Corte Sagital . . . . .	100
5.7.3	Corte Concatenado . . . . .	101
5.8	Considerações Finais . . . . .	102
<b>6</b>	<b>Conclusões</b> . . . . .	<b>105</b>
6.1	Desafios . . . . .	106
6.2	Trabalhos Futuros . . . . .	109
	<b>Bibliografia</b> . . . . .	<b>110</b>
<b>A</b>	<b>Anexo</b> . . . . .	<b>114</b>
A.1	Parecer do Comitê de Ética em Pesquisa . . . . .	114

## Lista de Figuras

1.1	Câmara pulpar e relação com a idade. . . . .	15
2.1	Corte transversal de um molar de um indivíduo adulto (BRITANNICA, 2023). Adaptado pelo autor. . . . .	19
2.2	Princípio de aquisição de imagens na TCFC, em que a fonte de raios X e o receptor de imagem giram em torno do paciente em um arco de 180 a 360 graus (VENKATESH; ELLURU, 2017). . . . .	21
2.3	Modelo de um neurônio artificial (Elaborado pelo autor). . . . .	22
2.4	Calculando os valores de saída de uma convolução (DUMOULIN; VISIN, 2016). . . . .	25
2.5	Visão geral de uma arquitetura genérica de rede neural convolucional e do processo de treinamento (YAMASHITA et al., 2018). . . . .	26
3.1	Visão geral dos modelos. <b>(a)</b> MDAE, <b>(b)</b> ADSE, <b>(c)</b> ADAE. (HAN et al., 2022). . . . .	30
3.2	Comparação entre os valores de MAE para cada um dos modelos (HAN et al., 2022). . . . .	30
3.3	Classificação de uma radiografia panorâmica usando a CNN para prever a faixa etária do paciente (KIM et al., 2021). . . . .	32
3.4	Agrupamento do resultado da previsão da faixa etária de cinco faixas etárias para as três faixas etárias (KIM et al., 2021). . . . .	32
3.5	Matriz de confusão dos resultados da estimativa de faixa etária (KIM et al., 2021). . . . .	33
3.6	Fluxo de trabalho para a análise de imagens da TCFC (ZHENG et al., 2020). . . . .	35
3.7	Gráficos de dispersão para todos os dentes de primeiros molares em dados de teste (ZHENG et al., 2020). . . . .	35
3.8	Arquitetura das duas redes propostas. Em vermelho o caminho para classificação da idade e em azul o caminho para predição do sexo (VILA-BLANCO et al., 2020). . . . .	37
3.9	Idade real comparada às previsões do modelo DASNet avaliadas para todo o conjunto de dados, $S_{all}$ (VILA-BLANCO et al., 2020). . . . .	38
3.10	Erro absoluto do DASNet na estimativa de idade cronológica ao treinar com conjuntos de dados reduzidos e avaliar o desempenho para diferentes faixas etárias (VILA-BLANCO et al., 2020). . . . .	39
3.11	Idade real comparada às previsões do DASNet para alguns conjuntos de dados reduzidos gerados para comparar com os métodos manuais de estimativa (VILA-BLANCO et al., 2020). . . . .	39
3.12	Estrutura básica da rede VGG16 (WANG et al., 2023). . . . .	41
3.13	Diagrama com a estrutura básica da rede ResNet101 (WANG et al., 2023). . . . .	41
4.1	Imagens ilustrativas de cortes do dente incisivo central superior de um mesmo indivíduo . . . . .	46
4.2	Imagem ilustrativa da junção dos cortes coronal e sagital em uma única imagem (corte concatenado). . . . .	47

4.3	Gráfico contendo a distribuição das 555 imagens de cada corte para cada uma das 8 classes. . . . .	48
4.4	Imagem ilustrativa de uma matriz de confusão gerada em um dos experimentos realizados. . . . .	51
5.1	Visão geral dos experimentos para cada conjunto de dados utilizando as 5 versões de modelo. . . . .	54
5.2	Matriz de confusão normalizada do conjunto de teste para o corte coronal, modelo <i>nano</i> com a melhor semente aleatória identificada. . . . .	56
5.3	Resultados do treinamento para o corte coronal, modelo <i>nano</i> com a melhor semente aleatória identificada. . . . .	57
5.4	Matriz de confusão normalizada do conjunto de teste para o corte coronal, modelo <i>Small</i> com a melhor semente aleatória identificada. . . . .	58
5.5	Resultados do treinamento para o corte coronal, modelo <i>Small</i> com a melhor semente aleatória identificada. . . . .	59
5.6	Matriz de confusão normalizada do conjunto de teste para o corte coronal, modelo <i>Medium</i> com a melhor semente aleatória identificada. . . . .	60
5.7	Resultados do treinamento para o corte coronal, modelo <i>Medium</i> com a melhor semente aleatória identificada. . . . .	61
5.8	Matriz de confusão normalizada do conjunto de teste para o corte coronal, modelo <i>Large</i> com a melhor semente aleatória identificada. . . . .	62
5.9	Resultados do treinamento para o corte coronal, modelo <i>Large</i> com a melhor semente aleatória identificada. . . . .	63
5.10	Matriz de confusão normalizada do conjunto de teste para o corte coronal, modelo <i>Extra Large</i> com a melhor semente aleatória identificada. . . . .	64
5.11	Resultados do treinamento para o corte coronal, modelo <i>Extra Large</i> com a melhor semente aleatória identificada. . . . .	65
5.12	Matriz de confusão normalizada do conjunto de teste para o corte sagital, modelo <i>Nano</i> com a melhor semente aleatória identificada. . . . .	66
5.13	Resultados do treinamento para o corte sagital, modelo <i>Nano</i> com a melhor semente aleatória identificada. . . . .	67
5.14	Matriz de confusão normalizada do conjunto de teste para o corte sagital, modelo <i>Small</i> com a melhor semente aleatória identificada. . . . .	68
5.15	Resultados do treinamento para o corte sagital, modelo <i>Small</i> com a melhor semente aleatória identificada. . . . .	69
5.16	Matriz de confusão normalizada do conjunto de teste para o corte sagital, modelo <i>Medium</i> com a melhor semente aleatória identificada. . . . .	70
5.17	Resultados do treinamento para o corte sagital, modelo <i>Medium</i> com a melhor semente aleatória identificada. . . . .	71
5.18	Matriz de confusão normalizada do conjunto de teste para o corte sagital, modelo <i>Large</i> com a melhor semente aleatória identificada. . . . .	72
5.19	Resultados do treinamento para o corte sagital, modelo <i>Large</i> com a melhor semente aleatória identificada. . . . .	73
5.20	Matriz de confusão normalizada do conjunto de teste para o corte sagital, modelo <i>Extra Large</i> com a melhor semente aleatória identificada. . . . .	74
5.21	Resultados do treinamento para o corte sagital, modelo <i>Extra Large</i> com a melhor semente aleatória identificada. . . . .	75
5.22	Matriz de confusão normalizada do conjunto de teste para o corte concatenado, modelo <i>Nano</i> com a melhor semente aleatória identificada. . . . .	76

5.23	Resultados do treinamento para o corte concatenado, modelo <i>Nano</i> com a melhor semente aleatória identificada. . . . .	77
5.24	Matriz de confusão normalizada do conjunto de teste para o corte concatenado, modelo <i>Small</i> com a melhor semente aleatória identificada. . . . .	78
5.25	Resultados do treinamento para o corte concatenado, modelo <i>Small</i> com a melhor semente aleatória identificada. . . . .	79
5.26	Matriz de confusão normalizada do conjunto de teste para o corte concatenado, modelo <i>Medium</i> com a melhor semente aleatória identificada. . . . .	80
5.27	Resultados do treinamento para o corte concatenado, modelo <i>Medium</i> com a melhor semente aleatória identificada. . . . .	81
5.28	Matriz de confusão normalizada do conjunto de teste para o corte concatenado, modelo <i>Large</i> com a melhor semente aleatória identificada. . . . .	82
5.29	Resultados do treinamento para o corte concatenado, modelo <i>Large</i> com a melhor semente aleatória identificada. . . . .	83
5.30	Matriz de confusão normalizada do conjunto de teste para o corte concatenado, modelo <i>Extra Large</i> com a melhor semente aleatória identificada. . . . .	84
5.31	Resultados do treinamento para o corte concatenado, modelo <i>Extra Large</i> com a melhor semente aleatória identificada. . . . .	85
5.32	<i>Boxplot</i> apresentando a variação das Acurácias de cada modelo a partir das 10 sementes aleatórias para o corte coronal. . . . .	87
5.33	<i>Boxplot</i> apresentando a variação dos valores de <i>Recall</i> Médio de cada modelo a partir das 10 sementes aleatórias para o corte coronal. . . . .	88
5.34	<i>Boxplot</i> apresentando a variação dos valores de Precisão Média de cada modelo a partir das 10 sementes aleatórias para o corte coronal. . . . .	89
5.35	<i>Boxplot</i> apresentando a variação dos valores de <i>F-Measure</i> Médio de cada modelo a partir das 10 sementes aleatórias para o corte coronal. . . . .	90
5.36	<i>Boxplot</i> apresentando a variação das Acurácias de cada modelo a partir das 10 sementes aleatórias para o corte sagital. . . . .	91
5.37	<i>Boxplot</i> apresentando a variação dos valores de <i>Recall</i> Médio de cada modelo a partir das 10 sementes aleatórias para o corte sagital. . . . .	92
5.38	<i>Boxplot</i> apresentando a variação dos valores de Precisão Média de cada modelo a partir das 10 sementes aleatórias para o corte sagital. . . . .	93
5.39	<i>Boxplot</i> apresentando a variação dos valores de <i>F-Measure</i> Médio de cada modelo a partir das 10 sementes aleatórias para o corte sagital. . . . .	94
5.40	<i>Boxplot</i> apresentando a variação das Acurácias de cada modelo a partir das 10 sementes aleatórias para o corte concatenado. . . . .	95
5.41	<i>Boxplot</i> apresentando a variação dos valores de <i>Recall</i> Médio de cada modelo a partir das 10 sementes aleatórias para o corte concatenado. . . . .	96
5.42	<i>Boxplot</i> apresentando a variação dos valores de Precisão Média de cada modelo a partir das 10 sementes aleatórias para o corte concatenado. . . . .	97
5.43	<i>Boxplot</i> apresentando a variação dos valores de <i>F-Measure</i> Médio de cada modelo a partir das 10 sementes aleatórias para o corte concatenado. . . . .	98
5.44	Mapa de destaque utilizando EigenCAM para imagens do corte coronal. . . . .	100
5.45	Mapa de destaque utilizando EigenCAM para imagens do corte sagital. . . . .	101
5.46	Mapa de destaque utilizando EigenCAM para imagens do corte concatenado. . . . .	102
5.47	Imagem ilustrativa das diferenças entre os cortes coronais e sagitais. . . . .	103
6.1	Comparação entre estruturas dentárias semelhantes de indivíduos de faixas etárias diferentes, com foco na câmara pulpar . . . . .	108
6.2	Câmara pulpar reduzida mesmo em idades mais jovens . . . . .	109

## Lista de Tabelas

3.1	Acurácia dos modelos para determinados limiares de idade (WANG et al., 2023). . . . .	42
3.2	Comparação entre os trabalhos relacionados. . . . .	44
5.1	Comparação entre as versões do modelo para o Corte Coronal . . . . .	86
5.2	Comparação entre as versões do modelo para o Corte Sagital . . . . .	91
5.3	Comparação entre as versões do modelo para o Corte Concatenado . . . .	95
5.4	Comparação entre os melhores modelos gerados para cada tipo de corte . .	99

## Lista de Abreviações

AP	Aprendizado Profundo
ANN	<i>Artificial Neural Network</i>
CNN	<i>Convolutional Neural Network</i>
DCC	Departamento de Ciência da Computação
MAE	<i>Mean Absolute Error</i>
TC	Tomografia Computadorizada
TCFC	Tomografia Computadorizada de Feixe Cônico
UFJF	Universidade Federal de Juiz de Fora

# 1 Introdução

## 1.1 Apresentação do tema

Os dentes constituem a parte mais dura do esqueleto humano e as modificações sofridas por eles são pouco influenciadas por fatores externos, sendo resistentes a carbonizações, traumas, efeitos de putrefação e tempo, além de não serem muito influenciados por condições nutricionais, médicas e de vida a que o indivíduo foi submetido (STAR et al., 2011). O processo de desenvolvimento dentário segue um padrão específico com diferentes estágios ao longo da vida, cada um de forma a deixar uma marca na estrutura dentária (ABOSHI; TAKAHASHI; KOMURO, 2010).

Por conta disso, o dente se torna uma valiosa fonte de informação para a estimativa de idade dos indivíduos e é amplamente utilizado pela ciência forense. Sua análise pode incluir a inspeção de características, como o desgaste do esmalte e a redução do tamanho da câmara pulpar pelo depósito de dentina secundária para determinar a idade aproximada de um indivíduo.

Porém, métodos manuais de análise de idade por meio do dente podem ser demorados e requerem a experiência de um profissional treinado. Desse modo, estimativas a curto prazo e precisas se tornam fundamentais, assim como modelos de aprendizado profundo (AP) podem ajudar a acelerar o processo e torná-lo mais preciso.

## 1.2 Contextualização

Assim como na computação, a área odontológica passa por uma constante evolução tecnológica, trazendo novidades que ajudam a melhorar a qualidade dos diagnósticos e, conseqüentemente, dos tratamentos. Além disso, um dos avanços mais significativos na odontologia foi a introdução de diferentes tipos de exames por imagem (VENKATESH; ELLURU, 2017).

Nesse cenário, tem-se a radiografia panorâmica como uma das mais comumente

utilizadas. Nesta, é possível ter uma visão de toda a área da cavidade bucal, ou seja, todos os dentes das arcadas superior e inferior em apenas uma imagem. Além disso, há exames por imagem mais recentes como a tomografia computadorizada de feixe cônico (TCFC). Esta é capaz de fornecer imagens tridimensionais da região maxilomandibular a partir de vários planos, oferecendo, também, uma imagem mais rica em detalhes, que melhora a acurácia do diagnóstico de alterações que acometem dentes e tecido ósseo.

Ademais, o surgimento das redes neurais artificiais, termo mais referenciado na literatura como *Artificial Neural Network* (ANN), representa uma evolução significativa na área de modelos computacionais. Inspiradas no funcionamento do cérebro humano, elas são capazes de aprender a partir de dados e executar tarefas complexas, como classificação, reconhecimento de padrões e previsão.

Nesse contexto, uma das mais recentes inovações em ANNs são as redes neurais convolucionais, mais referenciadas como *Convolutional Neural Network* (CNN) em trabalhos literários. Utilizando-se de camadas convolucionais para processar e analisar imagens, elas são capazes de extrair características e padrões relevantes, tornando-as bastante utilizadas em tarefas de classificação, detecção e reconhecimento de objetos em imagens (GOODFELLOW; BENGIO; COURVILLE, 2016).

Diante disso, neste trabalho serão utilizadas imagens geradas a partir da TCFC pelo departamento de odontologia da Universidade Federal de Juiz de Fora (UFJF). Assim, com os dados gerados, modelos de aprendizado profundo, alinhados com a capacidade das CNNs de extraírem características das imagens, poderão ser treinados.

Desse modo, assim como nos trabalhos de Zheng et al. (2020) e Han et al. (2022) em que modelos de aprendizado profundo foram utilizados para auxiliar na estimativa da idade a partir de imagens radiográficas, o presente trabalho poderá avaliar os modelos propostos e contribuir com novas soluções para a estimativa de idade a partir da análise da câmara pulpar de dentes.

### 1.3 Descrição do Problema

Um dente é composto por duas partes principais: a coroa, que é a parte visível acima da gengiva e a raiz, que é a parte que fica intraóssea e não visível clinicamente (BRITAN-

NICA, 2023).

A coroa é coberta pelo esmalte que protege a dentina, que é uma camada mais interna e que compõe a maior parte do dente. O centro do dente formado pela câmara pulpar, que contém a polpa dentária, rica em nervos e vasos sanguíneos, sendo responsável por fornecer nutrientes ao dente e transmitir sensações de dor e temperatura para o cérebro.

Cada uma dessas partes pode se relacionar com a tentativa de estimar a idade de um indivíduo na ciência forense. Por exemplo, após a formação completa do dente, inicia-se a deposição de dentina secundária nas paredes internas da câmara pulpar, reduzindo as dimensões dessa cavidade ao longo da vida. Essa deposição de dentina secundária, visível em exames por imagem (radiografias e tomografias), pode auxiliar na estimativa de idade de indivíduos adultos. Na Figura 1.1a, é possível observar a câmara pulpar abrangendo uma área considerável do dente. Com a progressão da idade, como ilustrado na Figura 1.1c, nota-se a redução da dimensão da câmara pulpar.

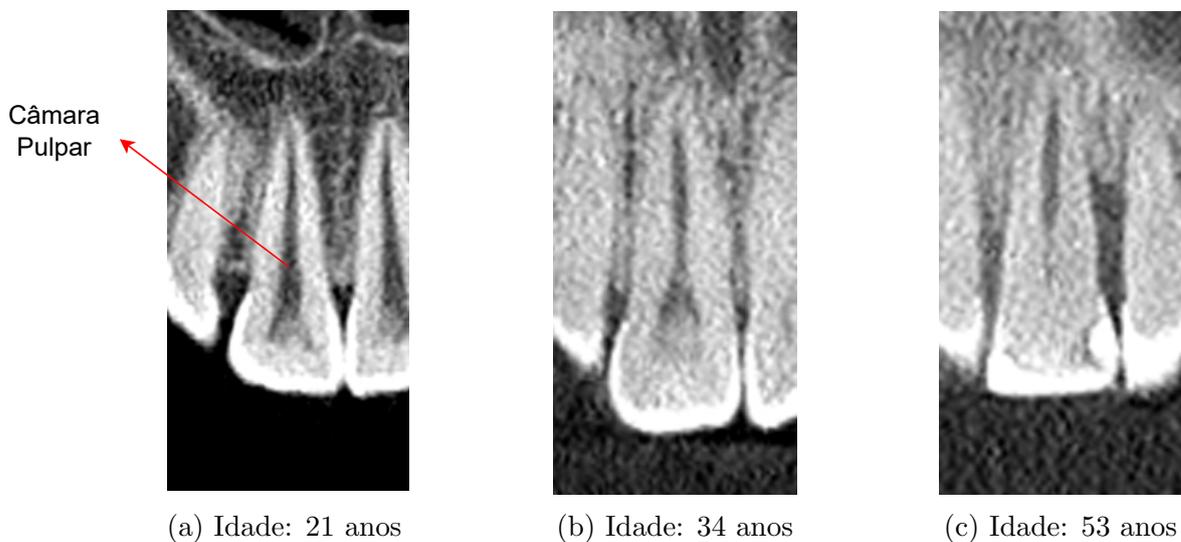


Figura 1.1: Câmara pulpar e relação com a idade.

Assim, este trabalho tem como foco a estimativa da idade cronológica, a partir da avaliação da câmara pulpar usando de modelos de aprendizado profundo, por meio de exames por imagem. Nesse contexto, a tomografia computadorizada de feixe cônico é uma ferramenta útil, pois permite uma visualização precisa e detalhada da estrutura interna do dente. Além disso, trabalhos da literatura têm apresentado resultados satisfatórios na estimativa de idade ao utilizar imagens de TCFC (KAZMI et al., 2019; PINCHI et al.,

2015). Com relação à pesquisa de Zheng et al. (2020), esta utilizou-se de AP para auxiliar nessa estimativa.

Em suma, nesta pesquisa, com o auxílio de exames de TCFC, imagens dos dentes serão geradas a partir dos seus cortes coronais e sagitais. Desse modo, elas servirão para o treinamento de modelos de aprendizado profundo que, posteriormente, serão capazes de realizar classificação da idade dos indivíduos.

## 1.4 Justificativa / Motivação

A estimativa da idade de um indivíduo é de grande importância para a ciência forense, assim como em outros campos, já que a determinação da idade de vítimas e suspeitos é necessária para a identificação em casos como desastres em massa, crimes organizados e casos de abuso (WILLEMS; MOULIN-ROMSEE; SOLHEIM, 2002; WILLEMS, 2001).

Além disso, houve um grande aumento de pesquisas destinadas a criar modelos que possam estimar a idade de pessoas a partir de dados obtidos por imagens radiográficas de seus dentes.

Porém, a literatura brasileira na área de aprendizado profundo e, principalmente, em relação à criação de modelos para prever a idade de um indivíduo a partir de seus dentes ainda necessita de maiores estudos. Logo, este trabalho pretende contribuir para a área ao propor três modelos que possam estimar a idade por meio de análise de imagens de TCFC com foco em analisar a câmara pulpar do dente. Cada um dos modelos será responsável por trabalhar com um tipo de corte dentário: coronal, sagital e o corte concatenado que fará a junção dos outros dois cortes em uma única imagem.

Além disso, este trabalho irá utilizar um tipo de AP que são as redes neurais convolucionais. Elas têm sido amplamente utilizadas na literatura e no campo de reconhecimento de imagens. Na análise de imagens médicas, por exemplo, são capazes de extrair características difíceis de serem detectadas por um ser humano, contribuindo para o problema de estimativa de idade por meio de radiografias (HAN et al., 2022). Ademais, esta pesquisa conta com a criação de seu próprio conjunto de dados de imagens de TCFC, favorecendo futuros trabalhos com o intuito de utilizar de AP para a estimativa de idade de seres humanos.

Portanto, em resumo, este trabalho se motiva ao desenvolvimento de modelos de aprendizado profundo, utilizando-se das redes neurais convolucionais, além da criação do seu próprio banco de imagens de TCFC. Ademais, este é um estudo inicial na estimativa da idade dos indivíduos deste grupo de pesquisa e, apesar da base de dados utilizada ainda não ser muito extensa, o resultados alcançados são promissores.

## 1.5 Objetivos

Esse trabalho possui o seguinte objetivo principal: desenvolver e analisar modelos de aprendizado profundo para estimar a idade através de imagens da câmara pulpar dos dentes incisivos centrais superiores.

Além disso, esse estudo tem como objetivos específicos:

- iniciar a criação de uma base de dados de imagens coletadas na UFJF;
- desenvolver modelos de aprendizado profundo utilizando-se de redes neurais convolucionais para a identificação da idade;
- disponibilizar os códigos e modelos gerados.

Dessa forma, a partir desses objetivos, tem-se que este trabalho também poderá contribuir com os demais fatores:

- incentivar a colaboração entre o Departamento de Ciência da Computação e outros domínios da Universidade;
- colaboração com a ciência forense para auxiliar na identificação da idade de indivíduos.

## 2 Fundamentação Teórica

Como o estudo presente neste trabalho aborda conceitos multidisciplinares, sua fundamentação teórica, além de abranger o campo da ciência da computação, também está relacionada à área da saúde, em particular, o domínio odontológico.

Neste caso, dois conceitos importantes da odontologia serão tratados neste capítulo. O primeiro será a morfologia do dente humano, para que se possa esclarecer a forma e estrutura dos dentes. Já o segundo conceito se relaciona à tomografia computadorizada de feixe cônico, com informações sobre como essa tecnologia é capaz de produzir imagens tridimensionais de alta resolução dos tecidos e estruturas bucais, possibilitando o uso dessas imagens no presente estudo e em outros trabalhos da literatura.

Por fim, também serão abordados dois conceitos importantes da ciência da computação para esta pesquisa: Redes Neurais Artificiais e Redes Neurais Convolucionais. Assim, será possível compreender como as ANNs funcionam e obter conhecimento sobre a capacidade das CNNs de analisar imagens de forma eficiente, identificando padrões relevantes e gerando seus resultados.

### 2.1 Morfologia dentária e sua relação com a idade

Como pode ser observado na Figura 2.1 todos os dentes possuem a mesma estrutura básica, consistindo em três camadas principais: esmalte, dentina e polpa. O esmalte é a camada mais externa, formada por minerais inorgânicos e sendo o tecido mais duro do corpo humano. A dentina, que é menos dura que o esmalte, compõe a maior parte do dente e é formada por células e tecido mineralizado. A polpa é a camada mais interna do dente e é composta por células, pequenos vasos sanguíneos e nervos. Ela se estende por quase toda a extensão do dente, desde a coroa até a raiz, e é responsável por nutrir o dente (BRITANNICA, 2023).

A raiz do dente localiza-se dentro do osso alveolar, abaixo da gengiva e é coberta pelo cimento, uma substância semelhante ao osso, mas menos dura que a dentina. O

cemento serve como meio de fixação para as fibras (ligamento periodontal) que mantêm o dente estável no osso alveolar (BRITANNICA, 2023).

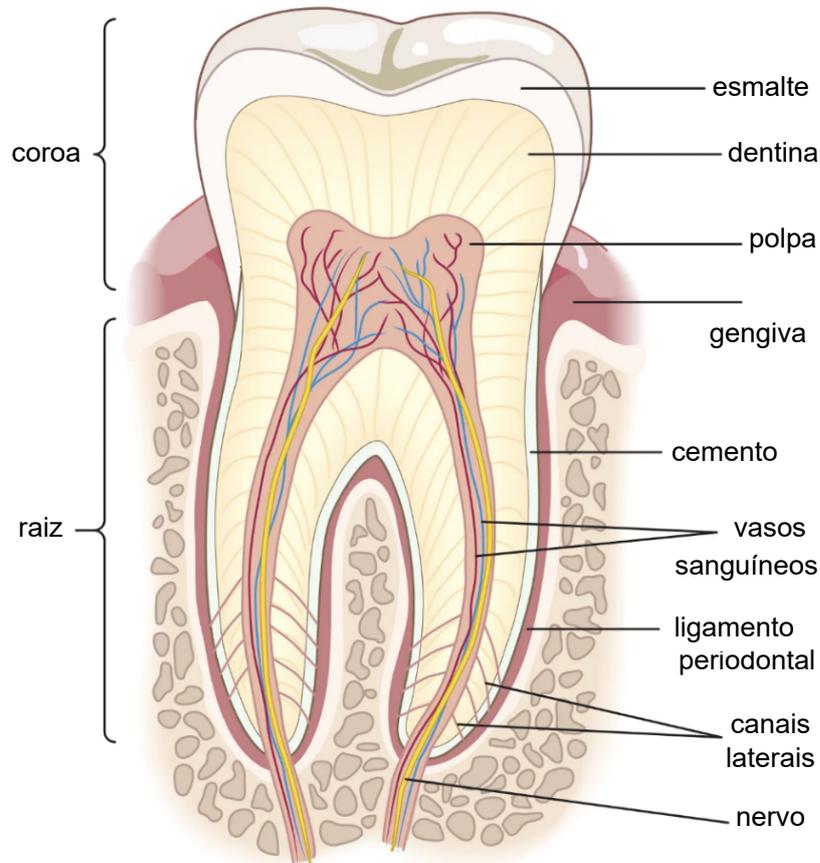


Figura 2.1: Corte transversal de um molar de um indivíduo adulto (BRITANNICA, 2023). Adaptado pelo autor.

Diante do exposto, essas estruturas dentárias acabam apresentando alterações ao longo do tempo, sendo que essas alterações podem ser usadas para estimar a idade em investigações forenses (GUPTA et al., 2014). Por exemplo, Yang, Jacobs e Willems (2006) mostram resultados promissores para estimativa de idade dentária usando imagens de TCFC em indivíduos vivos. Nesse caso, a pesquisa teve o intuito estabelecer uma correlação entre a idade cronológica de um indivíduo e a relação entre o volume da polpa de um dente específico.

Além disso, Shah, Ranghani e Limdiwala (2020) mostram em seu estudo como a dentina também pode ser analisada para se estimar a idade. Avaliando-se o comprimento da translucidez da dentina em seções retificadas de dentes, eles foram capazes de obterem resultados promissores na previsão da idade.

## 2.2 Tomografia Computadorizada de Feixe Cônico

Nesta seção, será apresentada a tomografia computadorizada de feixe cônico. Ela será utilizada neste trabalho com o intuito de produzir imagens tridimensionais de alta resolução dos dentes humanos.

GN Hounsfield, em 1972 introduziu a varredura axial transversal computadorizada, o que levou ao surgimento da tomografia computadorizada (TC). Porém, o alto custo, o acesso limitado e a alta exposição à radiação foram os principais inconvenientes para a subutilização da TC na odontologia (GOLDMAN, 2007).

Já na década de 90, a tomografia computadorizada de feixe cônico foi introduzida por Arai et al. (1999), assim como Mozzo et al. (1998), sendo que esses dois grupos trabalharam de forma independente. Essa tecnologia ofereceu a possibilidade da exploração 3D e imagens geradas com maior precisão do que as imagens 2D em aplicações orais e maxilofaciais. Serão apresentadas, ainda nesta seção, vantagens que fizeram com que a TCFC fosse introduzida rapidamente no campo da odontologia.

Alguns estudos literários descrevem o procedimento básico da TCFC (KUMAR et al., 2015; VENKATESH; ELLURU, 2017). De forma técnica, a TCFC utiliza de um feixe de raios X cônico com uma única varredura de até 360 graus, como ilustrado na Figura 2.2, onde a fonte de raios X e sensores para detecção se movem simultaneamente ao redor da cabeça do paciente, que é estabilizada por meio de um equipamento próprio. Esse processo permite a obtenção de um conjunto de imagens, variando de 180 a 1024, que são utilizadas para criar representações cefalométricas 2D.

As reconstruções 2D adquiridas pela TCFC são processadas pelo computador por meio do algoritmo de Feldkamp modificado. Esse algoritmo reconstrói as imagens em um volume anatômico que pode ser visualizado nos planos axial, coronal e sagital, todos em proporção 1:1 e ortogonais entre si (FELDKAMP; DAVIS; KRESS, 1984).

Desse modo, a TCFC acaba por apresentar vantagens para a imagem oral e maxilofacial quando comparada à TC convencional (PHAROAH; WHITE, 2009). Ela possui a capacidade de reduzir a exposição desnecessária do paciente e minimizar a radiação espalhada que degradaria a qualidade da imagem (PHAROAH; WHITE, 2009; PALOMO; PALOMO, 2009). Além disso, garante uma maior precisão da imagem atendendo à ne-

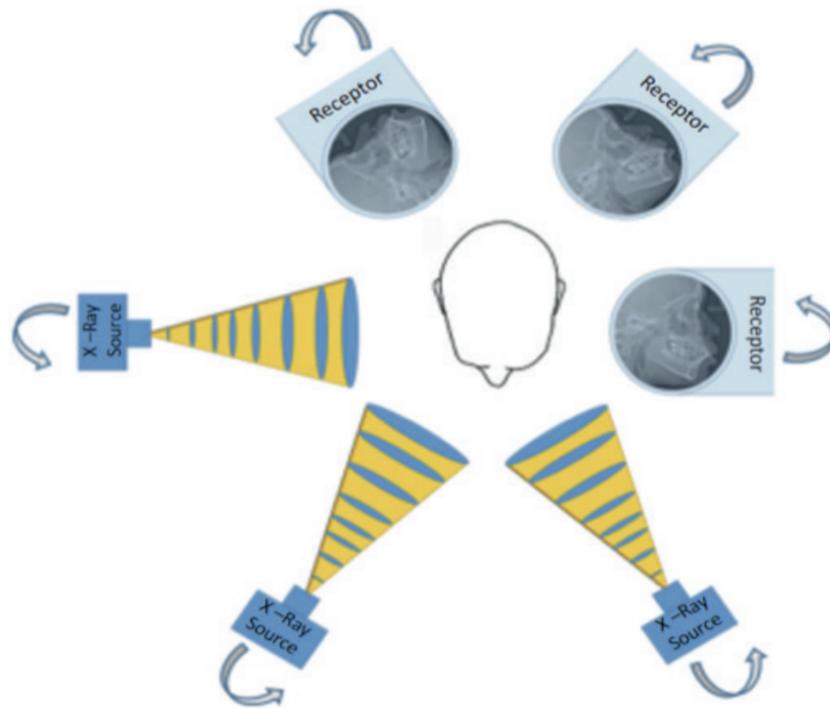


Figura 2.2: Princípio de aquisição de imagens na TCFC, em que a fonte de raios X e o receptor de imagem giram em torno do paciente em um arco de 180 a 360 graus (VENKATESH; ELLURU, 2017).

cessidade de exatidão para avaliação do local (SCARFE; FARMAN; SUKOVIC, 2006; RAZAVI et al., 2010).

Outra vantagem é seu tempo de varredura rápido, isso devido à rotação única que é utilizada na tecnologia da TCFC para adquirir imagens de base para imagens 3D (PHAROAH; WHITE, 2009). Ademais, este método proporciona reduções significativas na dose de radiação, que variam de 98,5% a 76,2%, em comparação com a tomografia computadorizada convencional (SCHULZE et al., 2004).

Diante disso, tem-se que a TCFC é uma grande aliada em aplicações da odontologia forense, especialmente na identificação da idade de um indivíduo. Nesse caso, o esmalte dentário é uma estrutura naturalmente resistente às mudanças, permitindo que os dentes resistem mesmo em situações desfavoráveis; enquanto que a câmara pulpar e a dentina apresentam alterações fisiológicas e patológicas à medida que a idade avança, ocorrendo a redução do tamanho da câmara pulpar pelo depósito de dentina secundária. A fim de quantificar essas alterações, muitas vezes é necessária a extração e seção dos dentes, o que pode não ser uma opção viável. Desse modo, a tomografia computadorizada

de feixe cônico surge como uma alternativa não invasiva para essa finalidade (YANG; JACOBS; WILLEMS, 2006).

## 2.3 Redes Neurais

Uma rede neural artificial, concebida como uma abstração do funcionamento do cérebro humano, é um programa computacional desenvolvido para simular a maneira como o cérebro humano processa informações. Analogamente ao cérebro humano, as ANNs são compostas por uma quantidade significativa de elementos computacionais de processamento simples, denominados neurônios artificiais (JAIN; MAO; MOHIUDDIN, 1996).

Além disso, elas assemelham-se ao cérebro nos aspectos de aquisição de conhecimento através do ambiente e na definição dos pesos sinápticos, que representam a força de conexão entre os neurônios e são usados para armazenar o conhecimento adquirido (HAYKIN, 2001).

A unidade de processamento básica de uma rede neural vista na Figura 2.3 é, assim como no sistema nervoso, o neurônio. Podem ser identificados três elementos básicos em um neurônio de um modelo neuronal artificial:

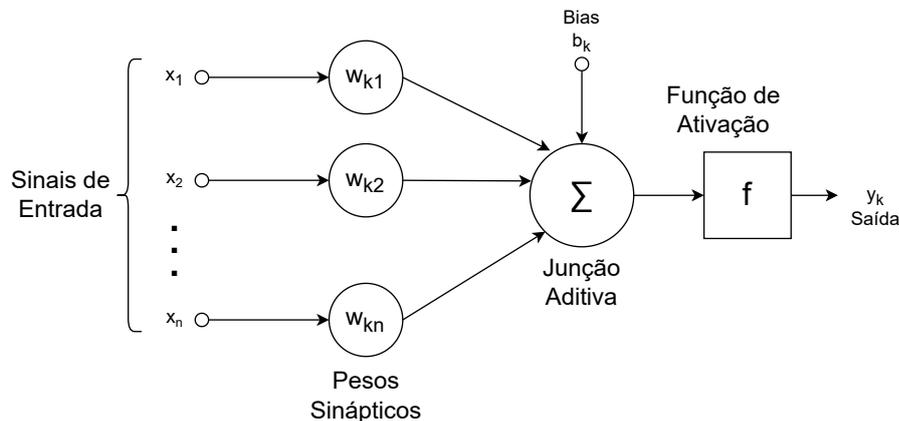


Figura 2.3: Modelo de um neurônio artificial (Elaborado pelo autor).

1. conjunto de sinapses, que são os elos de conexão entre os sinais de entrada e a junção aditiva onde cada sinapse é caracterizada por um peso próprio que ao contrário de um neurônio biológico pode assumir valores negativos. O sinal de entrada é multiplicado pelo peso e então é encaminhado à junção aditiva;

2. junção aditiva, ou combinador linear, é a parte que faz a soma dos sinais vindos das sinapses após a multiplicação feita e que encaminha o resultado a uma função de ativação. A esse resultado pode ser adicionado o parâmetro *bias*, um valor constante capaz de ajustar a saída do neurônio, controlando seu nível de ativação, independentemente das entradas;
3. a função de ativação é uma função matemática que é aplicada ao valor de ativação do neurônio (resultado da soma ponderada das entradas), determinando o valor de sua saída.

As funções de ativação são componentes importantes em redes neurais artificiais e estão presentes em todos os neurônios artificiais. Elas transformam a soma ponderada dos sinais de entrada em uma saída não-linear, permitindo que a rede modele relações complexas entre os dados de entrada e saída para resolver tarefas desafiadoras de aprendizado de máquina. Existem várias funções de ativação comumente usadas, como a sigmóide, *Rectified Linear Unit* (RELU) e tangente hiperbólica. A escolha da função de ativação depende da tarefa de aprendizado e das características dos dados de entrada e saída (GOODFELLOW; BENGIO; COURVILLE, 2016).

Finalmente, a partir das unidades de processamento simples, que são os neurônios artificiais, é possível agrupá-los em camadas formando uma rede neural artificial. Essas camadas são responsáveis por processar as informações de entrada de forma sequencial, onde cada camada realiza um processamento específico das informações e passa o resultado adiante para a próxima camada.

As ANNs podem ser construídas com várias camadas. Além disso, a forma como as informações são processadas pode variar de acordo com a organização e conectividade dessas camadas. Por exemplo, uma rede neural convolucional, que será apresentada na próxima seção, é utilizada, principalmente, para processamento de imagens, enquanto uma rede neural recorrente é frequentemente usada para processar sequências de dados, como, por exemplo, áudio ou texto (GOODFELLOW; BENGIO; COURVILLE, 2016). Um maior detalhamento dessas camadas e do processo de treinamento de uma CNN serão apresentados na próxima seção.

## 2.4 Redes Neurais Convolucionais

A rede neural convolucional é um tipo de rede neural profunda bastante famoso e comumente utilizado no campo de aprendizado profundo (LI et al., 2021; JHONG et al., 2020; ZHOU, 2020). Além disso, elas têm sido amplamente aplicadas em vários campos, como visão computacional (FANG et al., 2020), processamento de fala (PALAZ; MAGIMAI-DOSS; COLLOBERT, 2019), reconhecimento facial (LI; DENG; CHIANG, 2020), entre outros.

As CNNs foram bem detalhadas por Goodfellow, Bengio e Courville (2016). Porém, primeiramente, é importante apresentar uma visão geral do que é o pixel e como uma imagem é formada já que esse tipo de rede neural é utilizada para imagens. A imagem é formada por uma grade de pixels, sendo que cada um desses quadrados formam a imagem como um todo quando combinados. Por exemplo, se uma imagem possui 32 pixels em sua altura e largura, ela terá no total 1024 pixels.

Um algoritmo de aprendizado profundo como a CNN é capaz de extrair características das imagens recebendo-as como entrada, aplicando-se convoluções recursivas para extração de padrões dos dados.

Uma das características importantes de uma CNN é o operador de convolução. A convolução é um processo fundamental no qual cada elemento da imagem é combinado com seus vizinhos, utilizando para isso um núcleo de filtro específico, também conhecido como *kernel*. Os núcleos de filtro, como matrizes bidimensionais, são usados na convolução para aplicar um filtro na imagem de entrada. Eles são compostos por valores numéricos que determinam o comportamento do filtro, como suavização, detecção de borda, realce de detalhes, entre outros. Esses tipos de filtro são mais detalhados por Gonzalez e Woods (2018), sendo eles o núcleo Gaussiano, que é usado para suavização; o núcleo de Sobel, que é usado para detecção de bordas; e o núcleo Laplaciano, que é usado para realce de detalhes.

Desse modo, a etapa de convolução tem a imagem como uma matriz e o núcleo de filtro que é outra matriz. Este núcleo, representando uma matriz de pesos, realiza cálculos por meio de operações de multiplicação e soma, permitindo a modificação de pixels específicos na imagem. Assim, a convolução na CNN consiste em uma operação

matricial que proporciona uma abordagem para a extração de características.

Depois que o núcleo é aplicado em uma imagem durante o processo de convolução, obtém-se uma nova imagem filtrada ou convoluída. Essa nova imagem pode ter diferentes características e realçar diferentes aspectos da imagem original, dependendo do tipo de núcleo utilizado e do objetivo da convolução. A aplicação do núcleo geralmente é usada para melhorar a qualidade das imagens, remover ruídos e destacar determinadas características, tornando as imagens mais adequadas para análise e interpretação.

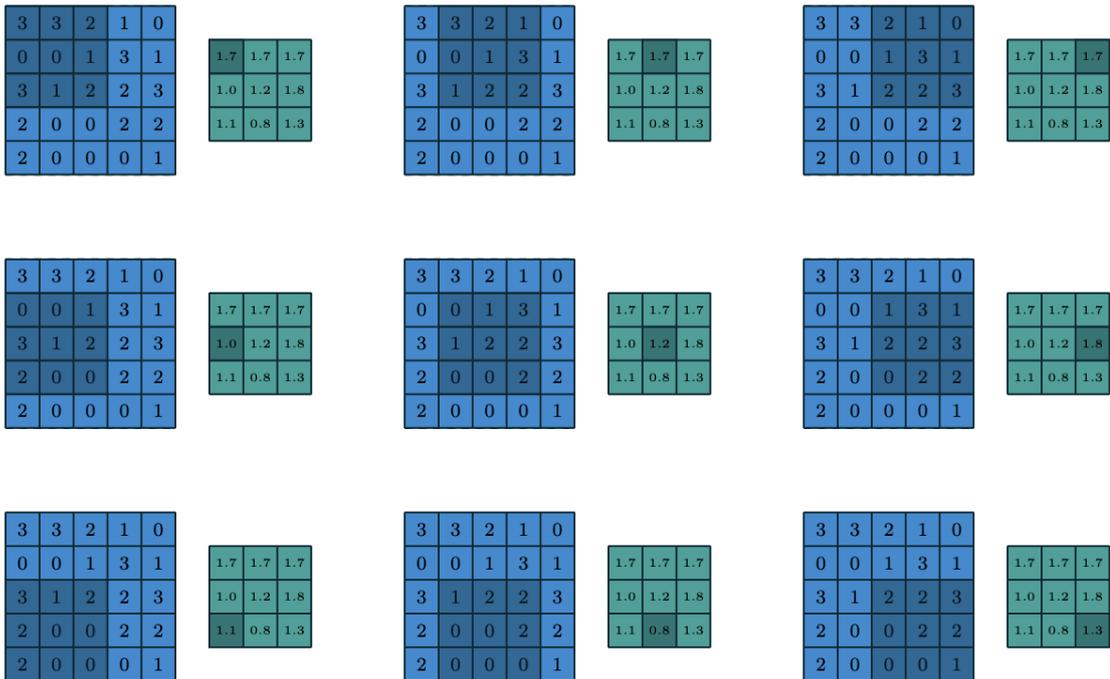


Figura 2.4: Calculando os valores de saída de uma convolução (DUMOULIN; VISIN, 2016).

Na Figura 2.4 há uma ilustração do núcleo de filtro sendo aplicado. Ele inicia no canto superior esquerdo da imagem, e realiza uma multiplicação em cada parte da matriz. Ao realizar essa multiplicação, o resultado é colocado em uma matriz de saída. Esta matriz atua como um mapa de características. Além disso, graças a essa operação de convolução, a imagem fica menor, reduzindo a sua resolução e, portanto, facilitando seu processamento. Apesar de alguns pixels serem desconsiderados, o importante é a detecção das características principais de interesse, já que o mapa de características pode preservá-las. Em sua camada de convolução, uma rede neural possui vários núcleos de

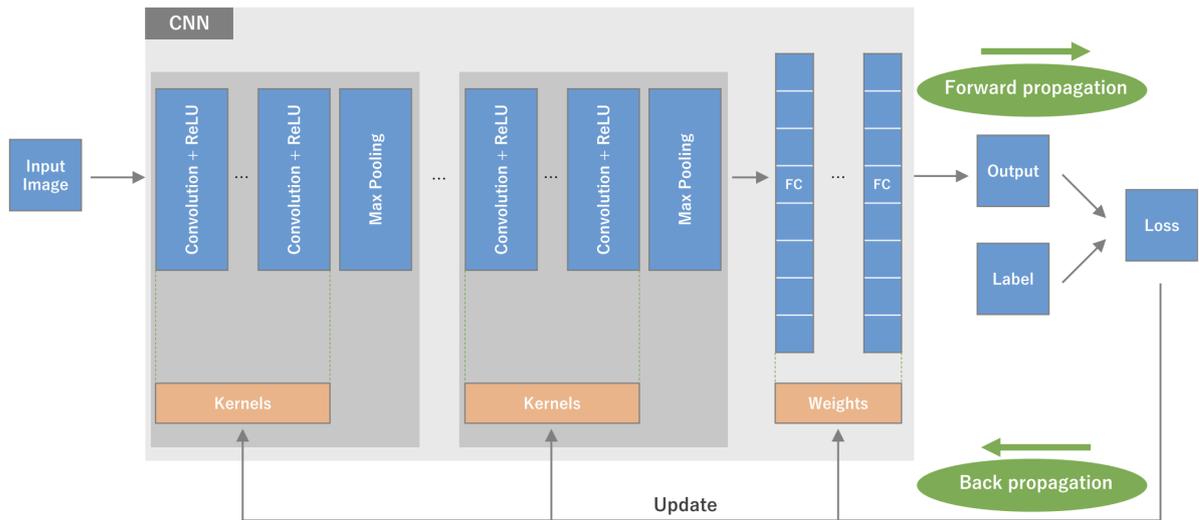


Figura 2.5: Visão geral de uma arquitetura genérica de rede neural convolucional e do processo de treinamento (YAMASHITA et al., 2018).

filtro responsáveis para detecção de características.

Além de camadas de convolução, há também a etapa de *pooling* em uma CNN. Esta etapa é aplicada em mapas de características, onde uma região de vizinhança é reduzida a um único valor, resultando na redução do número de parâmetros. Além de contribuir para a eficiência computacional, o *pooling* também tem o benefício de aumentar a capacidade de generalização da rede, tornando-a menos sensível a variações menores nos dados de entrada.

Existem vários tipos de *pooling*. Um exemplo é o *Max Pooling*, onde o valor máximo é selecionado em cada região de vizinhança.

Ademais, há a etapa de achatamento, também conhecida como *flattening*, que desempenha o papel de transformar a matriz após a extração de características em um formato vetorial. Este vetor é, então, utilizado como entrada para a camada densa da CNN. Na camada densa, a rede neural aprende a identificar os pesos corretos para assegurar a precisão da previsão final.

Por fim, na etapa de retropropagação, também conhecida como *backpropagation*, utiliza-se a retropropagação do erro para atualizar os pesos da rede neural. Ele ajusta os pesos dos filtros convolucionais e das camadas totalmente conectadas de modo a minimizar a diferença entre as previsões da rede e as saídas esperadas. O algoritmo calcula o gradiente da função de perda em relação aos pesos da rede, propagando-o a partir da ca-

mada de saída até a camada de entrada, atualizando assim os pesos de cada camada. Na Figura 2.5 pode ser observada a arquitetura de uma CNN e todas as etapas mencionadas anteriormente. (YAMASHITA et al., 2018).

## 2.5 Considerações Finais

Nas seções anteriores, foram apresentados conceitos importantes para o entendimento do trabalho. Cada conceito aborda os seguintes tópicos: Morfologia dentária e sua relação com a idade, Tomografia Computadorizada de Feixe Cônico, Redes Neurais e Redes Neurais Convolucionais.

Em suma, a estrutura dentária é composta por três camadas principais: esmalte, dentina e polpa. Essas camadas apresentam alterações ao longo do tempo que podem ser usadas para estimar a idade em investigações forenses. A dentina e a polpa são áreas que têm se mostrado promissoras para essa finalidade.

Além disso, foi visto como a TCFC é uma tecnologia que oferece imagens tridimensionais de alta precisão e resolução, com menor exposição à radiação e tempo de varredura rápido em comparação com a TC convencional. Essas vantagens fizeram com que a TCFC fosse introduzida rapidamente no campo da odontologia, especialmente na odontologia forense, onde tem sido utilizada para identificar a idade de um indivíduo de forma não invasiva. Com isso, é possível afirmar que a TCFC é uma importante aliada na área da odontologia, proporcionando benefícios tanto para profissionais quanto para pacientes.

Com relação às redes neurais artificiais foi apontado como são baseadas em neurônios capazes de aprender a partir de dados e realizar tarefas complexas de aprendizado de máquina. Além disso, também foi mencionado como as CNNs podem ser utilizadas em tarefas de visão computacional, a partir da extração de características das imagens.

## 3 Trabalhos Relacionados

Este capítulo tem como foco apresentar trabalhos relacionados ao tema desta monografia. Os trabalhos a seguir têm como objetivo a estimativa da idade dentária por meio de imagens radiográficas, seja por meio de radiografias panorâmicas ou por meio da TCFC. Além disso, eles buscam explorar o potencial do aprendizado de máquina nesse contexto. Para isso, fazem uso das CNNs para extrair características relevantes das radiografias e obter estimativas da idade dos indivíduos.

Desse modo, essa abordagem visa superar as limitações dos métodos convencionais, que dependem de conhecimento especializado e são subjetivos, proporcionando uma solução automatizada para a determinação da idade dentária.

Logo, trabalhos como os de Kim et al. (2021) e Zheng et al. (2020) são capazes de mostrar como a câmara pulpar do dente é uma região capaz de fornecer características importantes para a determinação da idade, sendo que ela é a parte dentária foco do presente trabalho. Ademais, outros estudos como os de Han et al. (2022) e Vila-Blanco et al. (2020) também retratam como o AP é capaz de se sair melhor na estimativa de idade em comparação com os métodos manuais, estes realizados por especialistas.

Para cada um dos trabalhos relacionados são apresentadas as seguintes informações: descrição do projeto, resultados encontrados, pontos positivos e negativos. Adicionalmente, ao final, é feita uma comparação entre os estudos.

### 3.1 Precisão na estimativa de idade por meio de aprendizado profundo: Interferência humana ou autonomia?

Até o momento, a partir de todas as informações que já foram apresentadas no presente trabalho, sabe-se como a determinação da idade pode ser benéfica em diversos contextos como: a área forense para fins de investigação, no diagnóstico de condições médicas e no

### 3.1 Precisão na estimativa de idade por meio de aprendizado profundo: Interferência humana ou autonomia?29

planejamento de tratamentos em ortodontia e pediatria. Com isso em mente, o trabalho de Han et al. (2022) é capaz de mostrar como métodos convencionais de estimativa de idade dentária são dependentes de especialistas e subjetivos, resultando em uma abordagem mais demorada. Além disso, também é apresentado como essa questão pode ser facilmente solucionada por meio da aplicação de técnicas de aprendizado de máquina, que oferecem uma abordagem mais eficiente e precisa.

Para isso, foram analisados três modelos para a estimativa de idade dos indivíduos. Além disso, todas as análises foram feitas a partir de imagens de radiografia panorâmica. No primeiro modelo, denominado “*Manual dental age estimation model - MDAE model*”, é realizada uma análise do estágio de desenvolvimento dos dentes por profissionais com alto conhecimento sobre o assunto. Após a anotação dos estágios dentários, eles são aplicados em uma fórmula de regressão linear múltipla para estimar a idade.

Os outros dois modelos utilizam-se de aprendizado profundo para auxiliar nas previsões, utilizando-se de redes neurais convolucionais para este fim. Desse modo, o segundo modelo, nomeado “*Automated dental stage evaluation model - ADSE model*”, ainda foi baseado nos estágios de desenvolvimento dentário definidos manualmente pelos profissionais. Este modelo teve o papel de analisar cada dente e, assim como os especialistas, determinar seu estágio de desenvolvimento de forma automática. Depois, os valores obtidos eram aplicados na mesma fórmula de regressão linear múltipla para a idade poder ser calculada.

Por fim, o último e terceiro modelo, intitulado “*End-to-end fully automated dental age estimation model - ADAE model*”, permite que a CNN extraia de forma autônoma potenciais características das imagens para estimativa de idade dentária. Uma visão geral dos modelos pode ser vista na Figura 3.1.

Como resultados, a principal métrica utilizada para comparação entre os modelos foi o *Mean Absolute Error* (MAE). A faixa de classificação das idades poderia variar de 5 até 25 anos, com incrementos de 1 ano em cada categoria. Com relação ao modelo MDAE, para as imagens no conjunto de teste com indivíduos entre 5 e 25 anos, o MAE foi de 1,67 anos. Já para o modelo ADSE, o MAE foi de 1,63 anos para o mesmo conjunto, indicando que este é capaz de realizar estimativas como as feitas pelos especialistas.

### 3.1 Precisão na estimativa de idade por meio de aprendizado profundo: Interferência humana ou autonomia?3

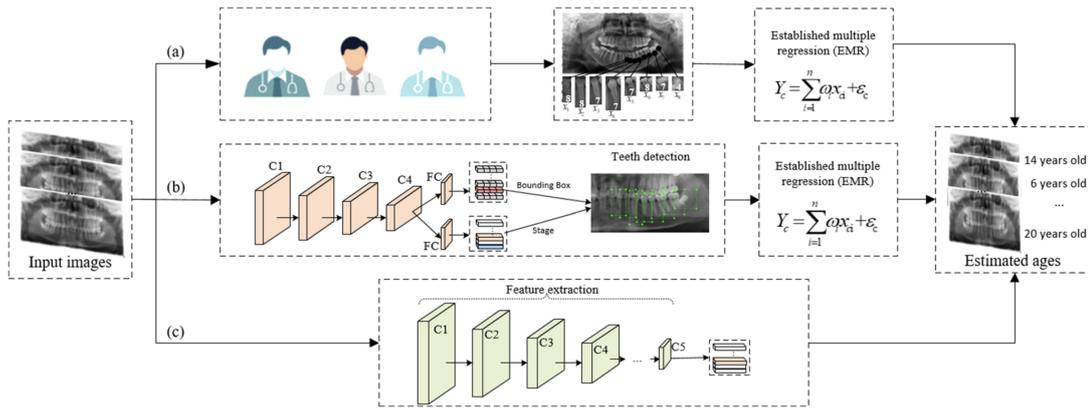


Figura 3.1: Visão geral dos modelos. (a) MDAE, (b) ADSE, (c) ADAE. (HAN et al., 2022).

Ademais, o modelo ADAE sem intervenção humana foi o que obteve os melhores resultados, aumentando consideravelmente a acurácia na previsão da idade dos indivíduos. Seu MAE foi de 0,83 anos para o conjunto de teste. De forma a indicar que ele foi capaz de extrair características mais apropriadas para a estimativa de idade analisando a imagem de toda a arcada dentária na radiografia panorâmica, ao invés de apenas analisar o estágio de desenvolvimento dentário. Uma comparação do MAE de cada um dos modelos pode ser vista a partir do gráfico na Figura 3.2.

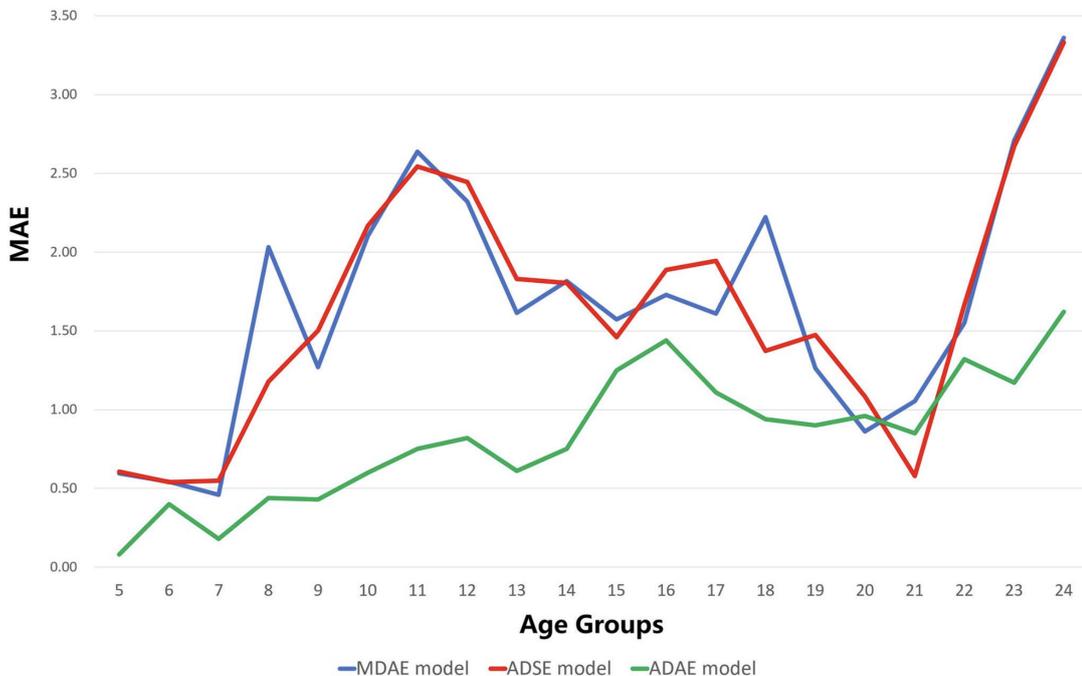


Figura 3.2: Comparação entre os valores de MAE para cada um dos modelos (HAN et al., 2022).

Logo, tem-se como pontos positivos da pesquisa de Han et al. (2022) o uso de CNN's para a estimativa da idade dos indivíduos e como elas podem ser bastante úteis nesse processo. Como pontos negativos, tem-se que a faixa de idade analisada não abrange indivíduos acima de 24 anos de idade.

## **3.2 Determinação do grupo etário utilizando inteligência artificial em imagens do primeiro molar**

Como segundo trabalho relacionado, o artigo de Kim et al. (2021) tem o objetivo de elaborar um sistema de diagnóstico baseado em inteligência artificial para a estimativa de grupos de idade dos indivíduos. Para isso, trabalhou-se com aprendizado profundo a partir de uma rede neural convolucional em que imagens de raios-X dos primeiros dentes molares são analisadas. Essas imagens dos dentes molares foram obtidas por meio de radiografia panorâmica.

Assim, para cada paciente, imagens dos seus 4 primeiros dentes molares foram obtidas. Ademais, neste estudo foram avaliados dois modelos para a estimativa de idade, sendo que a diferença para esses dois modelos está relacionada à divisão dos grupos de idades.

Primeiramente, os participantes foram divididos em 3 faixas etárias: crianças e adolescentes (0 a 19 anos), adultos jovens (20 a 49 anos) e idosos (idades acima de 50 anos). Desse modo, o primeiro modelo é responsável por estimar se um determinado indivíduo se encontra em um desses grupos. O método de estimativa é o mesmo para os dois modelos: para cada um dos 4 dentes é dada uma classificação da faixa etária, a classificação que recebe o maior número de votos é selecionada como sendo o grupo do indivíduo. Este método é ilustrado na Figura 3.3.

Como mencionado anteriormente, o segundo modelo difere apenas nas faixas etárias dos pacientes. Nesse caso, o grupo destinado aos pacientes de 20 a 49 anos foi dividido em 3 partes: 20-29 anos, 30-39 anos e 40-49 anos. Para fins de comparação, os resultados de predição para as cinco faixas etárias foram agrupados em três faixas etárias. Esse agrupamento pode ser observado na Figura 3.4.

### 3.2 Determinação do grupo etário utilizando inteligência artificial em imagens do primeiro molar<sup>32</sup>

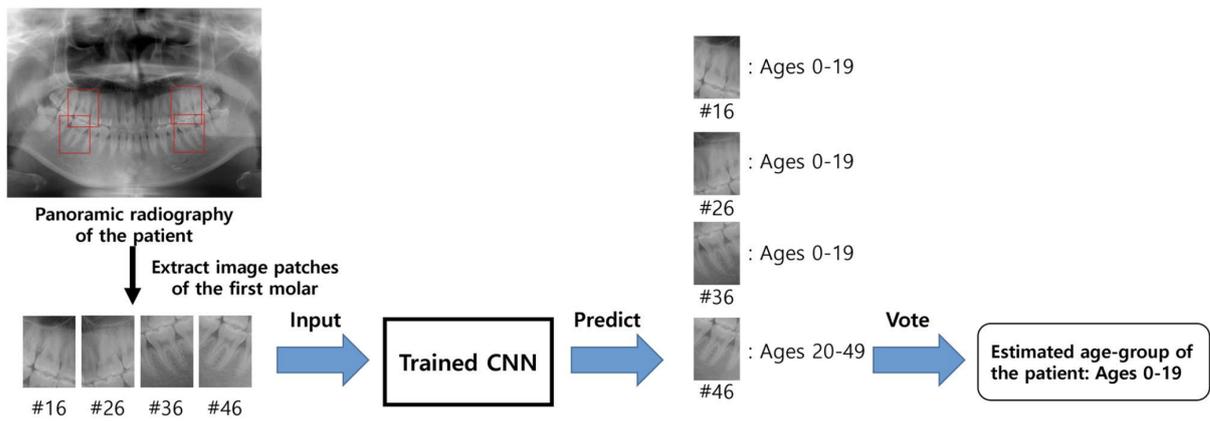


Figura 3.3: Classificação de uma radiografia panorâmica usando a CNN para prever a faixa etária do paciente (KIM et al., 2021).

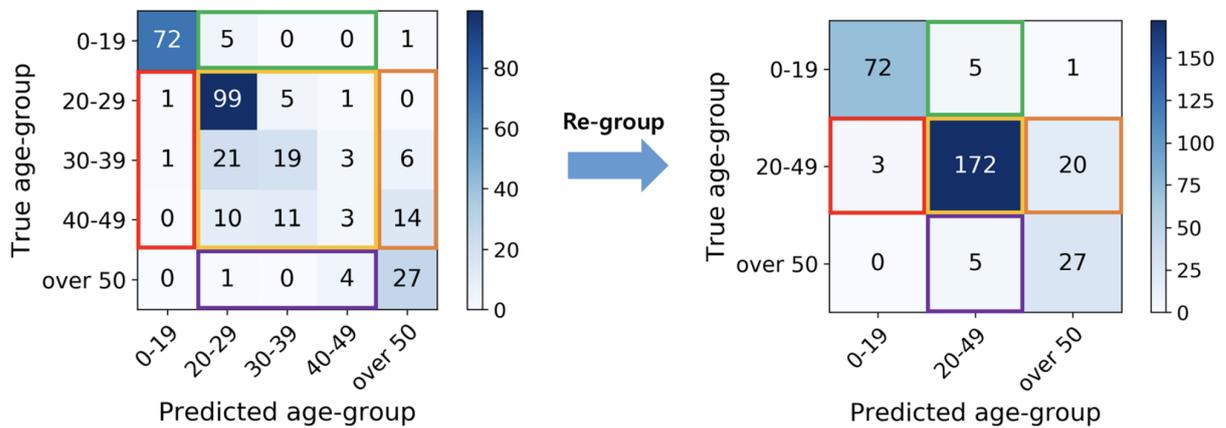


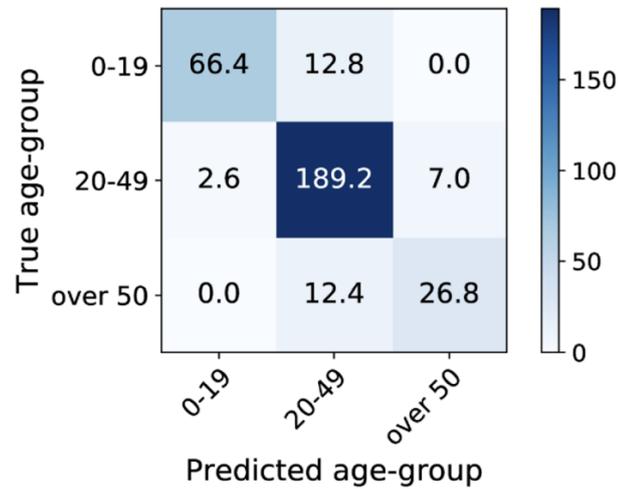
Figura 3.4: Agrupamento do resultado da previsão da faixa etária de cinco faixas etárias para as três faixas etárias (KIM et al., 2021).

Como resultados, curvas ROC (*Receiver Operating Characteristic*) e AUC (*Area Under the Curve*) foram usadas para investigar o desempenho do sistema. A precisão geral ao analisar os dentes de forma individual varia entre 87,04 e 88,33%. Porém, a precisão não diferiu significativamente de acordo com a localização do primeiro molar ou o número de faixas etárias usadas.

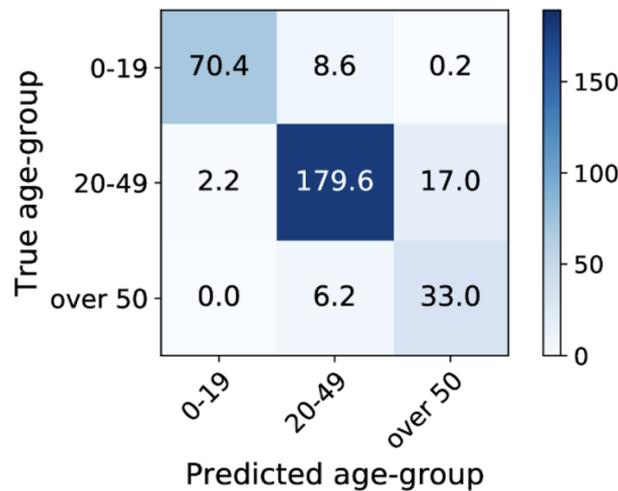
Além disso, as precisões da previsão de idade por paciente para os três grupos de idade e os cinco grupos de idade foram  $89,05 \pm 0,68\%$  e  $89,21 \pm 0,54\%$ , respectivamente. Também, não foram observadas diferenças significativas nos resultados dos dois modelos. Essas diferenças podem ser observadas comparando as matrizes de confusão da Figura 3.5. Também, foi aplicada a técnica de visualização Grad-CAM (*Gradient-weighted Class Activation Mapping*), permitindo entender quais características das imagens foram mais

### 3.2 Determinação do grupo etário utilizando inteligência artificial em imagens do primeiro molar<sup>33</sup>

relevantes para a classificação feita pelos modelos. E, determinou-se que, entre outros aspectos, a polpa dentária é uma característica visual importante na determinação da idade.



(a) Previsão para 3 faixas etárias



(b) Previsão para 5 faixas etárias

Figura 3.5: Matriz de confusão dos resultados da estimativa de faixa etária (KIM et al., 2021).

Logo, tem-se como pontos positivos deste projeto o uso de CNNs para a estimativa da idade dos indivíduos. Adicionalmente, foi apresentado como a câmara pulpar pôde ajudar e influenciar na classificação de um indivíduo em uma faixa de idade. Como pontos negativos, tem-se o número restrito de indivíduos que foram analisados. Ainda mais, outra limitação é que os dados são tendenciosos para adultos jovens (20 a 49 anos), já que o número de imagens utilizadas de indivíduos nesse grupo é muito maior em comparação às outras faixas etárias.

### 3.3 Estimativa de idade com base na segmentação da câmara pulpar em 3D

Como terceiro trabalho relacionado, o artigo de Zheng et al. (2020) tem os seguintes objetivos: desenvolver um método de segmentação automática da câmara pulpar dos primeiros dentes molares a partir de imagens 3D da TCFC e estimar as idades dos indivíduos através dos volumes pulpares calculados.

Assim, um pré-requisito para que o volume da câmara pulpar seja extraído é alcançar uma segmentação precisa das imagens 3D dos dentes. Essas imagens foram obtidas nos planos axial, sagital e coronal. Além disso, a polpa dentária foi segmentada manualmente para que o modelo de deep learning para segmentação pudesse ser treinado.

Para refinar a segmentação realizada pelo modelo de aprendizado profundo, o método de *level set* (CHAN; VESE, 2001; LI et al., 2005) foi implementado logo após as máscaras de segmentação previstas pelo modelo. Sendo que esse método é utilizado para auxiliar na definição de fronteiras entre diferentes regiões em uma imagem, o que permite uma delimitação mais precisa da estrutura da câmara pulpar.

Desse modo, após a segmentação das imagens 3D, é possível realizar a estimativa da idade do indivíduo. Para isso, o volume da câmara pulpar dos primeiros molares é calculado a partir das máscaras de segmentação. Em seguida, um modelo de regressão linear para previsão da idade é utilizado. Uma visão geral do fluxo pode ser vista na Figura 3.6. Ademais, neste modelo, a idade atua como variável dependente e o volume da câmara pulpar do primeiro molar como variável independente. Ele pode ser observado abaixo, onde  $V$  é o volume da câmara pulpar.

$$\text{AGE} = 134,6 - 32,7 \cdot \ln(V) \quad (3.1)$$

Para a análise dos resultados várias métricas foram utilizadas: o *Wilcoxon signed-rank test* foi utilizado para comparar idades humanas verdadeiras e estimadas em dados de teste com  $p > 0,05$ , sendo que um valor de  $p$  maior que 0,05 indica que não há diferença significativa. Também foram utilizados o erro médio absoluto (MAE), a raiz quadrada do erro médio (RMSE) e o coeficiente de correlação ( $r$ ) entre as idades verdadeira e estimada.

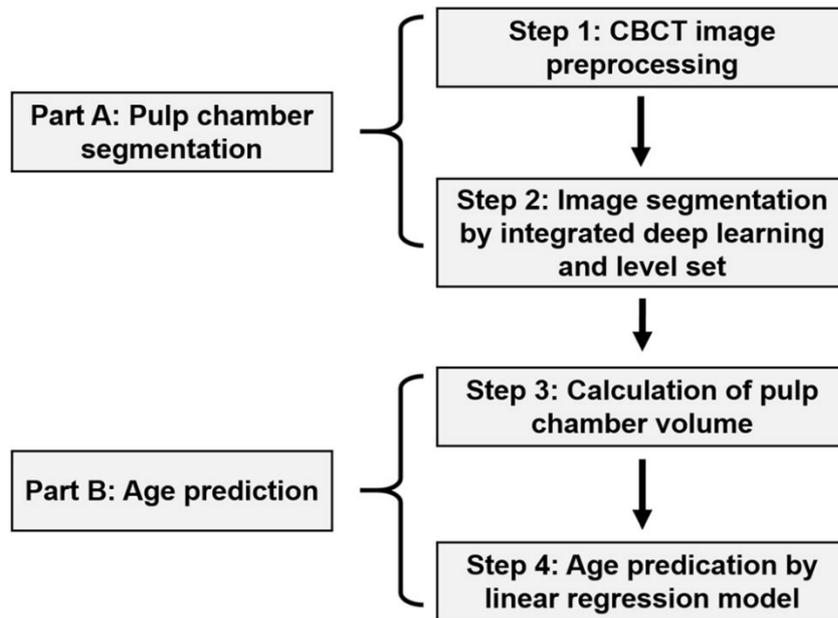
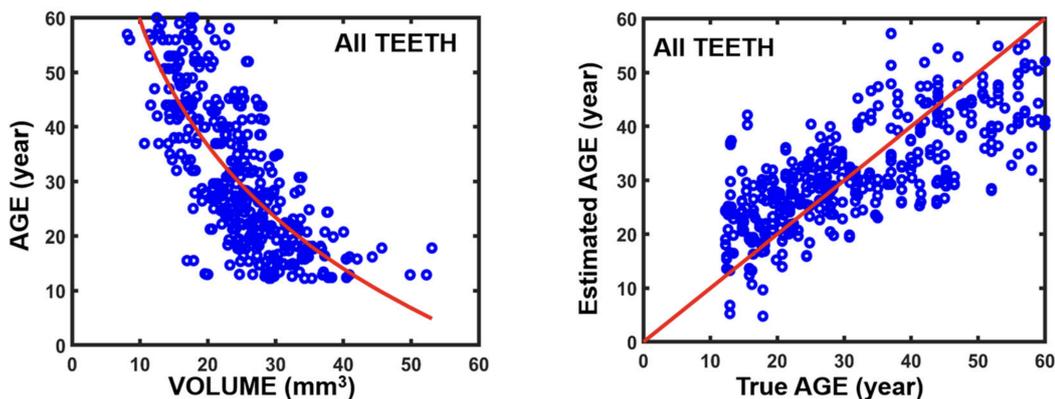


Figura 3.6: Fluxo de trabalho para a análise de imagens da TCFC (ZHENG et al., 2020).

Como resultados, não houve uma diferença significativa entre as idades humanas estimadas e verdadeiras. Para todas as métricas avaliadas, foram obtidos os seguintes valores: coeficiente de correlação  $r = 0,74$ ,  $p = 0,57$ , MAE = 7,26 anos e RMSE = 9,00 anos. Na Figura 3.7 é possível observar como o volume da câmara pulpar diminui com o aumento da idade, além de uma comparação entre as idades estimadas e a idade real dos indivíduos.



(a) Relação entre idade e volume da câmara pulpar

(b) Relação entre idade estimada e idade verdadeira

Figura 3.7: Gráficos de dispersão para todos os dentes de primeiros molares em dados de teste (ZHENG et al., 2020).

Por fim, tem-se como pontos positivos deste projeto a análise de imagens a partir da TCFC, além do foco na câmara pulpar do dente para a estimativa de idade, ambos

também alvo de interesse do presente trabalho. Como pontos negativos, tem-se o número reduzido de dados presente no *dataset* e o fato de que não foi utilizada uma CNN para classificação das idades dos indivíduos, mas sim para auxiliar na segmentação da câmara pulpar do dente.

### 3.4 Redes neurais profundas para estimativa da idade a partir de radiografia panorâmica

Como quarto trabalho relacionado, o artigo de Vila-Blanco et al. (2020) tem os seguintes objetivos: propor dois métodos para estimar a idade cronológica de um indivíduo a partir de imagens de radiografias panorâmicas. Os métodos propostos fazem uso de aprendizado profundo e utilizam redes neurais convolucionais.

Para isso, o primeiro método (DANet) consiste em uma CNN com um caminho sequencial para prever a idade. Já o segundo método (DASNet) adiciona uma segunda CNN para prever o sexo do indivíduo, utilizando das características particulares de cada sexo para promover uma melhora na classificação da idade.

Ambos os métodos foram testados em pessoas em que a idade variava de 4,5 até 89,2 anos. Além disso, também foram incluídas no conjunto de dados imagens radiológicas com uma qualidade ruim ou com características incomuns como cáries, próteses, implantes, entre outros, o que deixa a amostra mais próxima da realidade.

A arquitetura das duas redes desenvolvidas podem ser visualizadas na Figura 3.8. O modelo DANet (*Dental Age Net*) é observado pelos itens em vermelho. Já a segunda abordagem, o modelo DASNet (*Dental Age and Sex Net*), adicionou os itens em azul e possui uma rede para classificar as imagens de acordo com o gênero do indivíduo. Porém, o objetivo desta nova rede não foi a classificação do gênero em si, mas sim enviar as características do seu aprendizado com relação ao sexo para a rede que estava realizando as previsões da idade.

Ademais, além do conjunto de dados original chamado de  $S_{\text{all}}$ , outros conjuntos de dados foram gerados. Nesse caso, os novos conjuntos criados continham indivíduos em que a idade era menor ou igual aos seguintes valores: 40 ( $S_{40}$ ), 30 ( $S_{30}$ ), 25 ( $S_{25}$ ), 20  $S_{20}$  e

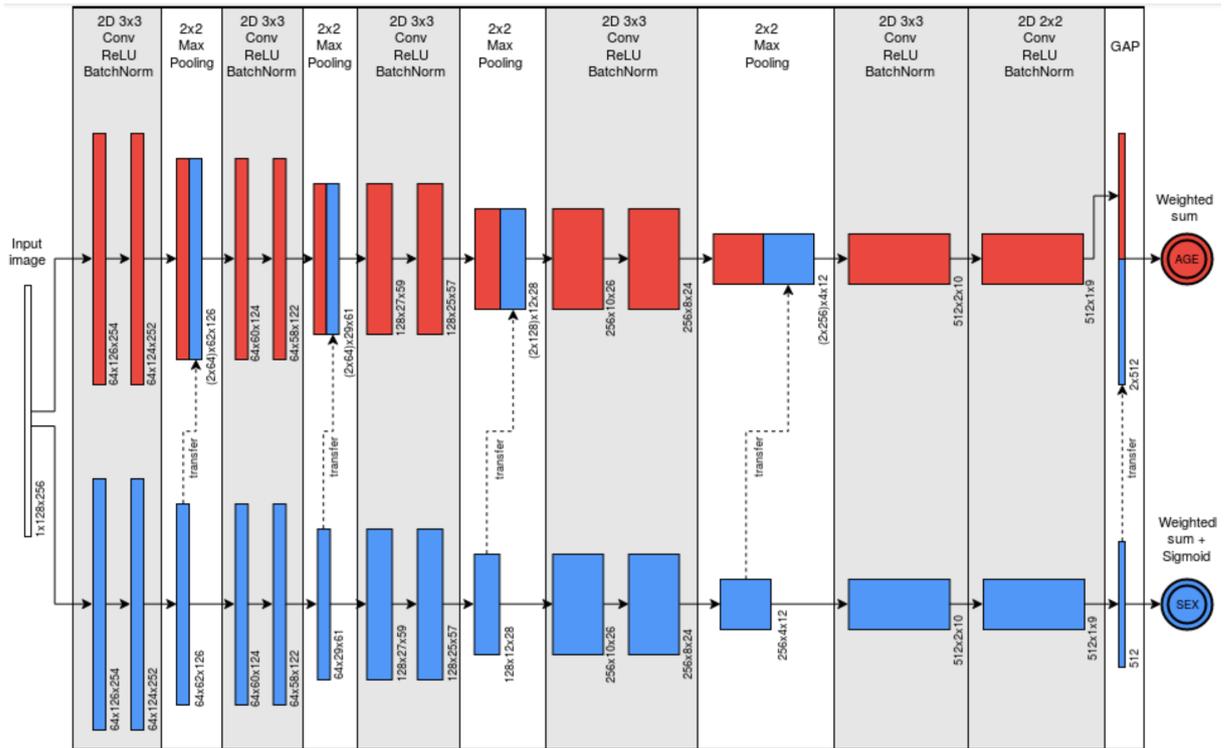


Figura 3.8: Arquitetura das duas redes propostas. Em vermelho o caminho para classificação da idade e em azul o caminho para previsão do sexo (VILA-BLANCO et al., 2020).

15 anos ( $S_{15}$ ). Assim, os modelos foram treinados e testados utilizando-se cada um desses agrupamentos.

Como resultados, o coeficiente  $R^2$  foi usado para medir a eficácia do modelo em capturar os padrões e tendências presentes nos dados observados. Em geral, houve alta correlação entre as idades estimadas pelos modelos DANet e DASNet e as idades reais ( $R^2 = 0,87$  e  $0,90$ , respectivamente).

Além disso, o DASNet forneceu melhores resultados do que o DANet no conjunto de dados  $S_{\text{all}}$ , apresentando um erro absoluto médio (AE) de cerca de 2 anos e 10 meses de idade e uma mediana de 0,12 anos. Na Figura 3.10, gerada pelas classificações realizadas pelo DASNet, pode ser visualizado como o nível de correlação foi bem maior para os indivíduos abaixo dos 20 anos de idade.

Adicionalmente, ao avaliar e comparar ambas as redes em relação aos conjuntos de dados reduzidos, os valores do AE diminuíram conforme a idade real dos sujeitos caiu. Na Figura 3.10, sendo esses valores menores no DASNet do que no DANet em todos os casos.

Além do mais, foi feita uma comparação do modelo DASNet com outros métodos

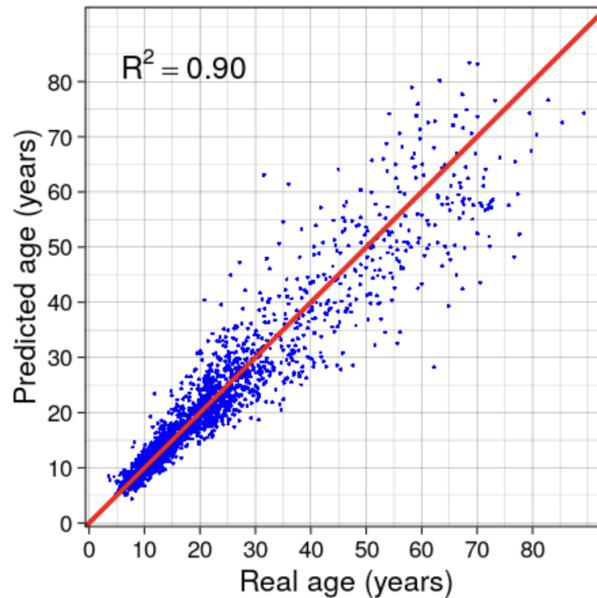


Figura 3.9: Idade real comparada às previsões do modelo DASNet avaliadas para todo o conjunto de dados,  $S_{\text{all}}$  (VILA-BLANCO et al., 2020).

clássicos da literatura em que a estimativa da idade é feita de forma manual Figura 3.11. Dessa forma, foram utilizados os conjuntos de dados dos trabalhos em que esses métodos foram utilizados como forma de comparação. E, considerando pelo menos uma das métricas possíveis de análise, o procedimento DASNet produziu resultados melhores que as abordagens manuais em todas as comparações.

Por fim, tem-se como pontos positivos deste projeto o uso de CNNs para a classificação da idade utilizando imagens radiográficas, além do uso de uma faixa etária extensa para os indivíduos, assim como a relevância que as características relacionadas ao sexo tiveram para os resultados. Como pontos negativos, tem-se a falta de foco em uma determinada região do dente para a estimativa, como a câmara pulpar, por exemplo.

### 3.5 Classificação de idade baseada em radiografia panorâmica através de método manual e CNN

Como quinto trabalho relacionado, o artigo de Guo et al. (2021) teve como objetivo comparar uma abordagem manual que utiliza modelos de regressão linear e uma rede neural convolucional de ponta a ponta para classificar a idade dentária com base na

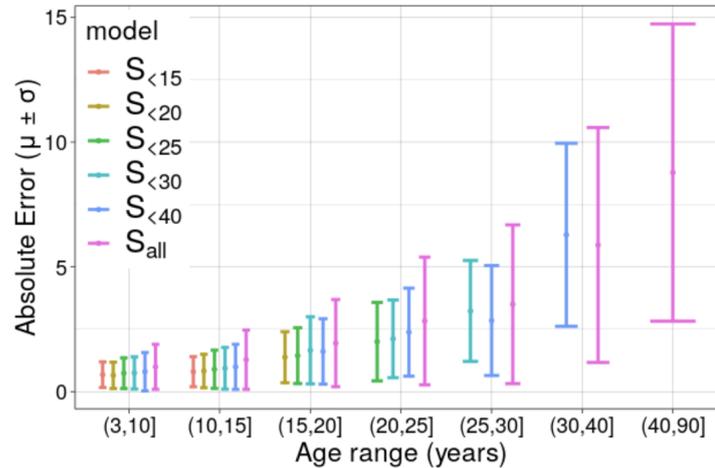


Figura 3.10: Erro absoluto do DASNet na estimativa de idade cronológica ao treinar com conjuntos de dados reduzidos e avaliar o desempenho para diferentes faixas etárias (VILA-BLANCO et al., 2020).

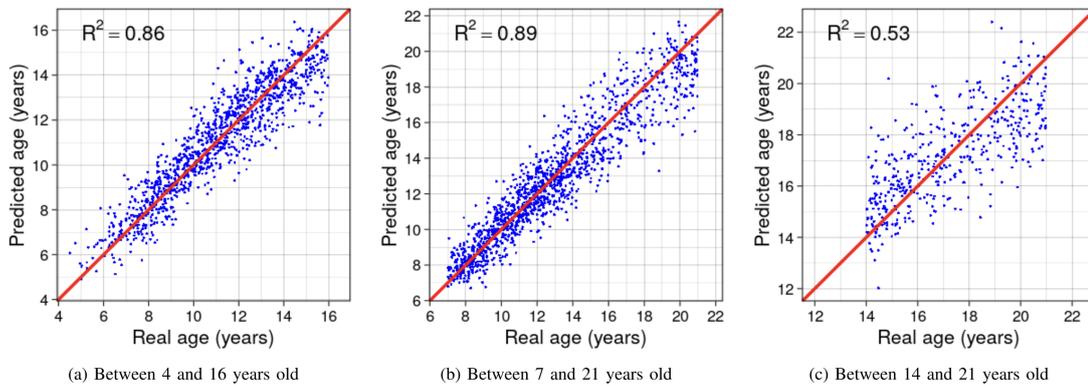


Figura 3.11: Idade real comparada às previsões do DASNet para alguns conjuntos de dados reduzidos gerados para comparar com os métodos manuais de estimativa (VILA-BLANCO et al., 2020).

análise de radiografias panorâmicas.

Para isso, um total de 10,257 imagens radiográficas foram obtidas e analisadas, sendo 4,579 homens e 5,678 mulheres com idades entre 5 e 24 anos. Assim, dividiu-se essas imagens em 20 grupos de faixas etárias. Por exemplo, o grupo 15 possui os pacientes com idades variando de 15,00 a 15,99 anos. As imagens foram bem distribuídas para cada grupo, cerca de 500 imagens para cada, e foram utilizadas para o treinamento da CNN.

Além disso, para o método manual, os estágios de desenvolvimento dentário foram estimados pelos especialistas com base no método de Demirjian. Ademais, para testar este método, foram criados modelos de regressão linear para os seguintes grupos de idade: 14, 16 e 18 anos. E, para fins comparativos, o método automático que faz uso da rede

neural convolucional também será testado nesses grupos de idade.

Como resultados, tem-se que a precisão dos métodos manuais é estável, com diferenças relativamente pequenas entre as faixas etárias de 14, 16 e 18 anos, sendo a maioria delas inferiores a 2%. Nesse caso, a acurácia do método manual foi de 92,5%, 91,3%, e 91,8% para essas faixas de idade, respectivamente.

Além do mais, com relação à CNN, um desempenho superior foi alcançado. A precisão para as faixas etárias de 14 e 16 anos ultrapassou 95%, e a precisão para a faixa etária de 18 anos alcançou 93,3%.

A fim de explorar o motivo de um melhor desempenho de classificação para o método com aprendizado de máquina, a técnica de visualização Grad-CAM foi aplicada. Desse modo, ela revelou um foco em características de baixa densidade nas imagens, como a cavidade pulpar, a membrana periodontal e as áreas entre os dentes. Logo, a CNN foi capaz de estabelecer uma correlação complexa entre múltiplas características e a idade cronológica, diferente do que é analisado pelo método manual.

Portanto, como pontos positivos, destaca-se o uso bem-sucedido de CNNs para classificação da idade dentária, proporcionando uma abordagem automatizada e precisa, segundo os autores. Como pontos negativos, tem-se a restrição da análise a uma população chinesa, sem generalização para outras populações étnicas, além da limitação da faixa etária testada.

### **3.6 Avaliação de duas redes neurais convolucionais na estimativa de idade dentária**

Como sexto trabalho relacionado, o artigo de Wang et al. (2023) teve como objetivo comparar duas redes neurais convolucionais para estimar a idade dentária com base na análise de radiografias panorâmicas. Uma dessas redes é a VGG16 e a outra é a ResNet101.

Para isso, foram coletadas 9586 (4054 do sexo masculino e 5532 do sexo feminino) imagens radiográficas de crianças e adolescentes chineses com idades entre 6 e 20 anos. Além disso, o grupo de estudo foi separado em 15 grupos de idade em um intervalo de 1 ano.

Ademais a arquitetura da VGG16 pode ser observada na Figura 1. A ResNet101 pode ser visualizada a partir da Figura 2 e a unidade *bottleneck* à direita é um componente presente nas camadas de 1 à 4 para reduzir a complexidade computacional e melhorar o desempenho da rede.

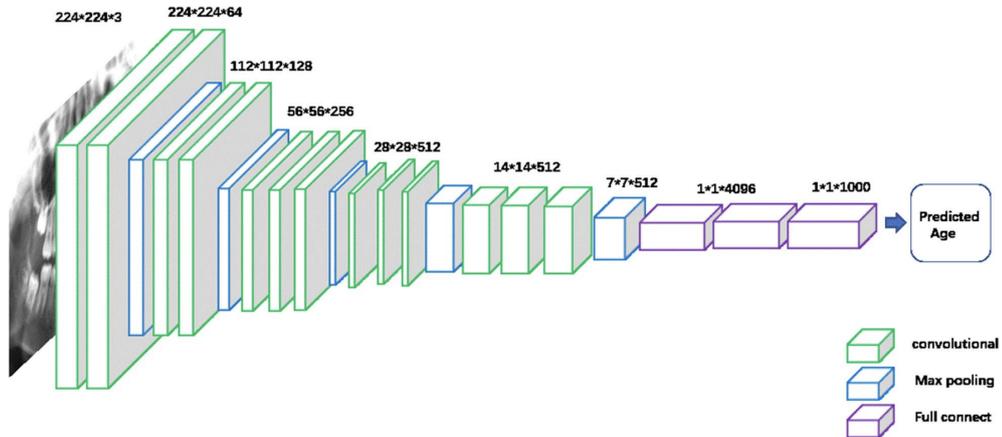


Figura 3.12: Estrutura básica da rede VGG16 (WANG et al., 2023).

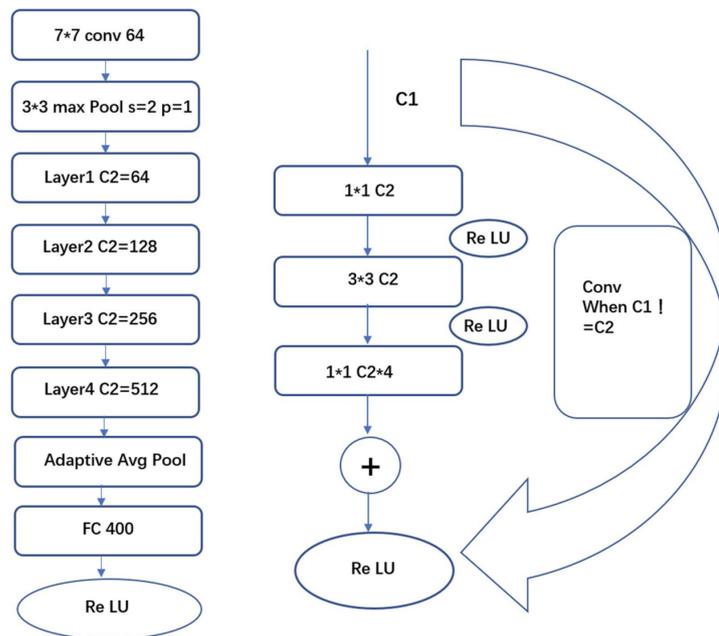


Figura 3.13: Diagrama com a estrutura básica da rede ResNet101 (WANG et al., 2023).

Como resultados, a acurácia, *recall*, precisão e *F-Measure* dos modelos foram utilizadas para avaliar a estimativa da idade. Um limiar de idade também foi empregado para avaliar os dois modelos de CNN.

Além do mais, observa-se na Tabela 3.1 que o modelo VGG16 obteve melhor

desempenho para problemas de classificação múltipla, com uma taxa de precisão de 79% para uma diferença de idade entre 1 e 1,99 anos. Nota-se também que a partir de um limiar de idade entre 2 e 2,99 anos o modelo obteve uma acurácia acima de 90%.

Limiar de Idade	Acurácia	
	VGG16	ResNet101
<1	0,48	0,31
<2	0,79	0,61
<3	0,91	0,79
<4	0,95	0,88
<5	0,97	0,95
<6	0,99	0,98
<7	0,99	0,99
<8	0,99	0,99
<9	1	1

Tabela 3.1: Acurácia dos modelos para determinados limiares de idade (WANG et al., 2023).

Adicionalmente, o modelo ResNet101 apresentou uma taxa de acurácia de 61% para a mesma faixa etária, ligeiramente inferior ao modelo VGG16. Em termos de acurácia em diferentes faixas etárias e valores de *F-Measure*, o modelo VGG16 apresentou um desempenho melhor em relação ao modelo ResNet101.

Logo, como pontos positivos, destaca-se o uso bem-sucedido de CNNs para classificação da idade dentária em uma grande amostra de radiografias, além de um comparativo entre diferentes arquiteturas para realizar esse processo. Como pontos negativos, tem-se a restrição da análise a uma população específica do leste chinês, sem generalização para outras populações étnicas, além da limitação da faixa etária testada.

### 3.7 Considerações Finais

Nas seções anteriores, foram apresentados os trabalhos relacionados com o projeto desta monografia. O foco de cada uma das seções para cada projeto foi apresentar a descrição do projeto, resultados encontrados, seus pontos positivos e negativos. Além disso, pode-se observar como todos os estudos compartilham o objetivo comum de utilizar técnicas de aprendizado profundo para a estimativa da idade dos indivíduos a partir de imagens radiográficas.

Adicionalmente, com exceção do trabalho de Zheng et al. (2020), todos os outros utilizam um modelo de CNN para realizar a classificação da idade dentária. Também, pode-se notar que a maioria dos trabalhos recorrem às radiografias panorâmicas para a obtenção das imagens dentárias. O trabalho de Zheng et al. (2020) é o único que utiliza imagens a partir da TCFC, estas imagens são dos cortes coronal, sagital e axial do dente, para posterior cálculo do volume da câmara pulpar.

Ademais, o foco de análise dos trabalhos variou entre as arcadas dentárias dos indivíduos e os primeiros dentes molares, sendo diferente do foco desta monografia que são dentes incisivos. Percebe-se, também, que quanto maior a extensão de idade analisada pelos modelos, mais difícil foi obter resultados precisos na estimativa de idade.

Por fim, os resultados desses trabalhos evidenciam como as CNNs e o aprendizado de máquina podem ser aplicados para aprimorar a acurácia, objetividade e eficiência na estimativa de idade e análise de imagens dentárias. Além disso, o trabalho desta monografia apresenta alguns diferenciais com relação às outras pesquisas como: uso de imagens da TCFC, foco na câmara pulpar dos dentes incisivos e a análise de uma faixa de idade extensa com inclusão de indivíduos adultos. A tabela 3.2 apresenta as informações mais relevantes de cada trabalho relacionado, proporcionando uma visão geral dos principais aspectos abordados por eles. Além disso, é possível comparar as características desses trabalhos com estudo realizado nesta monografia.

Trabalho	Método	Tipo de Imagem	Tamanho do Conjunto de Dados	Imagem	Faixa de Idade	Tipo de Classificação	Métricas
Han et al. (2022)	CNNs	Radiografia Panorâmica	10257	Arcada Dentária	5-25 anos	Estimativa da Idade	MAE
Kim et al. (2021)	CNN, Level Set, LR	Radiografia Panorâmica	1586	Primeiros Molares	0-60+ anos	Faixas de Idade	Acurácia
Zheng et al. (2020)	CNNs	TCFC	180	Primeiros Molares	10-60 anos	Estimativa da Idade	<i>Dice, r</i>
Vila-Blanco et al. (2020)	CNNs	Radiografia Panorâmica	2289	Arcada Dentária	4,5-89,2 anos	Estimativa da Idade	AE
Guo et al. (2021)	CNNs	Radiografia Panorâmica	10257	Arcada Dentária	5-24 anos	Estimativa da Idade	Precisão
Wang et al. (2023)	CNNs	Radiografia Panorâmica	9586	Arcada Dentária	6-20 anos	Estimativa da Idade	Acurácia
Esta monografia	CNNs	TCFC	1110	Dentes Incisivos	18-60 anos	Faixas de Idade	Acurácia

Tabela 3.2: Comparação entre os trabalhos relacionados.

## 4 Método Proposto

Este capítulo apresenta detalhes do processo adotado neste trabalho para a estimativa de idade por meio de imagens de câmara pulpar utilizando modelos de aprendizado profundo. Assim, são apresentadas informações sobre o conjunto de dados, os diferentes tipos de modelos utilizados no trabalho e os critérios de avaliação.

As imagens de TCFC utilizadas neste projeto estão armazenadas no Serviço de Radiologia Odontológica da Faculdade de Odontologia da UFJF (Minas Gerais, Brasil). Todas as imagens são adquiridas pelo mesmo tomógrafo (I-Cat<sup>®</sup>, *Imaging Sciences International*, Hatfield, Pensilvânia, EUA), com o seguinte protocolo de aquisição: 120 kV, 8 mA, 26,9 s de tempo de rotação, espessura de corte de 0,25 mm e FOV (*Field of View*) mínimo de 7 x 13 cm. Esse estudo foi aprovado pelo Comitê de Ética em Pesquisa (CEP) da UFJF sob o parecer n.º 5.485.995/2022, presente no Anexo A desta monografia.

Estão incluídos nesta base de dados exames de TCFC de participantes da região da Zona da Mata Mineira, de ambos os sexos, com idade entre 18 e 60 anos. Assim, especialistas fazem a análise dos exames e adotam os critérios que consistem em: exames que possuem uma boa qualidade para avaliação e que apresentam algum dente incisivo superior e/ou inferior hígido e com formação radicular completa (ápice fechado). Ademais, são excluídos os dentes que estejam cariados, restaurados, com tratamento endodôntico, reabsorções radiculares e com qualquer anomalia dentária ou lesão associada.

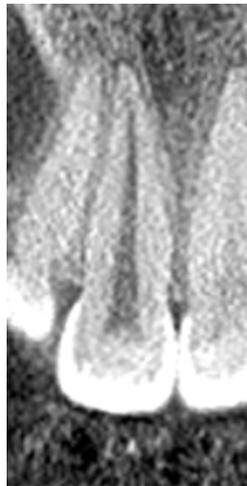
A estimativa da idade dos indivíduos pelos modelos de classificação será realizada a partir de faixas de idade de 5 em 5 anos. Muitos trabalhos definem faixas etárias para os estudos de estimativa de idade, alguns de 10 em 10 anos (KIM et al., 2021), outros de 5 em 5 anos (STAR et al., 2011; VOSSOUGH; MOVAHHEDIAN; GHAFUORI, 2021), além de estudos que consideram de 1 em 1 ano (HAN et al., 2022; WANG et al., 2023; VILA-BLANCO et al., 2020).

Desse modo, na base de dados é feita uma distribuição homogênea dos dentes incluídos em relação ao sexo (masculino e feminino). Com relação às faixas de idade, elas compõem um total de 8 faixas e são distribuídas da seguinte forma: Grupo 1, 18 a 22

anos; Grupo 2, 23 a 27 anos; Grupo 3, 28 a 32 anos; Grupo 4, 33 a 37 anos; Grupo 5, 38 a 42 anos; Grupo 6, 43 a 47 anos; Grupo 7, 48 a 52 anos; Grupo 8, 53 a 60 anos.

## 4.1 Descrição dos Conjuntos de Dados

As CNNs desenvolvidas são alimentadas com 3 tipos de imagens: cortes coronais, cortes sagitais e um último tipo de imagem, denominado concatenado, que faz uma união do corte coronal e do corte sagital. Assim, 3 conjuntos de dados são gerados e utilizados para cada um dos modelos de classificação em uma das 8 faixas etárias possíveis: Conjunto de Dados 1, análise de cortes coronais de indivíduos de ambos os sexos; Conjunto de Dados 2, análise de cortes sagitais; Conjunto de Dados 3, análise de cortes concatenados.



(a) Corte coronal



(b) Corte sagital

Figura 4.1: Imagens ilustrativas de cortes do dente incisivo central superior de um mesmo indivíduo

Um exemplo de corte coronal de um dente incisivo central superior pode ser visualizado na Figura 4.1a. Como esta imagem é de um corte coronal do dente, ela será utilizada pelo primeiro experimento correspondente ao seu respectivo conjunto de dados. Nessas imagens, pode-se observar a estrutura interna e mais escura do dente, representando a sua câmara pulpar.

Além disso, pode ser observado o corte sagital do mesmo dente e do mesmo indivíduo na Figura 4.1b. Ademais, constituindo o terceiro conjunto de dados, é possível observar na Figura 4.2 em que foi feita a junção dos cortes coronal e sagital, vistos anteriormente, em uma única imagem.



Figura 4.2: Imagem ilustrativa da junção dos cortes coronal e sagital em uma única imagem (corte concatenado).

Para a formação dos conjuntos de dados, foram obtidas um total de 1110 imagens, de modo que 555 imagens são de corte Coronal e as outras 555 são de corte Sagital. Cada par de imagens de cada corte corresponde a um mesmo indivíduo. Conseqüentemente, é possível gerar as 555 imagens com o tipo de corte concatenado.

A partir da Figura 4.3, pode ser observada como está a distribuição das classes do conjunto de dados para cada um dos cortes. Pode-se notar que há um desbalanceamento da base de dados, sendo que existe uma maior concentração de imagens nos grupos 1 e 2. Outros grupos possuem uma quantidade menor de dados. Por exemplo, o grupo 4 possui menos da metade do número de imagens do grupo 2 que é o que possui a maior quantidade de instâncias na base.

Por fim, para cada conjunto de dados, as imagens são divididas em conjuntos de dados de treinamento (70% das imagens), validação (15% das imagens) e teste (15% das imagens), com o intuito de realizar a geração dos modelos. As imagens dos cortes Coronal e Sagital possuem uma resolução de 150x300 pixels, enquanto as imagens do corte concatenado possuem uma resolução de 300x300 pixels.

## 4.2 Arquitetura dos Modelos

Com relação a arquitetura de CNN escolhida para a geração dos resultados que serão apresentados neste trabalho, treinamentos preliminares foram realizados com diferentes

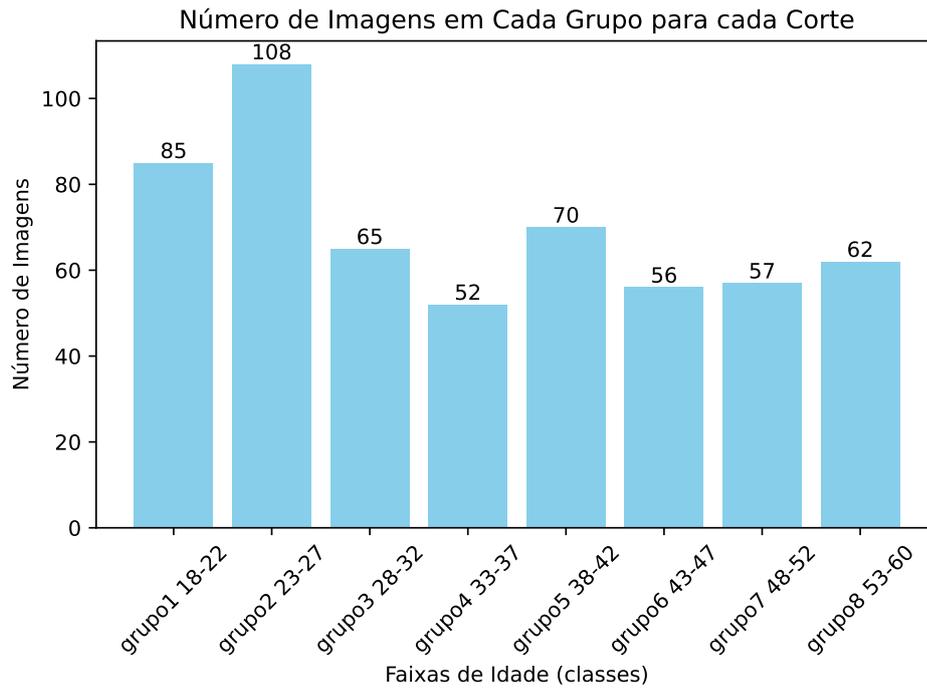


Figura 4.3: Gráfico contendo a distribuição das 555 imagens de cada corte para cada uma das 8 classes.

tipos de arquiteturas. Neste caso, os modelos gerados de cada uma delas eram utilizados na predição das 8 possíveis classes, representando as faixas de idade de um indivíduo, depois eram medidas as acurácias, precisões, assim como valores de *Recall* e *F-Measure* dos modelos para conjuntos de testes.

Uma dessas arquiteturas foi a *ResNet152* (HE et al., 2015). Essa arquitetura também foi utilizada no artigo de Kim et al. (2021), descrito na Seção 3.2. Outra arquitetura, também testada preliminarmente, foi a CoAtNet (DAI et al., 2021), desde sua versão 0 até a versão 4. Porém, os modelos gerados por essas arquiteturas não tiveram resultados satisfatórios em seus testes, sendo assim descartados.

Além disso, uma outra arquitetura que estava presente nos testes preliminares e que obteve resultados melhores em comparação às demais foi o YOLOv8 (JOCHER; CHAURASIA; QIU, 2023). Esta arquitetura foi lançada em janeiro de 2023 pela *Ultralytics*. Além disso, ela apresenta uma variedade de modelos, cada um otimizado para desempenhar funções específicas em visão computacional. Essas arquiteturas são concebidas para atender a uma ampla gama de necessidades, desde a detecção simples de objetos até tarefas mais desafiadoras, como segmentação de instâncias, identificação de objetos,

estimativa de pose e classificação.

Assim, para os testes que serão apresentados no próximo capítulo, foi escolhida a arquitetura YOLOv8, utilizando-se da sua tarefa de classificação. Além disso, ela apresenta 5 versões escaláveis para o seu modelo de classificação: YOLOv8n-cls (*nano*); YOLOv8s-cls (*small*); YOLOv8m-cls (*medium*); YOLOv8l-cls (*large*); YOLOv8x-cls (*extra large*). A versão *nano* é tratada como a mais rápida na realização de treinamentos e classificações, além de possuir menor número de camadas. Enquanto isso, a versão *extra large* é a mais robusta e com maior capacidade de precisão, mas a mais lenta entre elas com base nos testes realizados pela *Ultralytics* utilizando a base de dados *ImageNet* (RUSSAKOVSKY et al., 2015), que inclui cerca de 1000 classes.

### 4.3 Definição dos Hiperparâmetros

A definição dos hiperparâmetros do treinamento foi realizada de maneira empírica através de execuções preliminares e comparações entre cada um dos modelos gerados. Nestas comparações, os modelos foram produzidos utilizando os mesmos conjuntos de dados e o critério de avaliação mais importante que foi observada para decidir se houve melhoras ou não foi a acurácia, observando seus valores para o conjunto de validação.

Os seguintes hiperparâmetros foram os que tiveram um impacto positivo na geração dos modelos: taxa de aprendizado inicial de  $1 \times 10^{-3}$ ; tamanho do *batch* = 16; tamanho das imagens para 160px; otimizador como "auto" de modo que o modelo possa escolher o melhor otimizador ao longo do treinamento dentre as opções possíveis (*SGD*, *Adam*, *Adamax*, *AdamW*, *NAdam*, *RAdam*, *RMSProp*).

Na geração dos modelos, também foi utilizado aumento de dados, técnica aplicada para diversificar e expandir o conjuntos de dados de treinamento. Ela introduz várias transformações e modificações nas imagens originais, podendo variar dependendo da método utilizado.

Assim, para o treinamento foram realizados os seguintes aumento de dados: *HSV-Value* afetando a luminosidade da imagem; *translate* que desloca a imagem em diferentes direções; *scale* para dar zoom na imagem; *flip left-right* para virar a imagem na horizontal.

Alguns parâmetros de *augmentation*, ativos por padrão pelo YOLOv8, foram des-

cartados após testes preliminares, já que, quando desativados, os novos modelos gerados tiveram melhora na acurácia. São estes: *HSV-Hue* e *HSV-Saturation*, modificando a cor e intensidade da cor nas imagens, mas não tiveram efeitos nas imagens radiográficas como esperado; *mosaic* que retira pedaços de partes aleatórias das imagens e colam estes pedaços em uma nova imagem, não sendo interessante pois pode eliminar as regiões de interesse do dente que indicam a idade do indivíduo.

Nos experimentos, detalhados no próximo capítulo, os modelos foram treinados por 1000 épocas sem a utilização de pesos pré-treinados, empregando *Early Stopping* com valor igual a 150. Esta técnica tem como objetivo de interromper o processo de aprendizado do modelo assim que o seu desempenho sobre o conjunto de validação parar de melhorar por 150 épocas, evitando *overfitting*. O *overfitting* ocorre quando o modelo se ajusta excessivamente aos dados de treinamento, comprometendo sua capacidade de generalização para novos dados e resultando em um desempenho inferior em cenários do mundo real.

## 4.4 Critérios de Avaliação

Nesta seção, são apresentados os critérios de avaliação adotados aqui para avaliar o desempenho dos modelos de classificação desenvolvidos: *recall*, acurácia, precisão e *F-measure*.

Para auxiliar o cálculo desses critérios, são geradas matrizes de confusão, uma abordagem muito utilizada para avaliar modelos de classificação. Ela mostra as predições do modelo em relação às classes reais (REZENDE, 2003). Na Figura 4.4 é possível observar um exemplo de matriz de confusão gerada a partir dos experimentos, onde as linhas apresentam as classes reais e nas colunas as classes previstas pelo modelo.

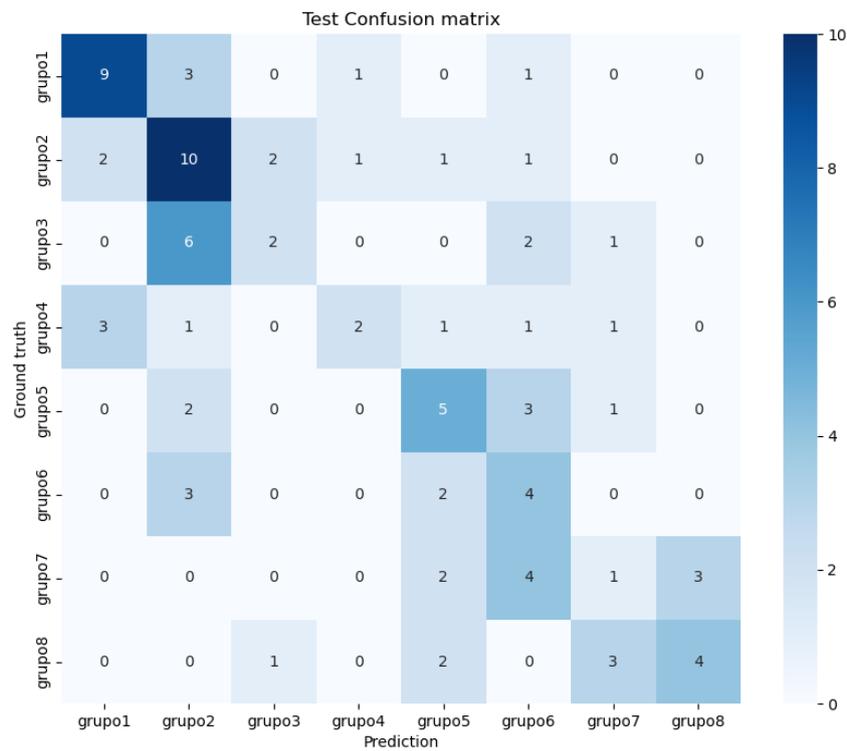


Figura 4.4: Imagem ilustrativa de uma matriz de confusão gerada em um dos experimentos realizados.

- Acurácia

É o principal critério de avaliação, sendo uma medida global que avalia a proporção de predições corretas em relação ao total de predições

$$\text{Acurácia} = \frac{\text{Número de Predições Corretas}}{\text{Total de Predições}} \quad (4.1)$$

- *Recall*

É uma medida que expressa a capacidade do modelo em identificar corretamente instâncias positivas.

$$\text{Recall} = \frac{\text{Verdadeiros Positivos}}{\text{Verdadeiros Positivos} + \text{Falsos Negativos}} \quad (4.2)$$

- Precisão

É uma medida que indica a proporção de instâncias classificadas como positivas que são realmente positivas.

$$\text{Precisão} = \frac{\text{Verdadeiros Positivos}}{\text{Verdadeiros Positivos} + \text{Falsos Positivos}} \quad (4.3)$$

- *F-measure*

É a média harmônica entre precisão e *Recall*. Desse modo, ela é útil quando há um desequilíbrio entre as classes.

$$F\text{-measure} = 2 \times \frac{\text{Precisão} \times \text{Recall}}{\text{Precisão} + \text{Recall}} \quad (4.4)$$

Utilizando esses critérios, é possível analisar o desempenho dos modelos propostos e contrastá-los com os trabalhos presentes na literatura. Embora os dados em cada um dos trabalhos variem, é possível identificar tendências e semelhanças, o que permite avaliar a eficácia e a inovação dos modelos desenvolvidos nesta monografia. Além disso, essa análise comparativa é capaz de revelar pontos fortes e limitações dos modelos propostos.

## 5 Resultados

Este capítulo é destinado a apresentar os resultados obtidos ao gerar modelos de classificação para cada um dos 3 conjunto de dados trabalhados nesta monografia. Sendo que, o mesmo dente de um determinado indivíduo está presente em cada um dos 3 conjunto de dados através do seu respectivo tipo de corte. Dessa forma, a partir dos experimentos, é possível comparar qual tipo de corte apresenta resultados melhores entre si.

Para cada conjunto de dados foram realizados experimentos com as 5 versões de modelo disponíveis para a arquitetura YOLOv8. Além disso, para cada uma das versões de modelo, foram realizados 10 experimentos com 10 sementes aleatórias distintas. Uma visão geral da estrutura desses experimentos pode ser observada na Figura 5.1 Assim, a fim de ilustrar, nas subseções de cada modelo serão apresentados resultados referentes ao desempenho da pior e melhor semente aleatória correspondente ao seu conjunto de dados. Já ao fim do capítulo, serão apresentadas comparações entre os modelos e resultados desses experimentos.

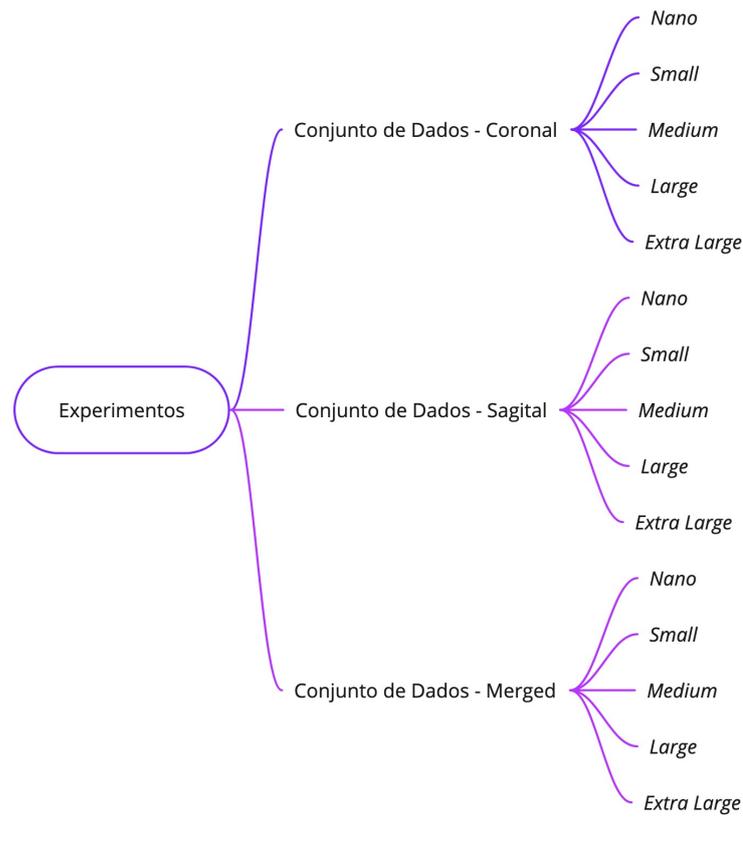


Figura 5.1: Visão geral dos experimentos para cada conjunto de dados utilizando as 5 versões de modelo.

## 5.1 Ambiente de Desenvolvimento

O modelo foi desenvolvido no ambiente do Kaggle, uma plataforma de ciência de dados que oferece um ambiente de desenvolvimento integrado (IDE) baseado na linguagem Python. Um dos principais fatores pela escolha dessa plataforma é que o Kaggle proporciona meios eficientes para o armazenamento de dados e construção dos conjuntos utilizados pelos modelos. Além disso, o Kaggle opera na nuvem, proporcionando maior disponibilidade, pois permite a execução remota de códigos e a obtenção dos resultados após a conclusão de todas as execuções. Ademais, também oferece horas semanais gratuitas para o uso de GPUs (KAGGLE, 2023).

## 5.2 Bibliotecas Utilizadas

Durante a implementação dos códigos, diversas bibliotecas foram utilizadas, sendo as principais a serem citadas são: Ultralytics, utilizada para geração dos modelos do YOLOv8 (JOCHER; CHAURASIA; QIU, 2023); Pandas, para manipulação de dados (TEAM, 2020); Matplotlib, usada para visualização (HUNTER, 2007); Splitfolders, para manipulação de conjunto de dados; Numpy, para manipulação de matrizes e vetores (HARRIS et al., 2020); e Seaborn, para auxiliar na criação de matrizes de confusão (WASKOM, 2021).

Além disso, o *notebook* com os códigos para o treinamento dos modelos, assim como realização dos experimentos, pode ser encontrado na URL <https://www.kaggle.com/code/alexvpp/odonto-train-yolo>. Como a base de dados com as imagens se encontra em fase de construção, ainda não será disponibilizada, mas será futuramente.

## 5.3 Conjunto de Dados 1 - Corte Coronal

### 5.3.1 Modelo *Nano*

A Figura 5.2 apresenta a matriz de confusão, exibindo os resultados da inferência no conjunto de teste para o modelo *Nano* com a melhor semente aleatória identificada. Além disso, os valores dos critérios de avaliação associados são os seguintes: Acurácia, 0,407; *Recall* médio, 0,358; Precisão média, 0,328; *F-measure* médio, 0,319.

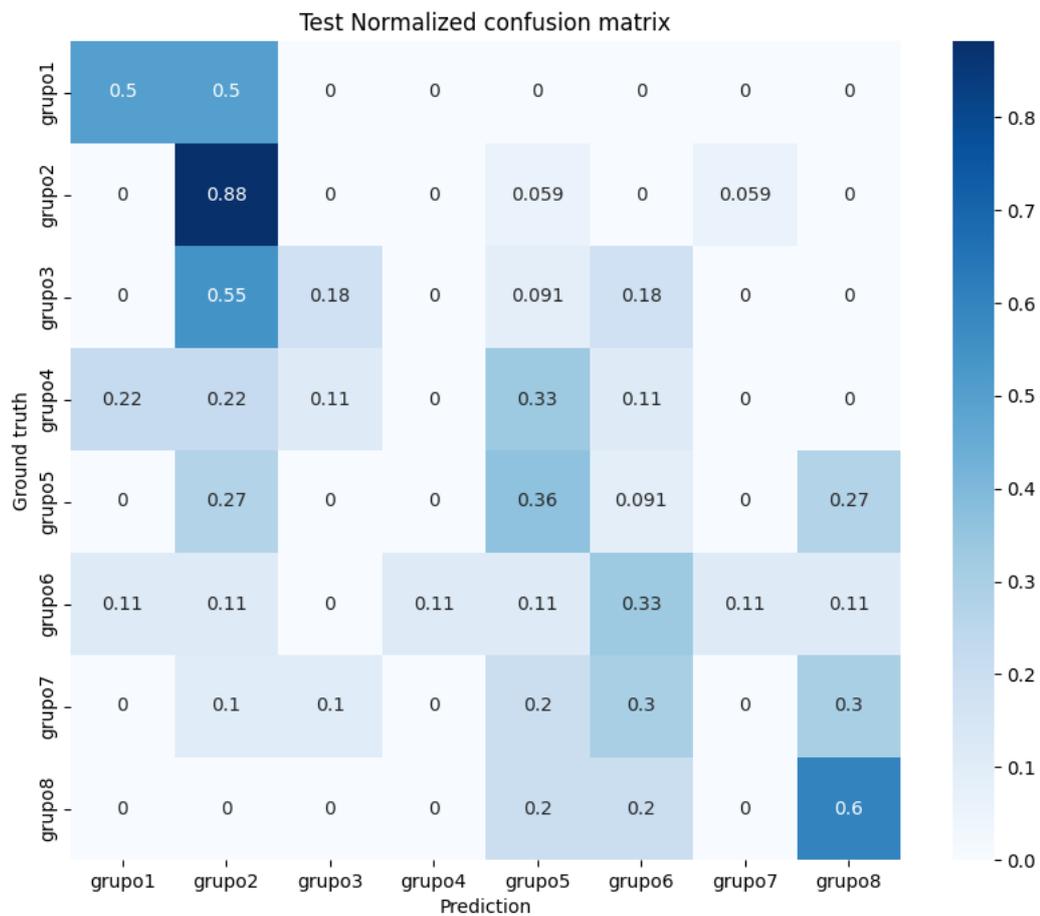


Figura 5.2: Matriz de confusão normalizada do conjunto de teste para o corte coronal, modelo *nano* com a melhor semente aleatória identificada.

Adicionalmente, na Figura 5.3, são apresentados os detalhes do seu treinamento em que a última época foi a 181 e a função de perda para o conjunto de treino teve seu menor valor em 0,257. Já a função de perda para o conjunto de validação obteve o menor valor em 1,858. Para este mesmo conjunto, é possível observar a variação dos valores de acurácia, onde a acurácia *Top 1*, que foca na classe mais provável, atingiu seu ponto máximo em 0,455. Ademais, a acurácia *Top 5*, que busca identificar se a classe correta entre as cinco principais classes previstas, teve o seu valor mais alto em 0,937.

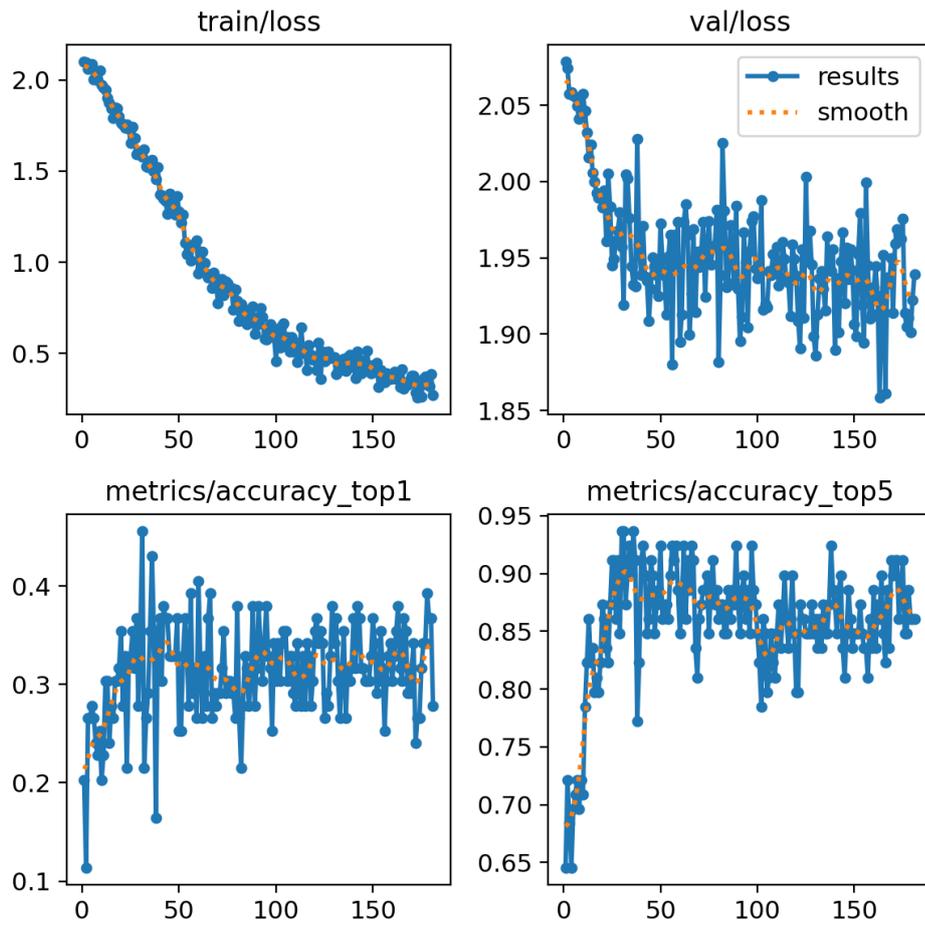


Figura 5.3: Resultados do treinamento para o corte coronal, modelo *nano* com a melhor semente aleatória identificada.

Com relação aos resultados de inferência do conjunto de teste para a pior semente aleatória identificada, os valores de cada critério foram: Acurácia, 0,286; *Recall*, 0,252; Precisão, 0,226; *F-measure*, 0,236.

### 5.3.2 Modelo *Small*

A Figura 5.4 apresenta a matriz de confusão, exibindo os resultados da inferência no conjunto de teste para o modelo *Small* com a melhor semente aleatória identificada. Além disso, os valores dos critérios de avaliação associados são os seguintes: Acurácia, 0,429; *Recall* médio, 0,386; Precisão média, 0,364; *F-measure* médio, 0,342.

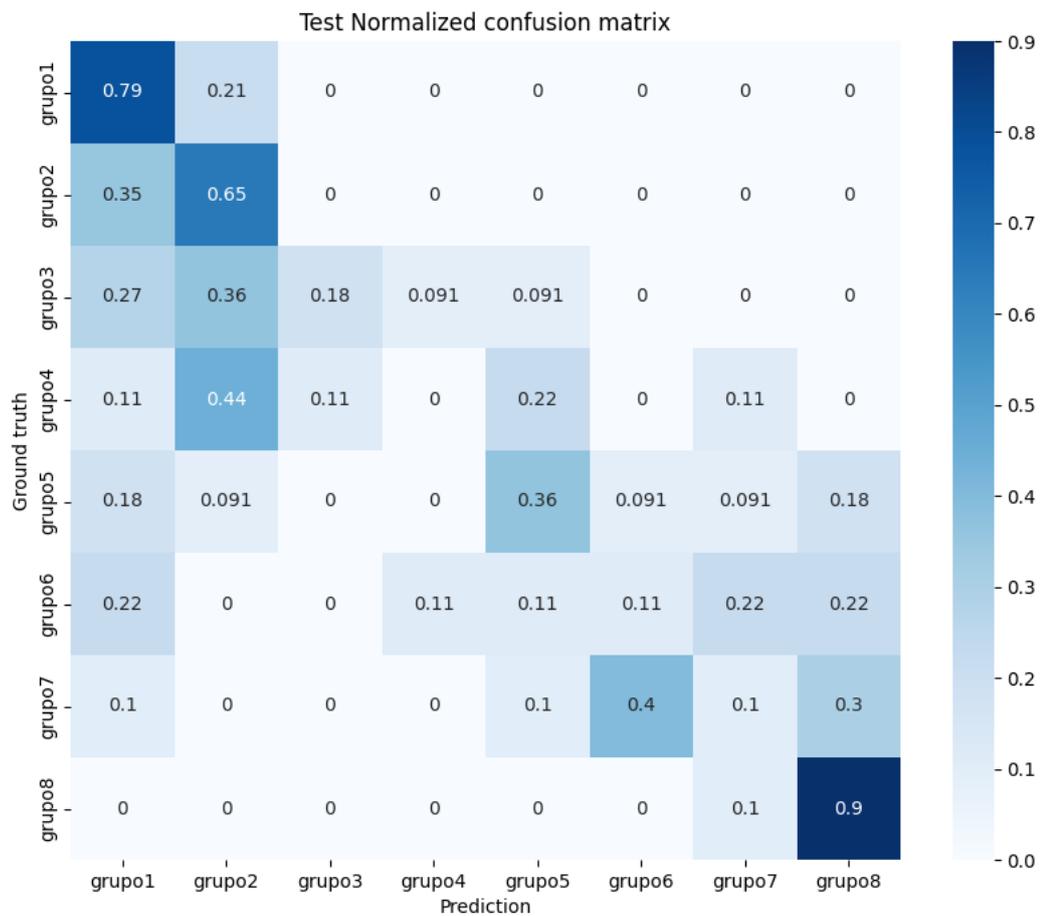


Figura 5.4: Matriz de confusão normalizada do conjunto de teste para o corte coronal, modelo *Small* com a melhor semente aleatória identificada.

Adicionalmente, na Figura 5.5, são apresentados os detalhes do seu treinamento em que a última época foi a 271 e a função de perda para o conjunto de treino teve seu menor valor em 0,171. Já a função de perda para o conjunto de validação obteve o menor valor em 1,816. Para este mesmo conjunto, é possível observar a variação dos valores de acurácia, onde a acurácia *Top 1* atingiu seu ponto máximo em 0,481. Ademais, a acurácia *Top 5* teve o seu valor mais alto em 0,937.

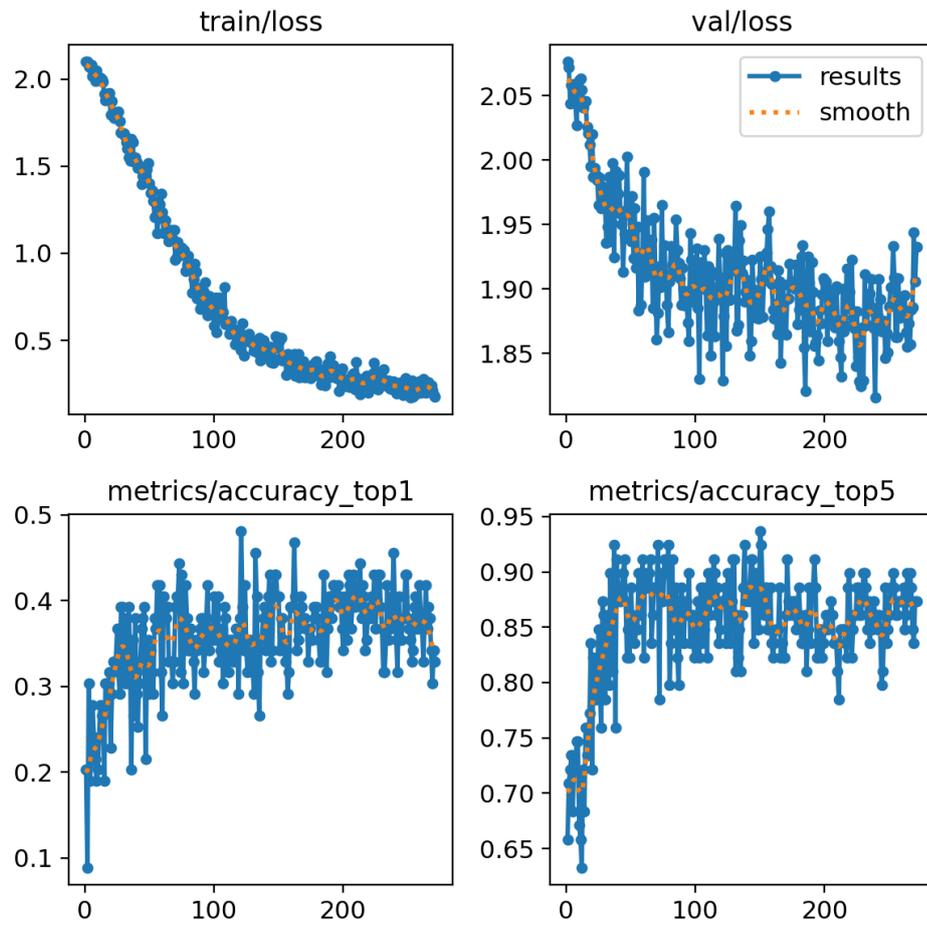


Figura 5.5: Resultados do treinamento para o corte coronal, modelo *Small* com a melhor semente aleatória identificada.

Com relação aos resultados de inferência do conjunto de teste para a pior semente aleatória identificada, os valores de cada critério foram: Acurácia, 0,264; *Recall* médio, 0,24; Precisão média, 0,27; *F-measure* médio, 0,241.

### 5.3.3 Modelo *Medium*

A Figura 5.6 apresenta a matriz de confusão, exibindo os resultados da inferência no conjunto de teste para o modelo *Medium* com a melhor semente aleatória identificada. Além disso, os valores dos critérios de avaliação associados são os seguintes: Acurácia, 0,396; *Recall* médio, 0,368; Precisão média, 0,360; *F-measure* médio, 0,345.

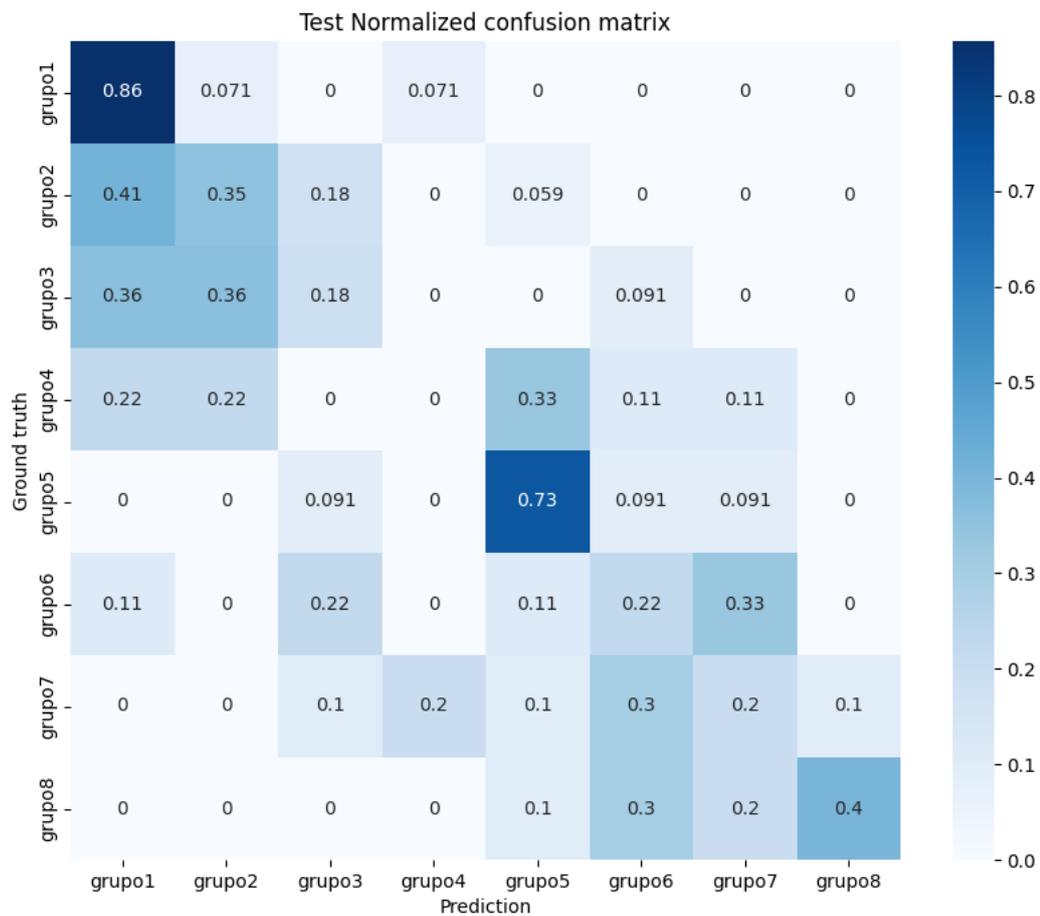


Figura 5.6: Matriz de confusão normalizada do conjunto de teste para o corte coronal, modelo *Medium* com a melhor semente aleatória identificada.

Adicionalmente, na Figura 5.7, são apresentados os detalhes do seu treinamento em que a última época foi a 410 e a função de perda para o conjunto de treino teve seu menor valor em 0,106. Já a função de perda para o conjunto de validação obteve o menor valor em 1,787. Para este mesmo conjunto, é possível observar a variação dos valores de acurácia, onde a acurácia *Top 1* atingiu seu ponto máximo em 0,49. Ademais, a acurácia *Top 5* teve o seu valor mais alto em 0,974.

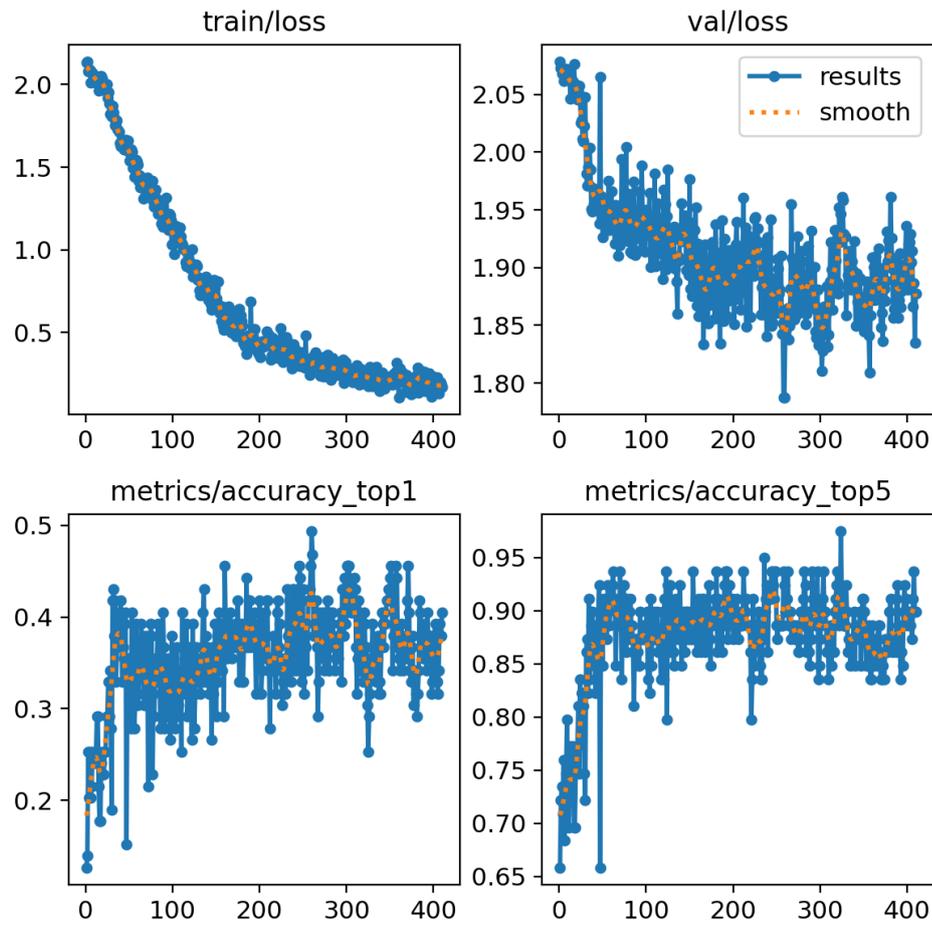


Figura 5.7: Resultados do treinamento para o corte coronal, modelo *Medium* com a melhor semente aleatória identificada.

Com relação aos resultados de inferência do conjunto de teste para a pior semente aleatória identificada, os valores de cada critério foram: Acurácia, 0,297; *Recall* médio, 0,276; Precisão média, 0,283; *F-measure* médio, 0,268.

### 5.3.4 Modelo *Large*

A Figura 5.8 apresenta a matriz de confusão, exibindo os resultados da inferência no conjunto de teste para o modelo *Large* com a melhor semente aleatória identificada. Acurácia, 0,429; *Recall* médio, 0,383; Precisão média, 0,398; *F-measure* médio, 0,386.

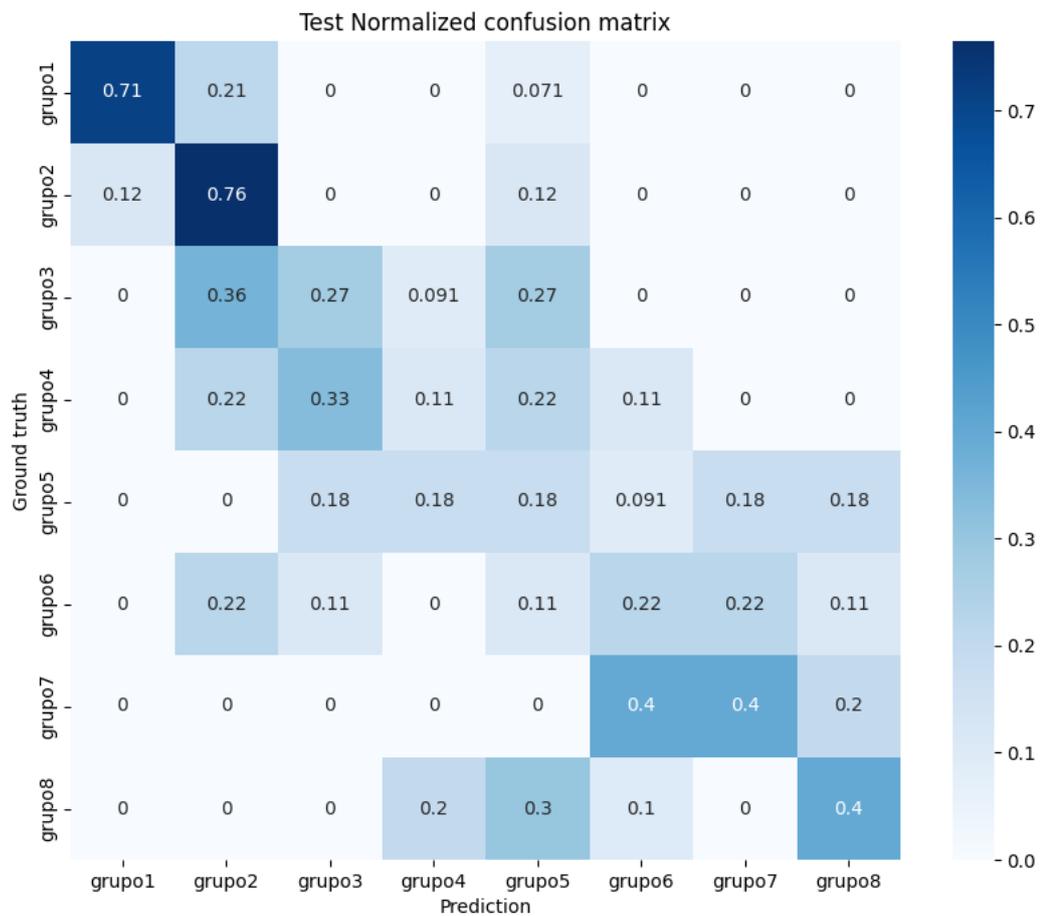


Figura 5.8: Matriz de confusão normalizada do conjunto de teste para o corte coronal, modelo *Large* com a melhor semente aleatória identificada.

Adicionalmente, na Figura 5.9, são apresentados os detalhes do seu treinamento em que a última época foi a 394 e a função de perda para o conjunto de treino teve seu menor valor em 0,094. Já a função de perda para o conjunto de validação obteve o menor valor em 1,846. Para este mesmo conjunto, é possível observar a variação dos valores de acurácia, onde a acurácia *Top 1* atingiu seu ponto máximo em 0,45. Ademais, a acurácia *Top 5* teve o seu valor mais alto em 0,962.

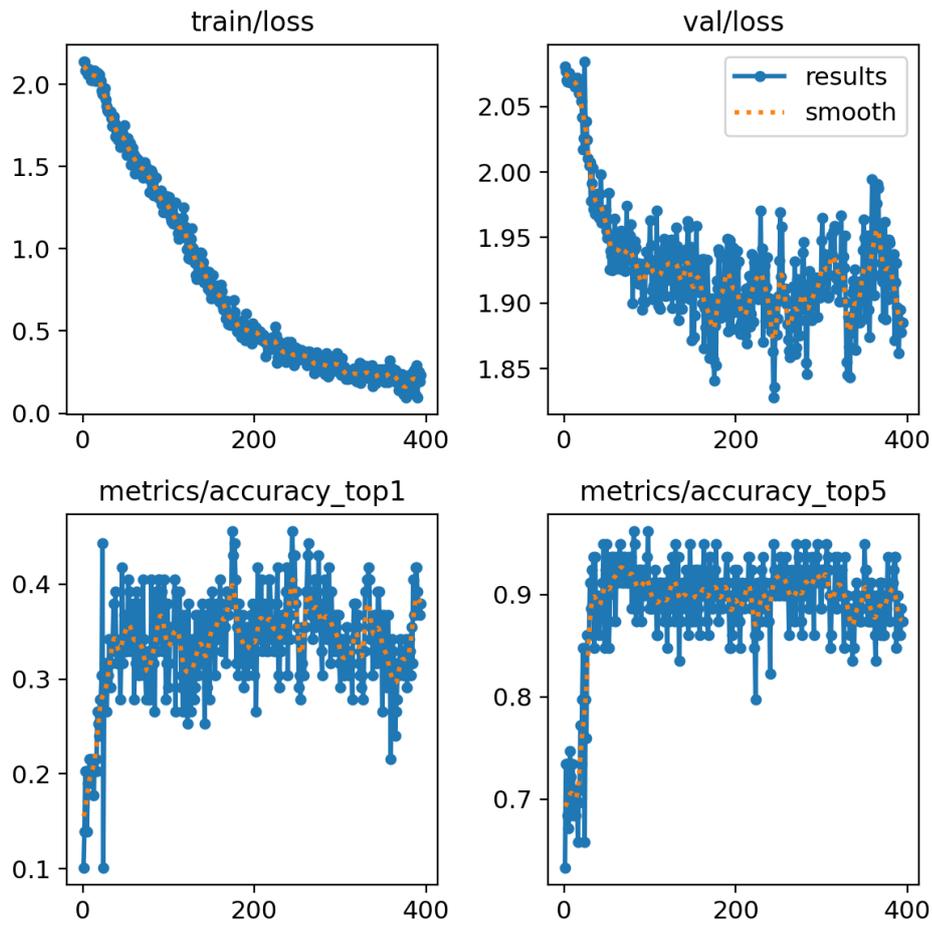


Figura 5.9: Resultados do treinamento para o corte coronal, modelo *Large* com a melhor semente aleatória identificada.

Com relação aos resultados de inferência do conjunto de teste para a pior semente aleatória identificada, os valores de cada critério foram: Acurácia, 0,286; *Recall* médio, 0,247; Precisão média, 0,221; *F-measure* médio, 0,223.

### 5.3.5 Modelo *Extra Large*

A Figura 5.10 apresenta a matriz de confusão, exibindo os resultados da inferência no conjunto de teste para o modelo *Extra Large* com a melhor semente aleatória identificada. Além disso, os valores dos critérios de avaliação associados são os seguintes: Acurácia, 0,396; *Recall* médio, 0,359; Precisão média, 0,369; *F-measure* médio, 0,355.

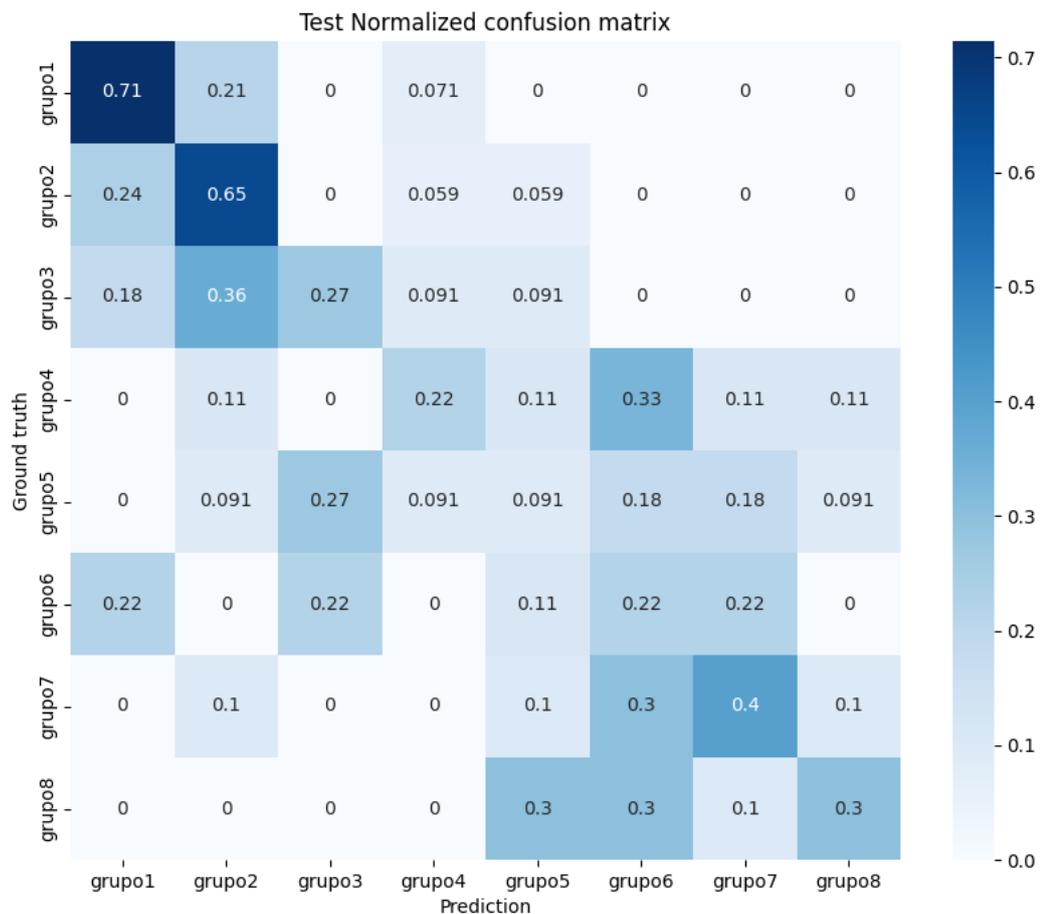


Figura 5.10: Matriz de confusão normalizada do conjunto de teste para o corte coronal, modelo *Extra Large* com a melhor semente aleatória identificada.

Adicionalmente, na Figura 5.11, são apresentados os detalhes do seu treinamento em que a última época foi a 463 e a função de perda para o conjunto de treino teve seu menor valor em 0,128. Já a função de perda para o conjunto de validação obteve o menor valor em 1,779. Para este mesmo conjunto, é possível observar a variação dos valores de acurácia, onde a acurácia *Top 1* atingiu seu ponto máximo em 0,512. Ademais, a acurácia *Top 5* teve o seu valor mais alto em 0,962.

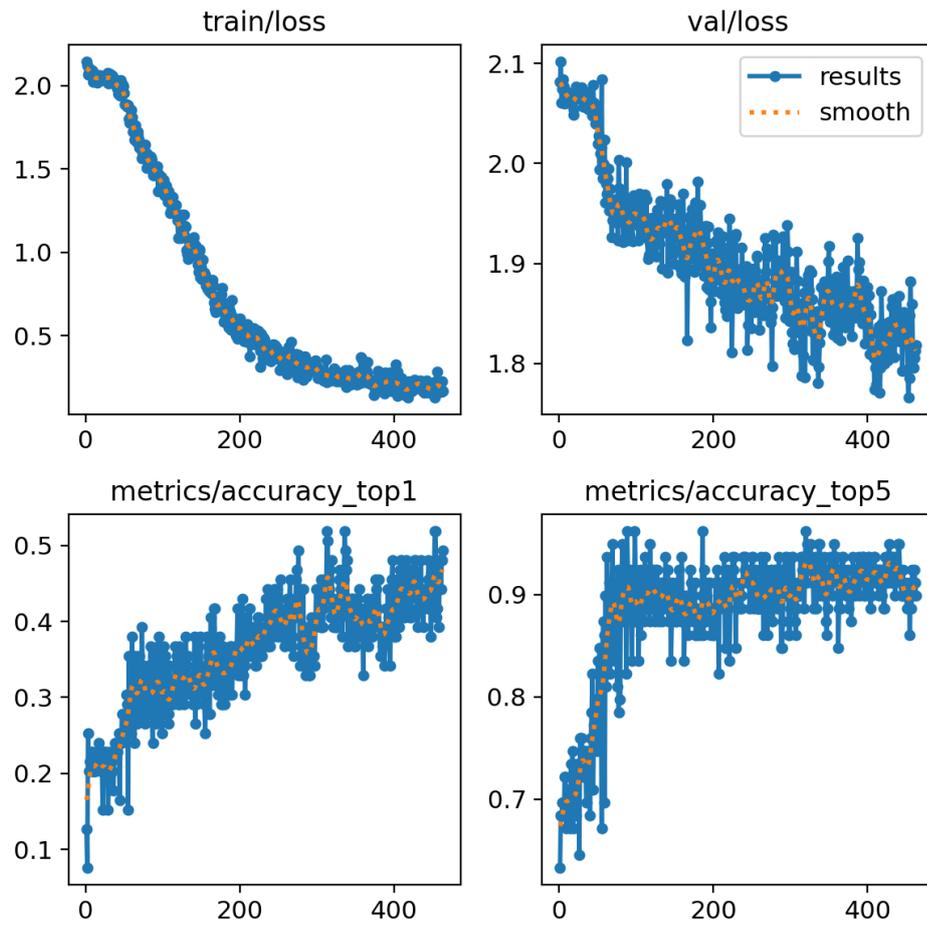


Figura 5.11: Resultados do treinamento para o corte coronal, modelo *Extra Large* com a melhor semente aleatória identificada.

Com relação aos resultados de inferência do conjunto de teste para a pior semente aleatória identificada, os valores de cada critério foram: Acurácia, 0,297; *Recall* médio, 0,267; Precisão média, 0,267; *F-measure* médio, 0,245.

## 5.4 Conjunto de Dados 2 - Corte Sagital

### 5.4.1 Modelo *Nano*

A Figura 5.12 apresenta a matriz de confusão, exibindo os resultados da inferência no conjunto de teste para o modelo *Nano* com a melhor semente aleatória identificada. Além disso, os valores dos critérios associados são os seguintes: Acurácia, 0,516; *Recall* médio, 0,505; Precisão média, 0,518; *F-measure* médio, 0,506.

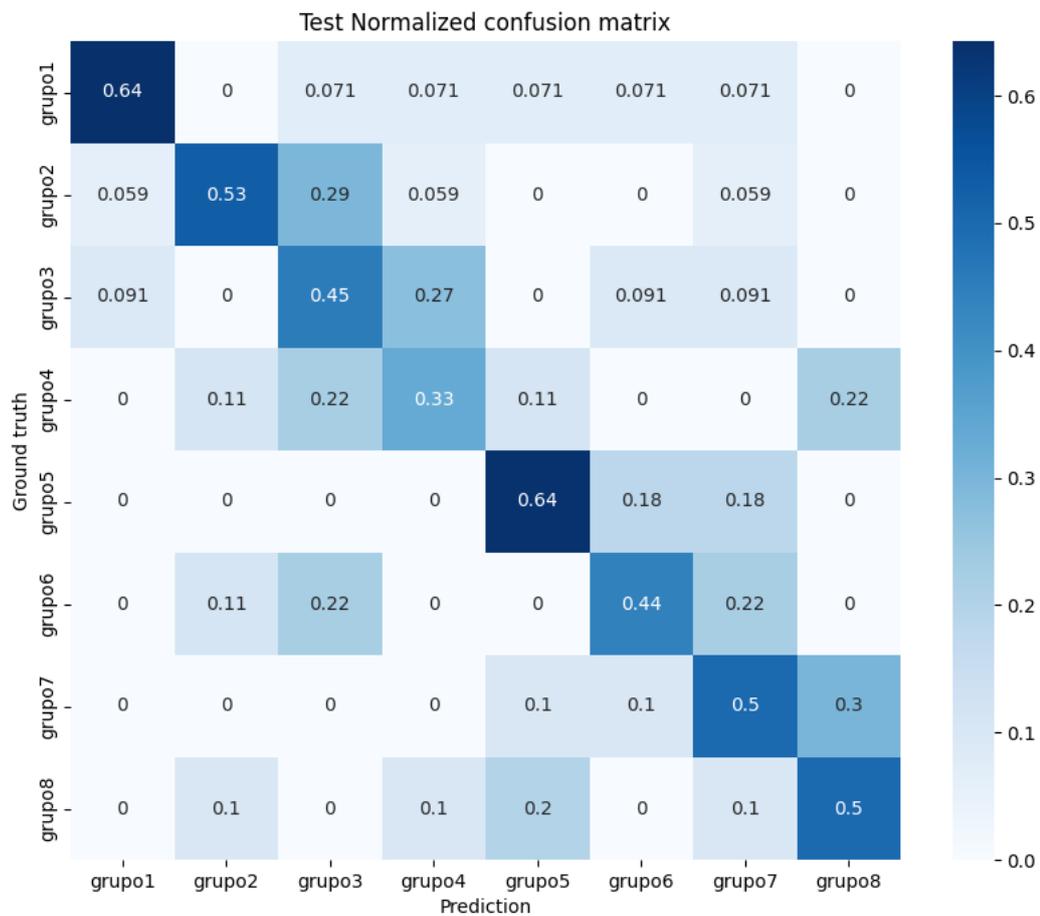


Figura 5.12: Matriz de confusão normalizada do conjunto de teste para o corte sagital, modelo *Nano* com a melhor semente aleatória identificada.

Adicionalmente, na Figura 5.13, são apresentados os detalhes do seu treinamento em que a última época foi a 521 e a função de perda para o conjunto de treino teve seu menor valor em 0,087. Já a função de perda para o conjunto de validação obteve o menor valor em 1,708. Para este mesmo conjunto, é possível observar a variação dos valores de acurácia, onde a acurácia *Top 1* atingiu seu ponto máximo em 0,58. Ademais, a acurácia *Top 5* teve o seu valor mais alto em 0,987.

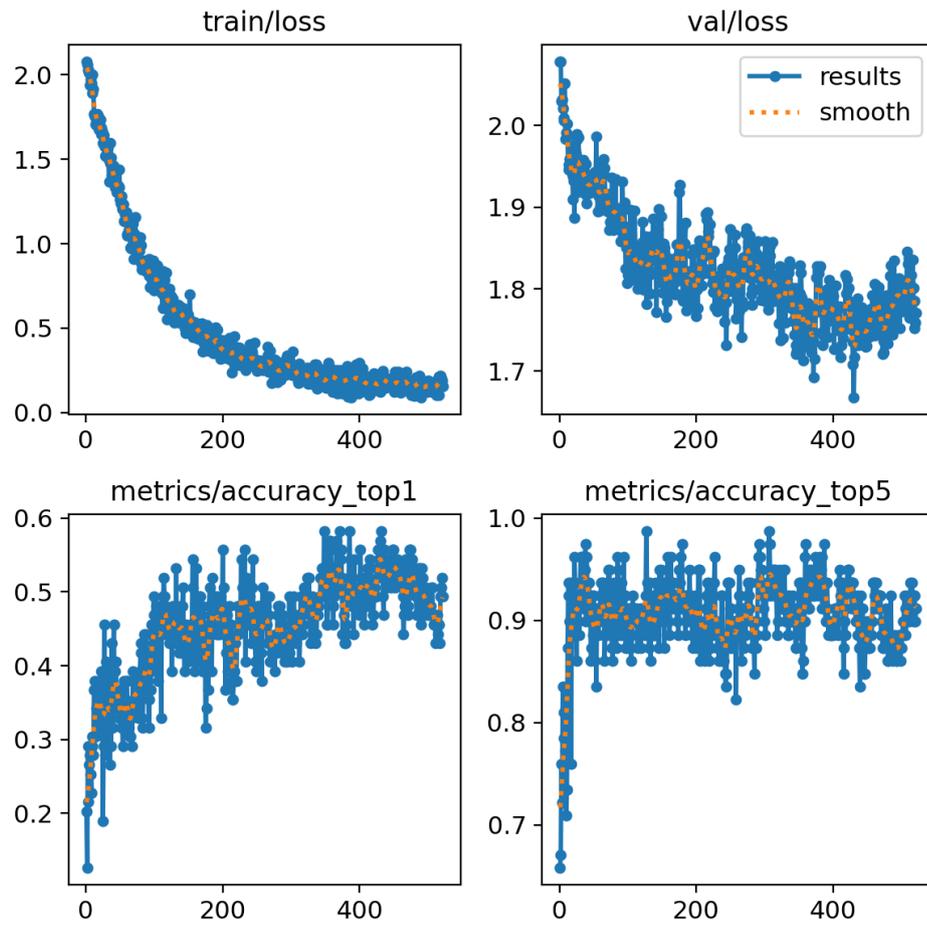


Figura 5.13: Resultados do treinamento para o corte sagital, modelo *Nano* com a melhor semente aleatória identificada.

Com relação aos resultados de inferência do conjunto de teste para a pior semente aleatória identificada, os valores de cada critério foram: Acurácia, 0,330; *Recall* médio, 0,301; Precisão média, 0,325; *F-measure* médio, 0,298.

### 5.4.2 Modelo *Small*

A Figura 5.14 apresenta a matriz de confusão, exibindo os resultados da inferência no conjunto de teste para o modelo *Small* com a melhor semente aleatória identificada. Além disso, os valores dos critérios de avaliação associados são os seguintes: Acurácia, 0,495; *Recall* médio, 0,459; Precisão média, 0,469; *F-measure* médio, 0,453.

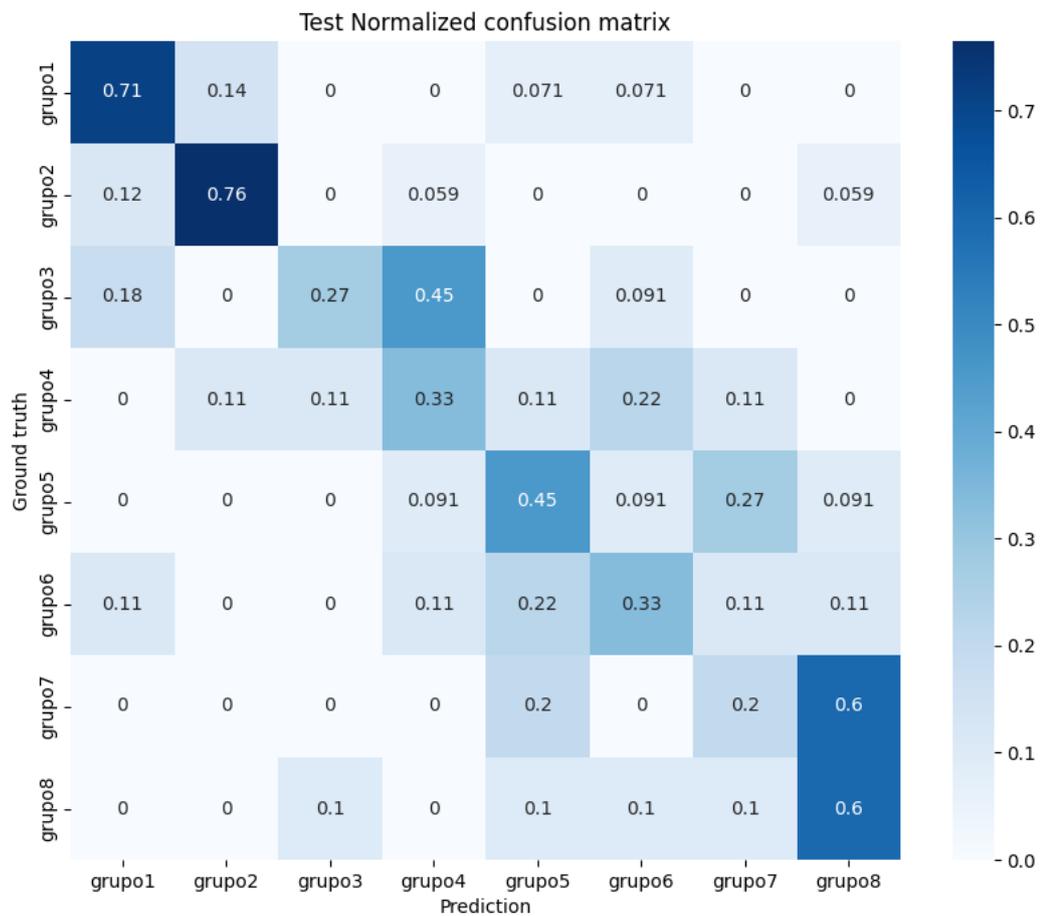


Figura 5.14: Matriz de confusão normalizada do conjunto de teste para o corte sagital, modelo *Small* com a melhor semente aleatória identificada.

Adicionalmente, na Figura 5.15, são apresentados os detalhes do seu treinamento em que a última época foi a 665 e a função de perda para o conjunto de treino teve seu menor valor em 0,059. Já a função de perda para o conjunto de validação obteve o menor valor em 1,723. Para este mesmo conjunto, é possível observar a variação dos valores de acurácia, onde a acurácia *Top 1* atingiu seu ponto máximo em 0,58. Ademais, a acurácia *Top 5* teve o seu valor mais alto em 0,987.

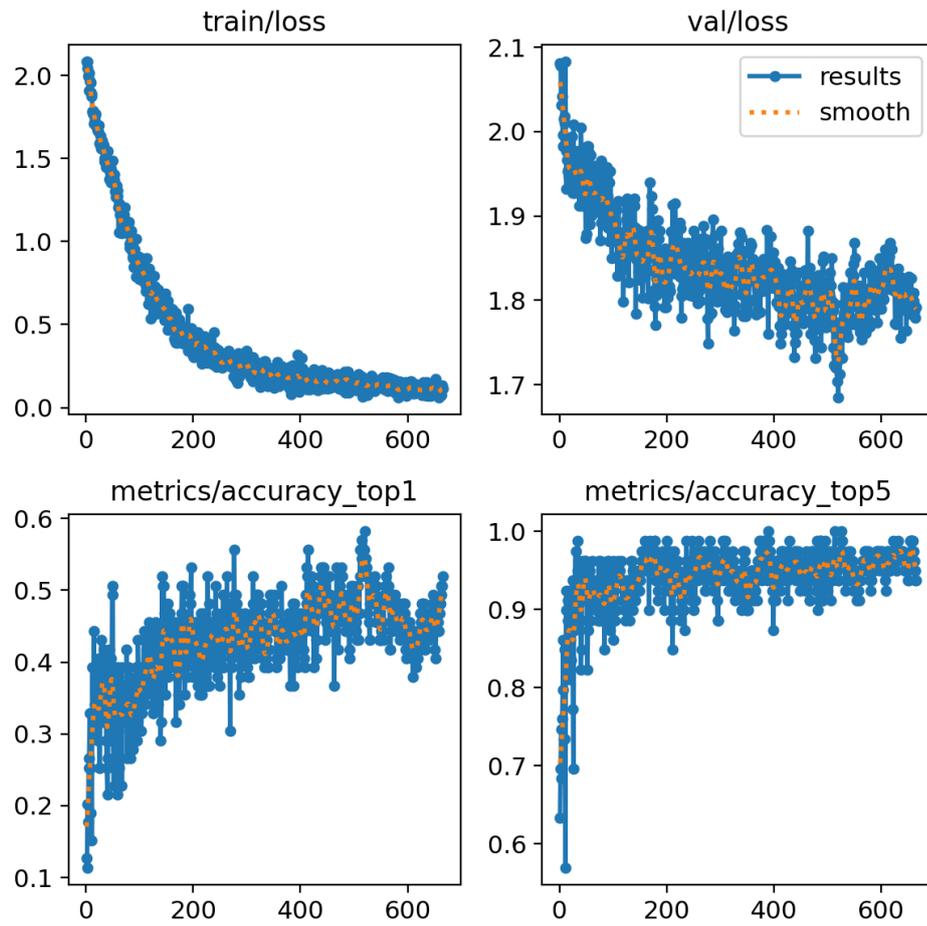


Figura 5.15: Resultados do treinamento para o corte sagital, modelo *Small* com a melhor semente aleatória identificada.

Com relação aos resultados de inferência do conjunto de teste para a pior semente aleatória identificada, os valores de cada critério foram: Acurácia, 0,385; *Recall* médio, 0,346; Precisão média, 0,343; *F-measure* médio, 0,330.

### 5.4.3 Modelo *Medium*

A Figura 5.16 apresenta a matriz de confusão, exibindo os resultados da inferência no conjunto de teste para o modelo *Medium* com a melhor semente aleatória identificada. Além disso, os valores dos critérios de avaliação associados são os seguintes: Acurácia, 0,407; *Recall* médio, 0,391; Precisão média, 0,433; *F-measure* médio, 0,389.

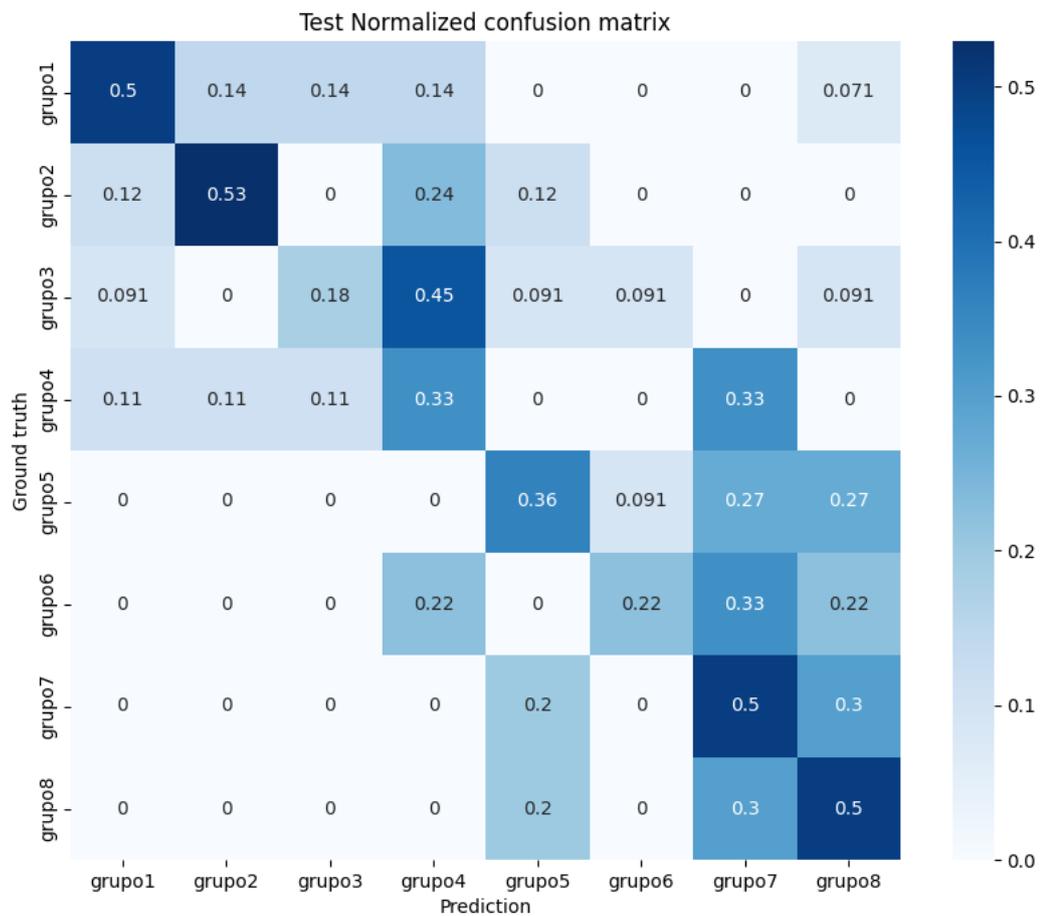


Figura 5.16: Matriz de confusão normalizada do conjunto de teste para o corte sagital, modelo *Medium* com a melhor semente aleatória identificada.

Adicionalmente, na Figura 5.17, são apresentados os detalhes do seu treinamento em que a última época foi a 515 e a função de perda para o conjunto de treino teve seu menor valor em 0,099. Já a função de perda para o conjunto de validação obteve o menor valor em 1,762. Para este mesmo conjunto, é possível observar a variação dos valores de acurácia, onde a acurácia *Top 1* atingiu seu ponto máximo em 0,55. Ademais, a acurácia *Top 5* teve o seu valor mais alto em 0,987.

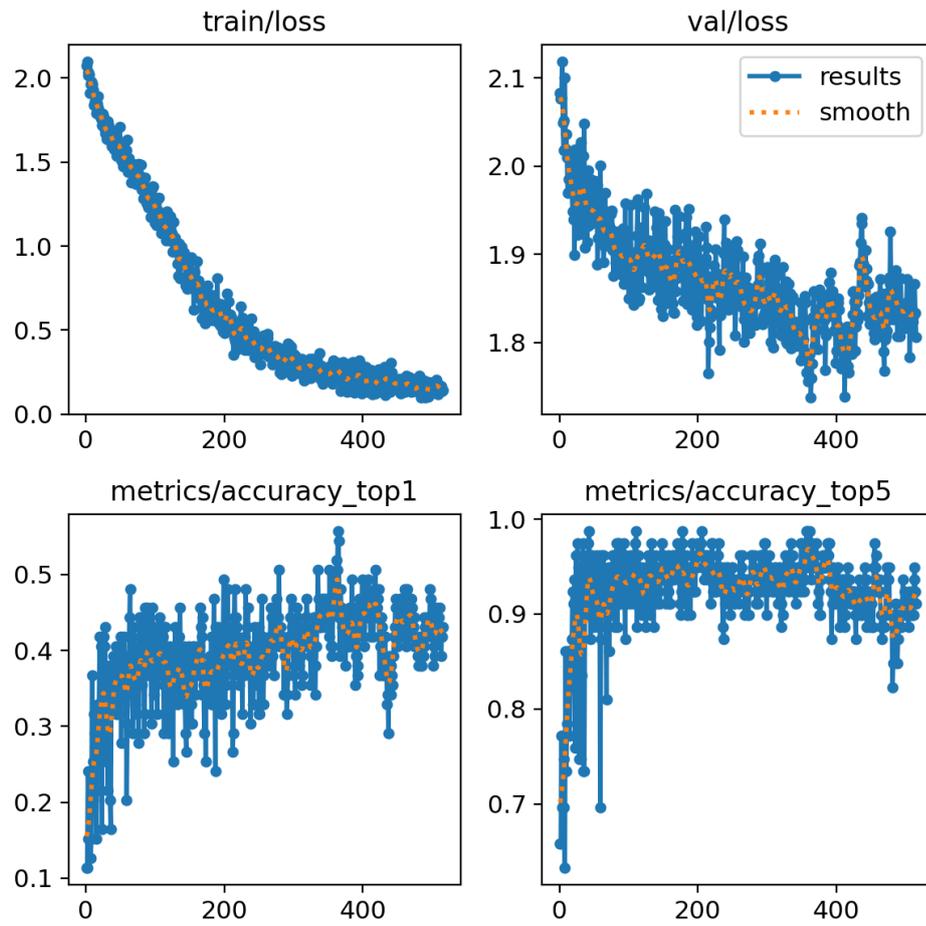


Figura 5.17: Resultados do treinamento para o corte sagital, modelo *Medium* com a melhor semente aleatória identificada.

Com relação aos resultados de inferência do conjunto de teste para a pior semente aleatória identificada, os valores de cada critério foram: Acurácia, 0,275; *Recall* médio, 0,232; Precisão média, 0,229; *F-measure* médio, 0,226.

#### 5.4.4 Modelo *Large*

A Figura 5.18 apresenta a matriz de confusão, exibindo os resultados da inferência no conjunto de teste para o modelo *Large* com a melhor semente aleatória identificada. Além disso, os valores dos critérios de avaliação associados são os seguintes: Acurácia, 0,385; *Recall* médio, 0,356; Precisão média, 0,352; *F-measure* médio, 0,344.

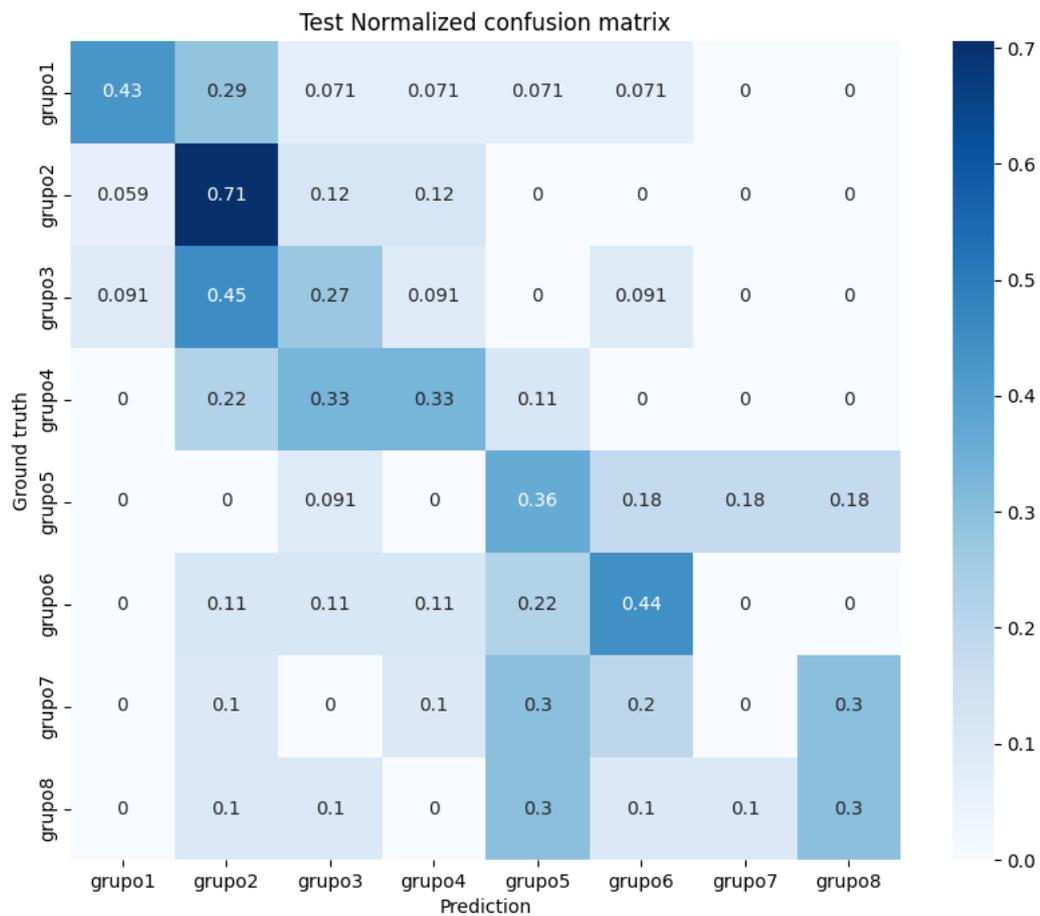


Figura 5.18: Matriz de confusão normalizada do conjunto de teste para o corte sagital, modelo *Large* com a melhor semente aleatória identificada.

Adicionalmente, na Figura 5.19, são apresentados os detalhes do seu treinamento em que a última época foi a 424 e a função de perda para o conjunto de treino teve seu menor valor em 0,146. Já a função de perda para o conjunto de validação obteve o menor valor em 1,806. Para este mesmo conjunto, é possível observar a variação dos valores de acurácia, onde a acurácia *Top 1* atingiu seu ponto máximo em 0,49. Ademais, a acurácia *Top 5* teve o seu valor mais alto em 0,987.

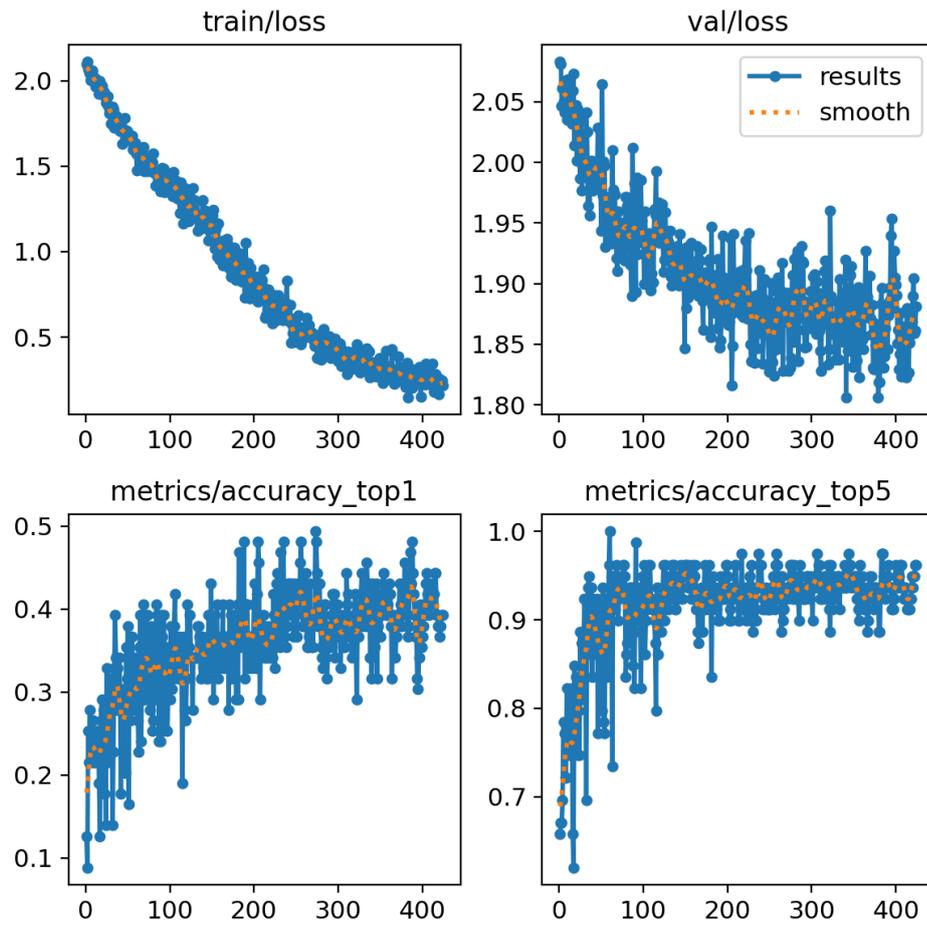


Figura 5.19: Resultados do treinamento para o corte sagital, modelo *Large* com a melhor semente aleatória identificada.

Com relação aos resultados de inferência do conjunto de teste para a pior semente aleatória identificada, os valores de cada critério foram: Acurácia, 0,297; *Recall* médio, 0,264; Precisão média, 0,221; *F-measure* médio, 0,231.

### 5.4.5 Modelo *Extra Large*

A Figura 5.20 apresenta a matriz de confusão, exibindo os resultados da inferência no conjunto de teste para o modelo *Extra Large* com a melhor semente aleatória identificada. Além disso, os valores dos critérios de avaliação associados são os seguintes: Acurácia, 0,429; *Recall* médio, 0,403; Precisão média, 0,480; *F-measure* médio, 0,428.

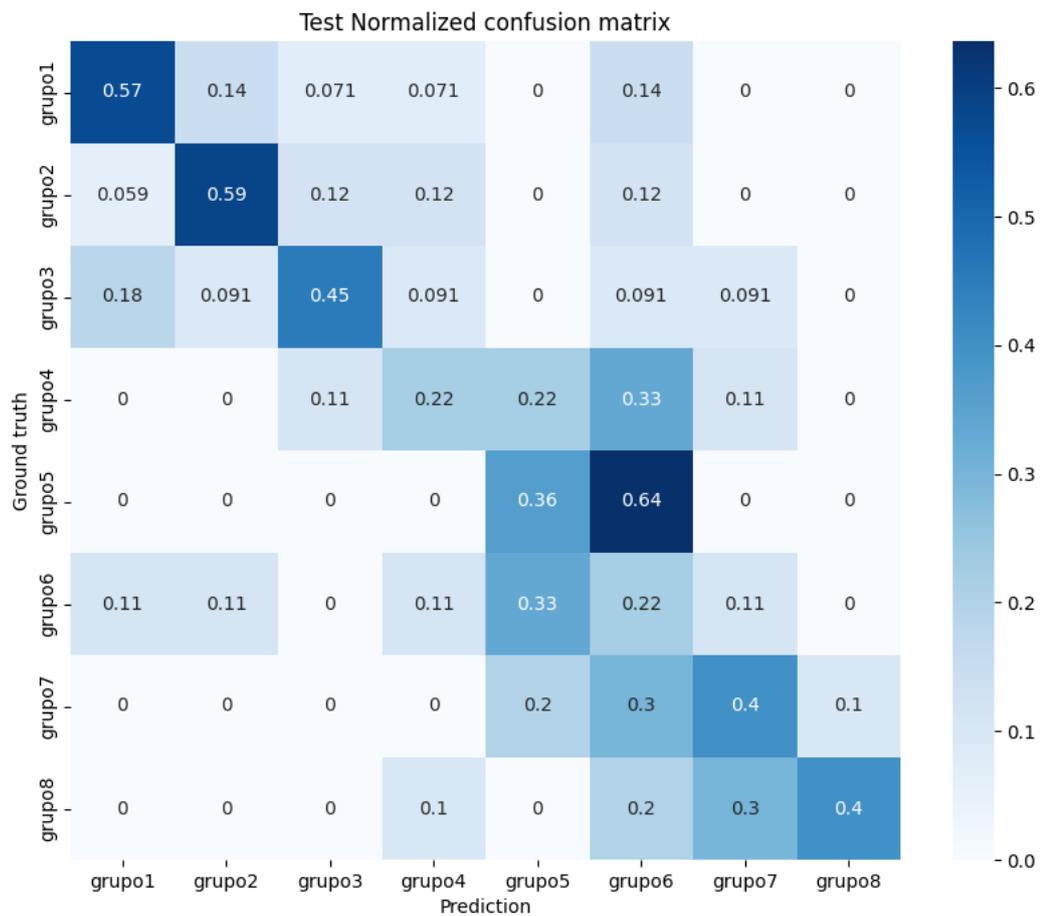


Figura 5.20: Matriz de confusão normalizada do conjunto de teste para o corte sagital, modelo *Extra Large* com a melhor semente aleatória identificada.

Adicionalmente, na Figura 5.21, são apresentados os detalhes do seu treinamento em que a última época foi a 464 e a função de perda para o conjunto de treino teve seu menor valor em 0,204. Já a função de perda para o conjunto de validação obteve o menor valor em 1,776. Para este mesmo conjunto, é possível observar a variação dos valores de acurácia, onde a acurácia *Top 1* atingiu seu ponto máximo em 0,53. Ademais, a acurácia *Top 5* teve o seu valor mais alto em 0,987.

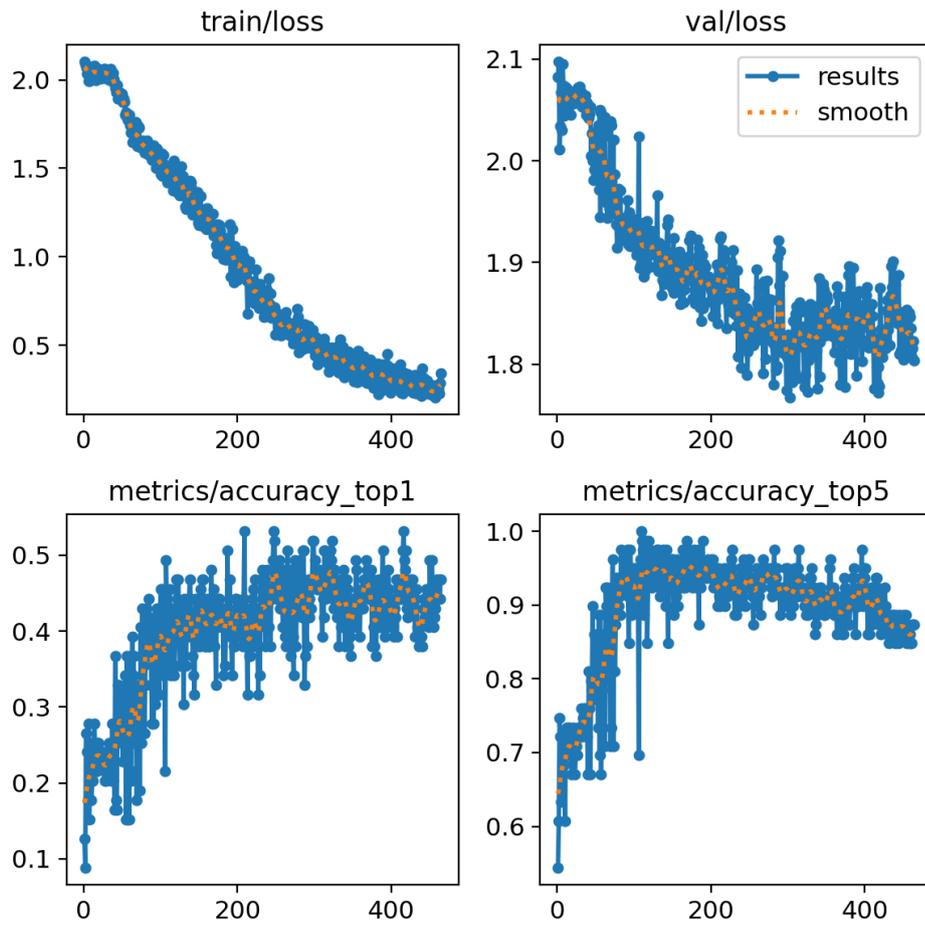


Figura 5.21: Resultados do treinamento para o corte sagital, modelo *Extra Large* com a melhor semente aleatória identificada.

Com relação aos resultados de inferência do conjunto de teste para a pior semente aleatória identificada, os valores de cada critério foram: Acurácia, 0,297; *Recall* médio, 0,287; Precisão média, 0,343; *F-measure* médio, 0,297.

## 5.5 Conjunto de Dados 3 - Corte Concatenado

### 5.5.1 Modelo *Nano*

A Figura 5.22 apresenta a matriz de confusão, exibindo os resultados da inferência no conjunto de teste para o modelo *Nano* com a melhor semente aleatória identificada. Além disso, os valores dos critérios de avaliação associados são os seguintes: Acurácia, 0,527; *Recall* médio, 0,491; Precisão média, 0,520; *F-measure* médio, 0,487.

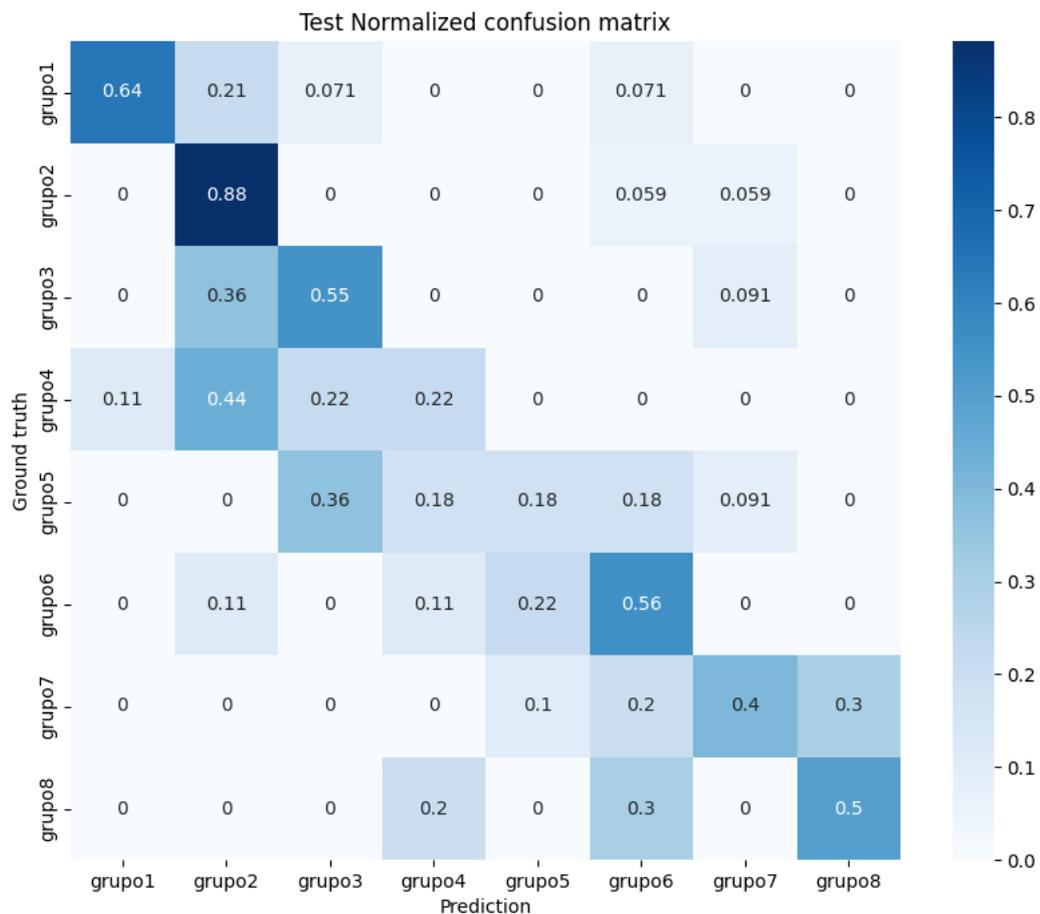


Figura 5.22: Matriz de confusão normalizada do conjunto de teste para o corte concatenado, modelo *Nano* com a melhor semente aleatória identificada.

Adicionalmente, na Figura 5.23, são apresentados os detalhes do seu treinamento em que a última época foi a 346 e a função de perda para o conjunto de treino teve seu menor valor em 0,126. Já a função de perda para o conjunto de validação obteve o menor valor em 1,738. Para este mesmo conjunto, é possível observar a variação dos valores de acurácia, onde a acurácia *Top 1* atingiu seu ponto máximo em 0,6. Ademais, a acurácia *Top 5* teve o seu valor mais alto em 0,974.

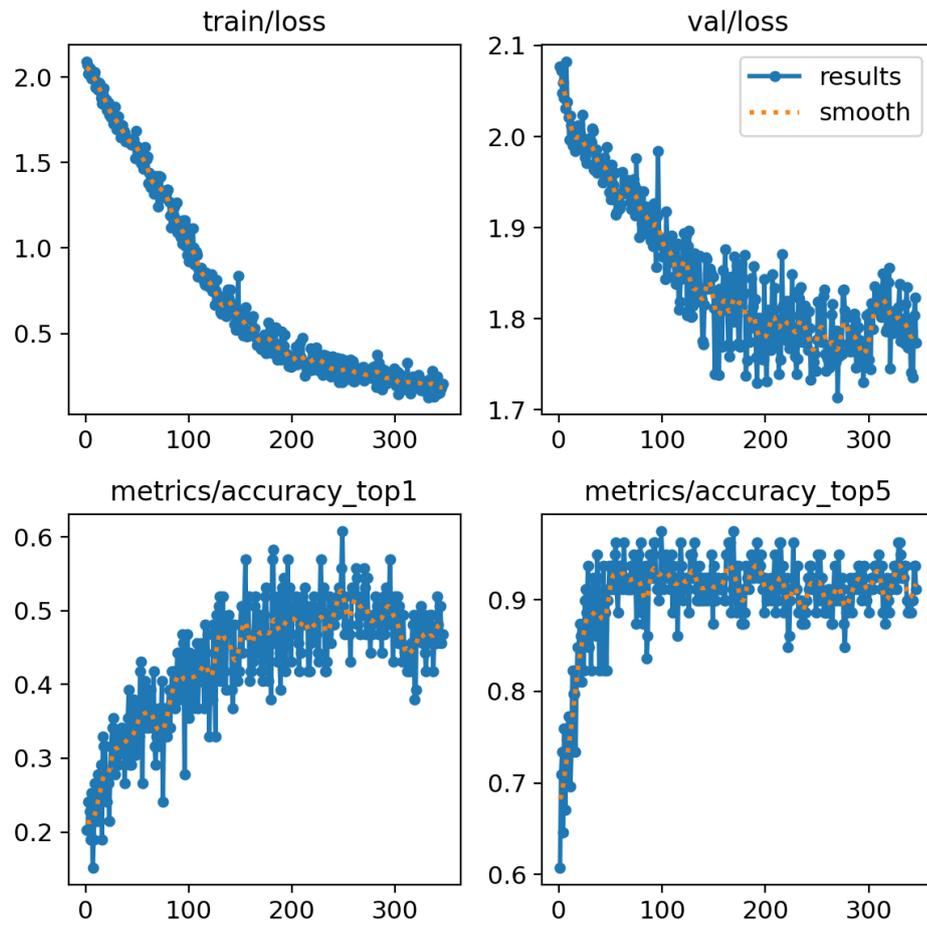


Figura 5.23: Resultados do treinamento para o corte concatenado, modelo *Nano* com a melhor semente aleatória identificada.

Com relação aos resultados de inferência do conjunto de teste para a pior semente aleatória identificada, os valores de cada critério foram: Acurácia, 0,319; *Recall* médio, 0,292; Precisão média, 0,314; *F-measure* médio, 0,295.

### 5.5.2 Modelo *Small*

A Figura 5.24 apresenta a matriz de confusão, exibindo os resultados da inferência no conjunto de teste para o modelo *Small* com a melhor semente aleatória identificada. Além disso, os valores dos critérios de avaliação associados são os seguintes: Acurácia, 0,462; *Recall* médio, 0,420; Precisão média, 0,429; *F-measure* médio, 0,421.

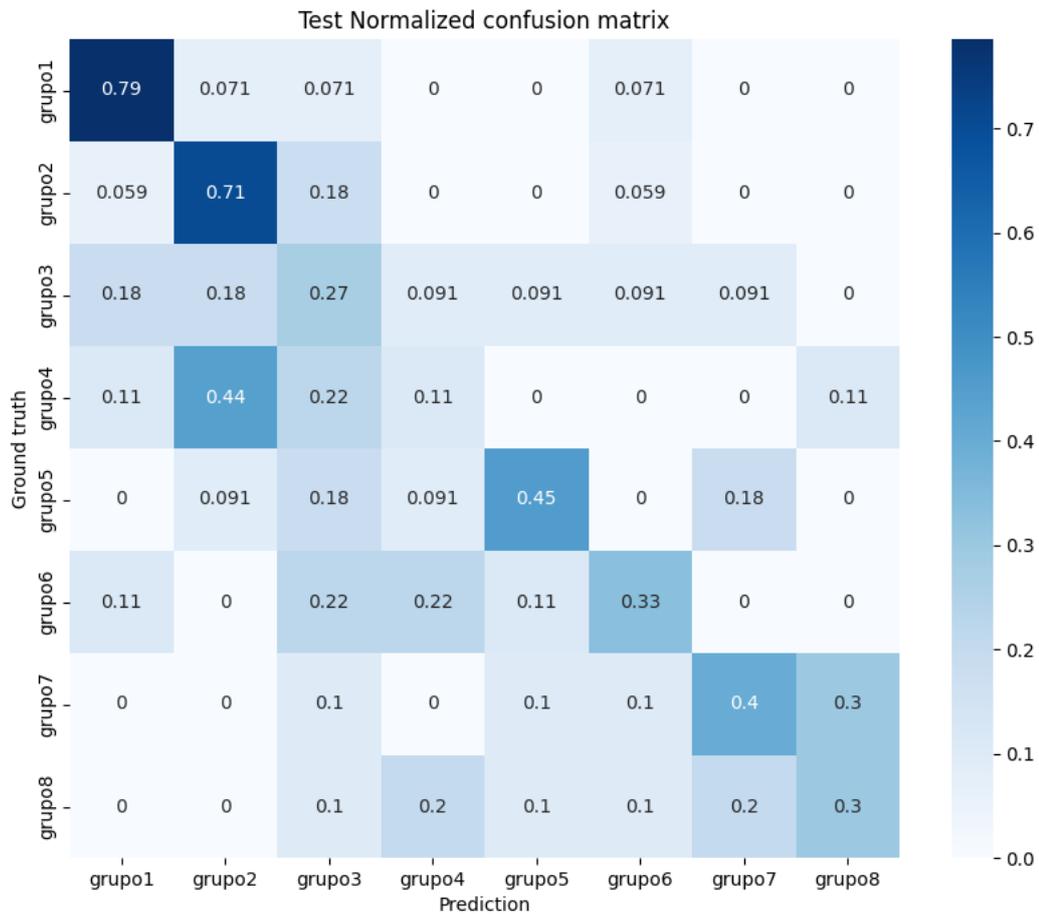


Figura 5.24: Matriz de confusão normalizada do conjunto de teste para o corte concatenado, modelo *Small* com a melhor semente aleatória identificada.

Adicionalmente, na Figura 5.25, são apresentados os detalhes do seu treinamento em que a última época foi a 345 e a função de perda para o conjunto de treino teve seu menor valor em 0,157. Já a função de perda para o conjunto de validação obteve o menor valor em 1,723. Para este mesmo conjunto, é possível observar a variação dos valores de acurácia, onde a acurácia *Top 1* atingiu seu ponto máximo em 0,56. Ademais, a acurácia *Top 5* teve o seu valor mais alto em 0,987.

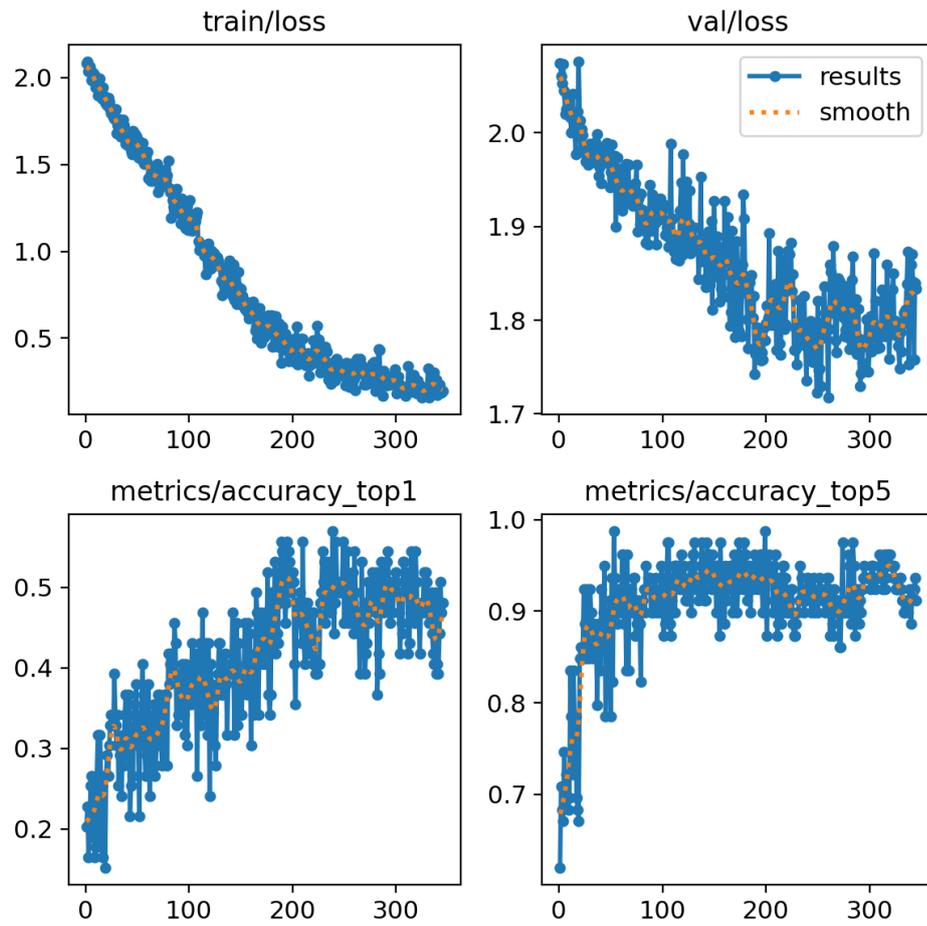


Figura 5.25: Resultados do treinamento para o corte concatenado, modelo *Small* com a melhor semente aleatória identificada.

Com relação aos resultados de inferência do conjunto de teste para a pior semente aleatória identificada, os valores de cada critério foram: Acurácia, 0,385; *Recall* médio, 0,359; Precisão média, 0,377; *F-measure* médio, 0,359.

### 5.5.3 Modelo *Medium*

A Figura 5.26 apresenta a matriz de confusão, exibindo os resultados da inferência no conjunto de teste para o modelo *Medium* com a melhor semente aleatória identificada. Além disso, os valores dos critérios de avaliação associados são os seguintes: Acurácia, 0,473; *Recall* médio, 0,431; Precisão média, 0,420; *F-measure* médio, 0,414.

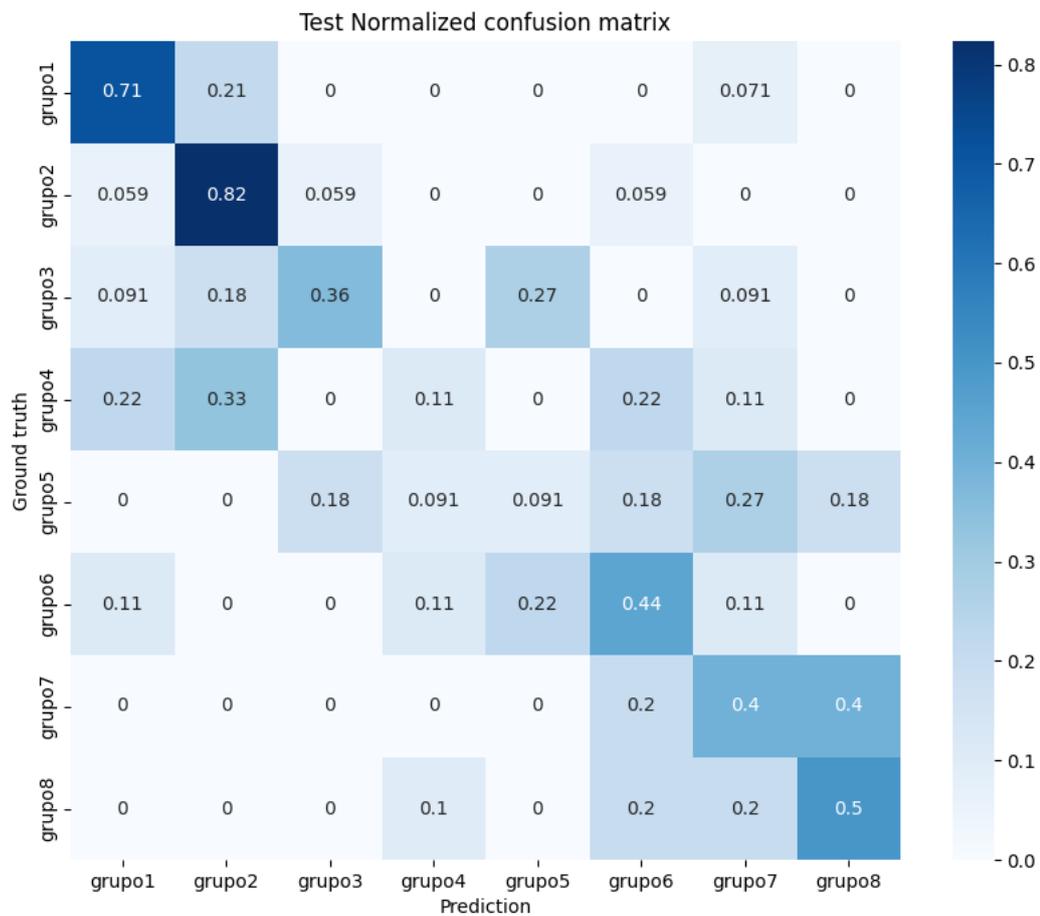


Figura 5.26: Matriz de confusão normalizada do conjunto de teste para o corte concatenado, modelo *Medium* com a melhor semente aleatória identificada.

Adicionalmente, na Figura 5.27, são apresentados os detalhes do seu treinamento em que a última época foi a 840 e a função de perda para o conjunto de treino teve seu menor valor em 0,045. Já a função de perda para o conjunto de validação obteve o menor valor em 1,722. Para este mesmo conjunto, é possível observar a variação dos valores de acurácia, onde a acurácia *Top 1* atingiu seu ponto máximo em 0,56. Ademais, a acurácia *Top 5* teve o seu valor mais alto em 0,987.

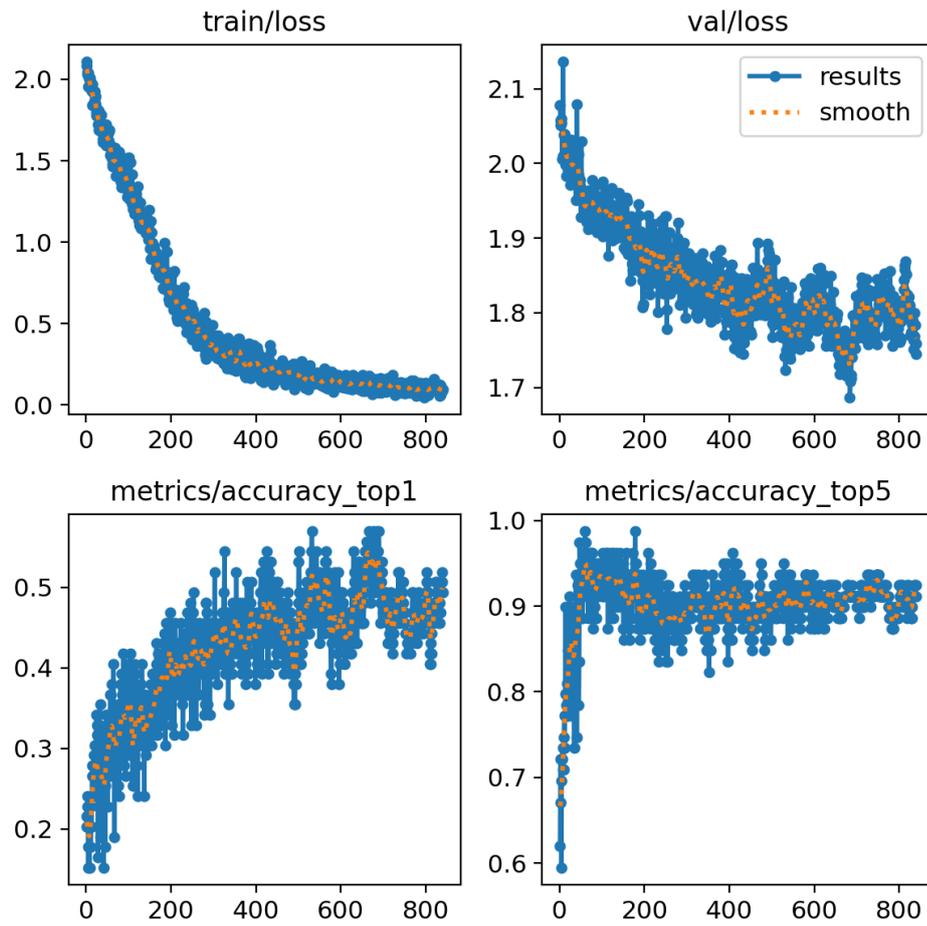


Figura 5.27: Resultados do treinamento para o corte concatenado, modelo *Medium* com a melhor semente aleatória identificada.

Com relação aos resultados de inferência do conjunto de teste para a pior semente aleatória identificada, os valores de cada critério foram: Acurácia, 0,352; *Recall* médio, 0,325; Precisão média, 0,358; *F-measure* médio, 0,322.

#### 5.5.4 Modelo *Large*

A Figura 5.28 apresenta a matriz de confusão, exibindo os resultados da inferência no conjunto de teste para o modelo *Large* com a melhor semente aleatória identificada. Além disso, os valores dos critérios de avaliação associados são os seguintes: Acurácia, 0,451; *Recall* médio, 0,438; Precisão média, 0,439; *F-measure* médio, 0,429.

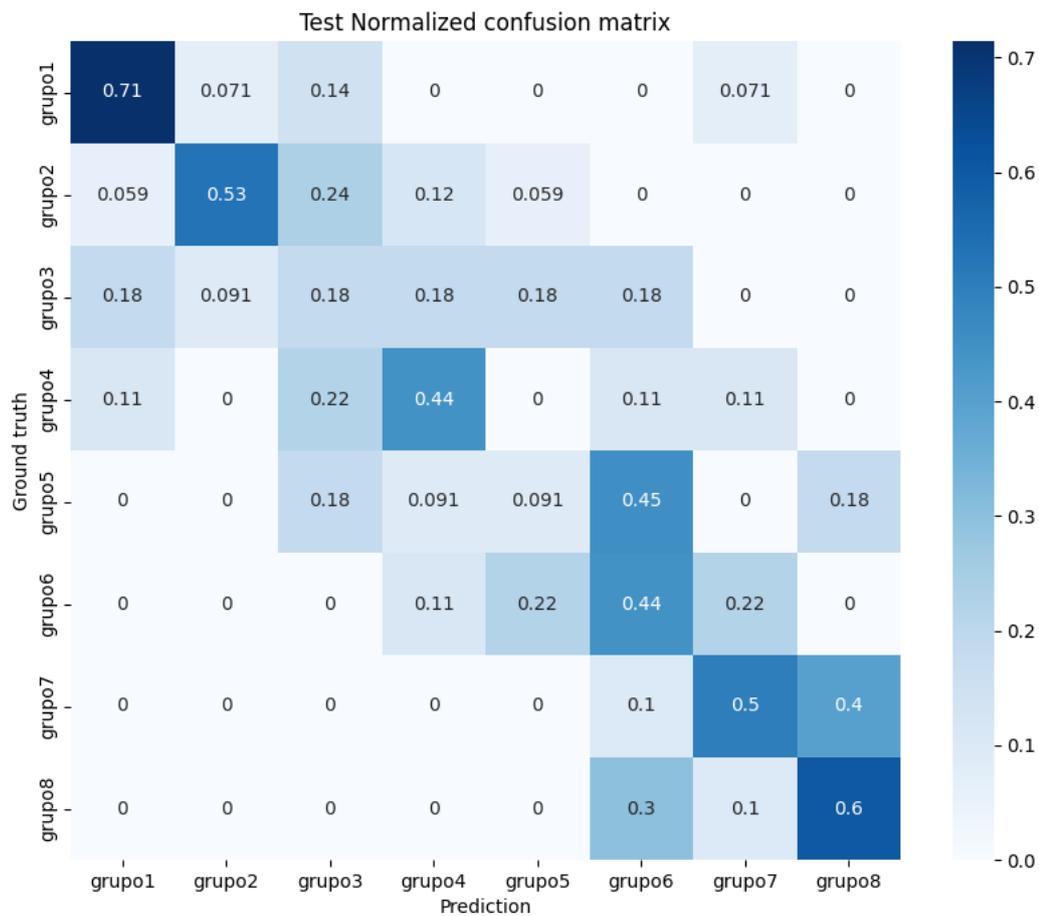


Figura 5.28: Matriz de confusão normalizada do conjunto de teste para o corte concatenado, modelo *Large* com a melhor semente aleatória identificada.

Adicionalmente, na Figura 5.29, são apresentados os detalhes do seu treinamento em que a última época foi a 500 e a função de perda para o conjunto de treino teve seu menor valor em 0,164. Já a função de perda para o conjunto de validação obteve o menor valor em 1,791. Para este mesmo conjunto, é possível observar a variação dos valores de acurácia, onde a acurácia *Top 1* atingiu seu ponto máximo em 0,51. Ademais, a acurácia *Top 5* teve o seu valor mais alto em 0,987.

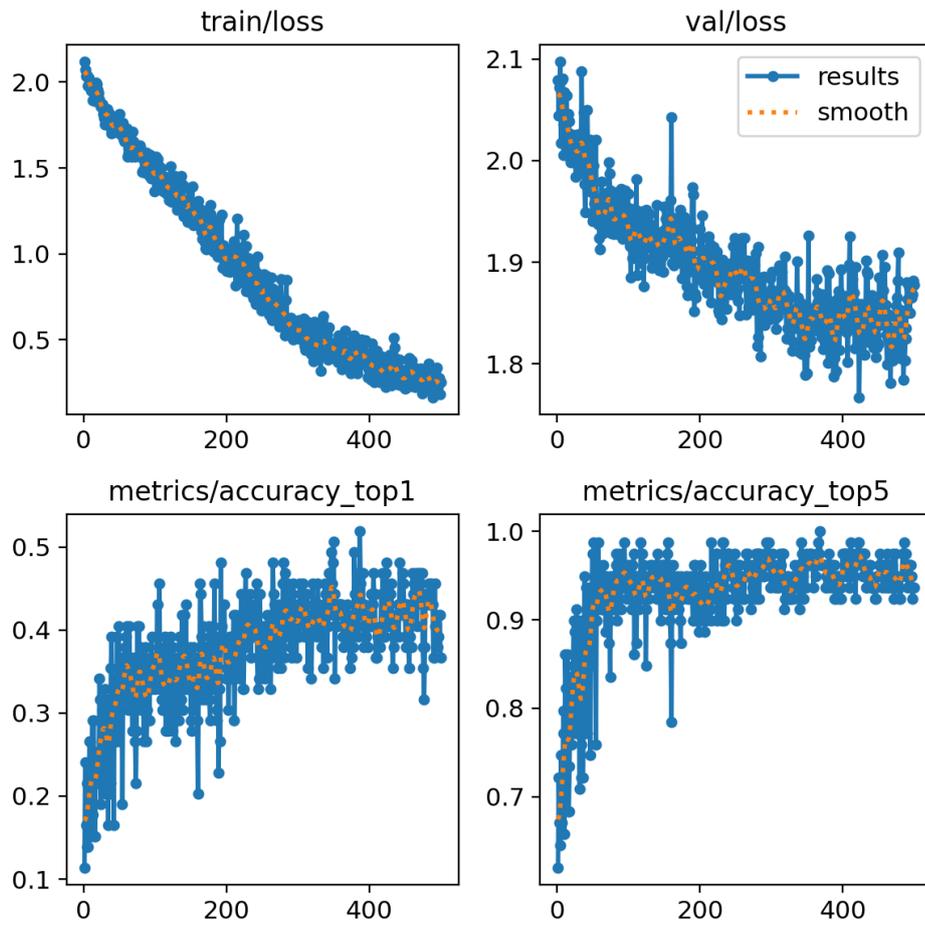


Figura 5.29: Resultados do treinamento para o corte concatenado, modelo *Large* com a melhor semente aleatória identificada.

Com relação aos resultados de inferência do conjunto de teste para a pior semente aleatória identificada, os valores de cada critério foram: Acurácia, 0,308; *Recall* médio, 0,278; Precisão média, 0,251; *F-measure* médio, 0,246.

### 5.5.5 Modelo *Extra Large*

A Figura 5.30 apresenta a matriz de confusão, exibindo os resultados da inferência no conjunto de teste para o modelo *Extra Large* com a melhor semente aleatória identificada. Além disso, os valores dos critérios de avaliação associados são os seguintes: Acurácia, 0,473; *Recall* médio, 0,443; Precisão média, 0,450; *F-measure* médio, 0,441.

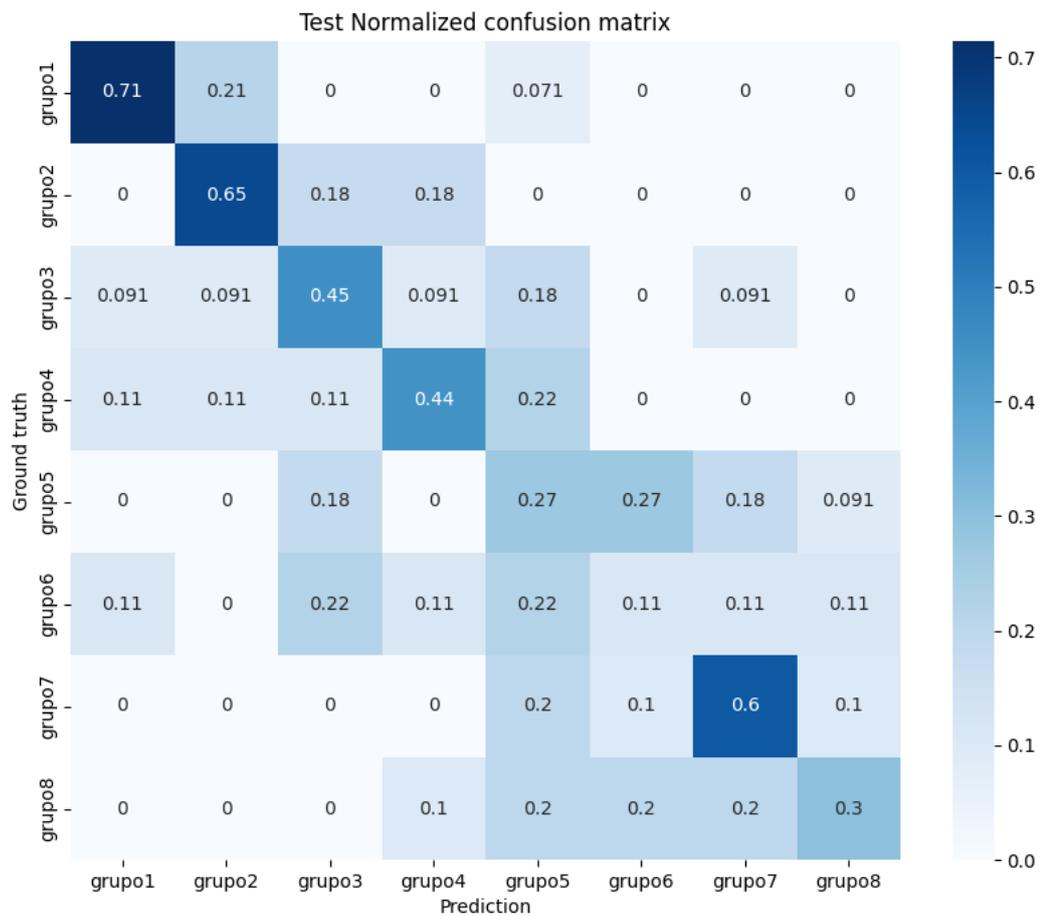


Figura 5.30: Matriz de confusão normalizada do conjunto de teste para o corte concatenado, modelo *Extra Large* com a melhor semente aleatória identificada.

Adicionalmente, na Figura 5.31, são apresentados os detalhes do seu treinamento em que a última época foi a 589 e a função de perda para o conjunto de treino teve seu menor valor em 0,137. Já a função de perda para o conjunto de validação obteve o menor valor em 1,778. Para este mesmo conjunto, é possível observar a variação dos valores de acurácia, onde a acurácia *Top 1* atingiu seu ponto máximo em 0,51. Ademais, a acurácia *Top 5* teve o seu valor mais alto em 0,987.

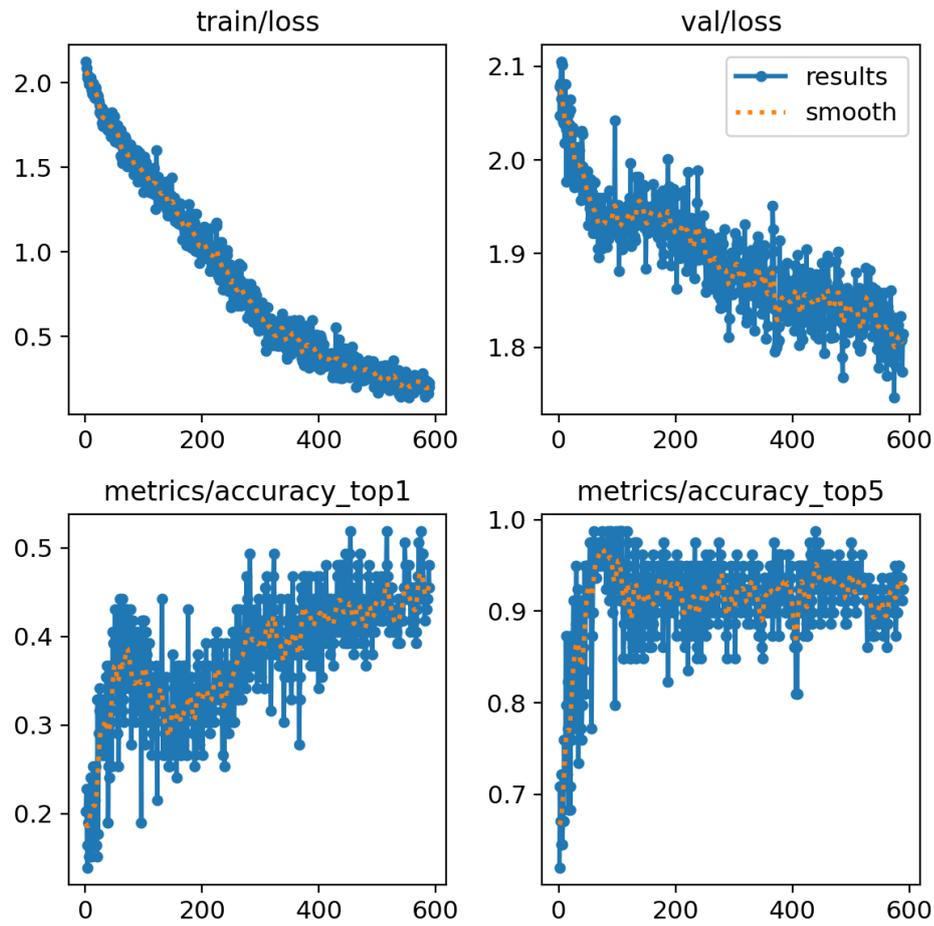


Figura 5.31: Resultados do treinamento para o corte concatenado, modelo *Extra Large* com a melhor semente aleatória identificada.

Com relação aos resultados de inferência do conjunto de teste para a pior semente aleatória identificada, os valores de cada critério foram: Acurácia, 0,330; *Recall* médio, 0,292; Precisão média, 0,258; *F-measure* médio, 0,269.

## 5.6 Comparação dos Modelos

### 5.6.1 Corte Coronal

Na Tabela 5.1, é possível realizar uma comparação entre as 5 versões de modelo para o corte coronal. Ela apresenta os resultados de acurácia, *Recall* Médio, Precisão Média e *F-Measure* Médio. Além disso, são apresentadas as médias e desvios padrão destes critérios, calculados a partir de 10 experimentos distintos, cada um conduzido com suas respectivas sementes aleatórias para cada uma das cinco versões. Ademais, os dados desses

experimentos foram obtidos utilizando o conjunto de teste do conjunto de dados.

Pode-se observar que o modelo *Nano* obteve a maior média para a Acurácia e *Recall* Médio, com os valores de 0,36 e 0,319, respectivamente. Já o modelo *Small* atingiu a maior média para a Precisão Média e *F-Measure* Médio, com os valores de 0,313 e 0,298, respectivamente.

Versão do Modelo	Acurácia		<i>Recall</i> Médio		Precisão Média		<i>F-Measure</i> Médio	
	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão
<i>Nano</i>	<b>0,36</b>	0,039	<b>0,319</b>	0,035	0,3	0,039	0,294	0,033
<i>Small</i>	0,358	0,053	0,316	0,049	<b>0,313</b>	0,041	<b>0,298</b>	0,045
<i>Medium</i>	0,347	<b>0,031</b>	0,310	0,033	<b>0,313</b>	<b>0,034</b>	0,287	<b>0,032</b>
<i>Large</i>	0,341	0,044	0,304	0,043	0,300	0,082	0,286	0,049
<i>Extra Large</i>	0,341	0,034	0,304	<b>0,029</b>	0,308	0,041	0,291	0,036

Tabela 5.1: Comparação entre as versões do modelo para o Corte Coronal

Na Figura 5.32, há os *boxplots* de cada versão de modelo, apresentando a variação dos valores de suas acurácias nos 10 experimentos. A caixa do *boxplot* representa onde a maioria dos dados está concentrada, enquanto a linha dentro dessa caixa representa a mediana, indicando uma tendência central dos dados. Já as linhas fora da caixa representam os valores máximo e mínimo, sendo que pontos além dessas linhas são considerados *outliers* e seguem um padrão diferente do restante dos dados.

Ademais, é possível observar como os modelos *Nano* e *Small* variam bastante em seus valores de acurácia, mas tendem a apresentar os maiores resultados, capazes de ultrapassar 0,425. Os modelos restantes tendem a apresentar valores reduzidos, sendo que o *Medium* e *Extra Large* têm menor variação de valores, além de estarem abaixo de 0,4.

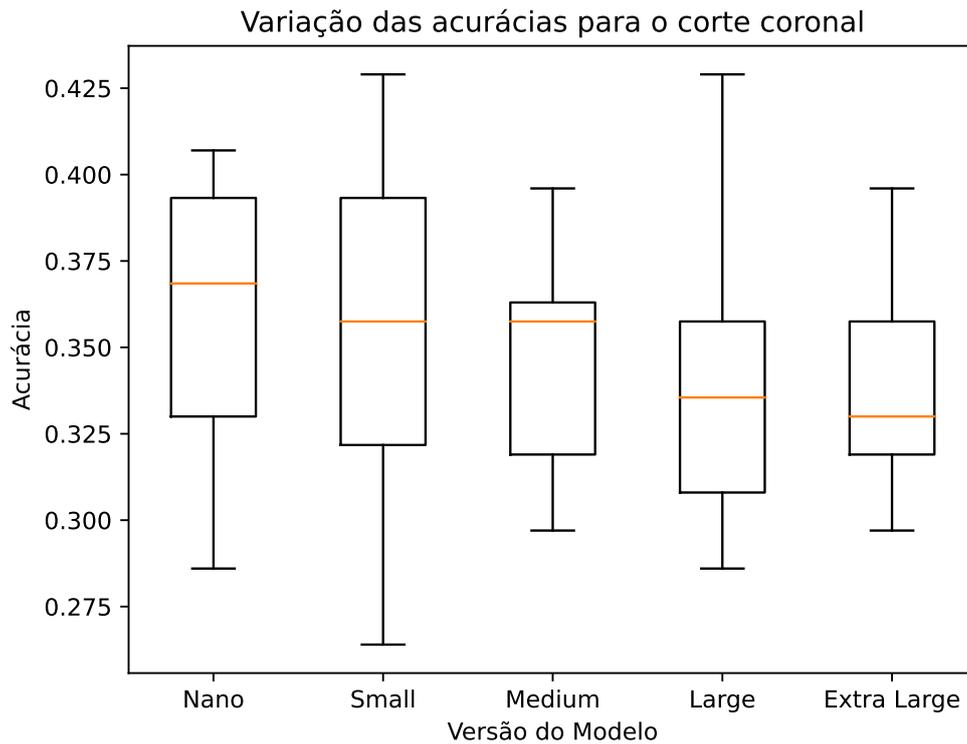


Figura 5.32: *Boxplot* apresentando a variação das Acurácias de cada modelo a partir das 10 sementes aleatórias para o corte coronal.

Na Figura 5.33, há os *boxplots* de cada versão de modelo, apresentando a variação dos valores de *Recall* Médio em seus 10 experimentos. Além disso, é possível observar como os modelos *Nano* e *Small* variam consideravelmente em seus resultados, oscilando de 0,24 até 0,38. Porém, juntamente com o modelo *Medium*, tendem a apresentar os maiores resultados. Os modelos restantes, *Large* e *Extra Large*, apresentam valores mais baixos, inferiores a 0,32, e pouca variação entre eles.

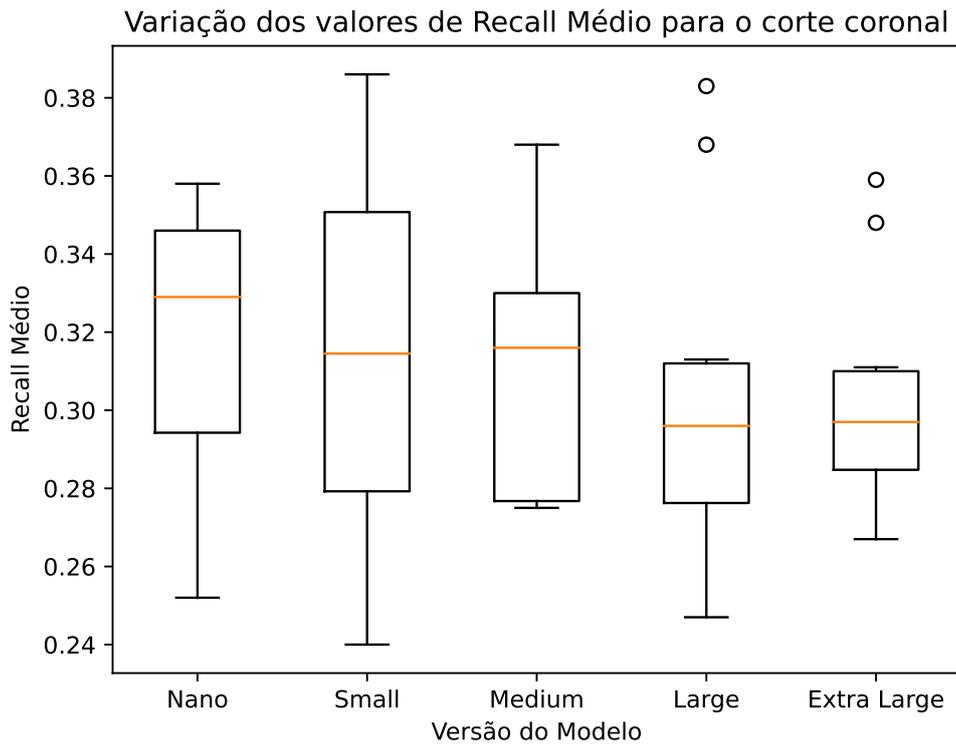


Figura 5.33: *Boxplot* apresentando a variação dos valores de *Recall* Médio de cada modelo a partir das 10 sementes aleatórias para o corte coronal.

Na Figura 5.34, há os *boxplots* de cada versão do modelo, apresentando a variação dos valores de *Precisão* Média em seus 10 experimentos. Adicionalmente, é possível observar uma tendência geral dos modelos de apresentarem precisões até um valor de 0,35, apesar de existir um *outlier* para o modelo *Large* com precisão média acima de 0,45.

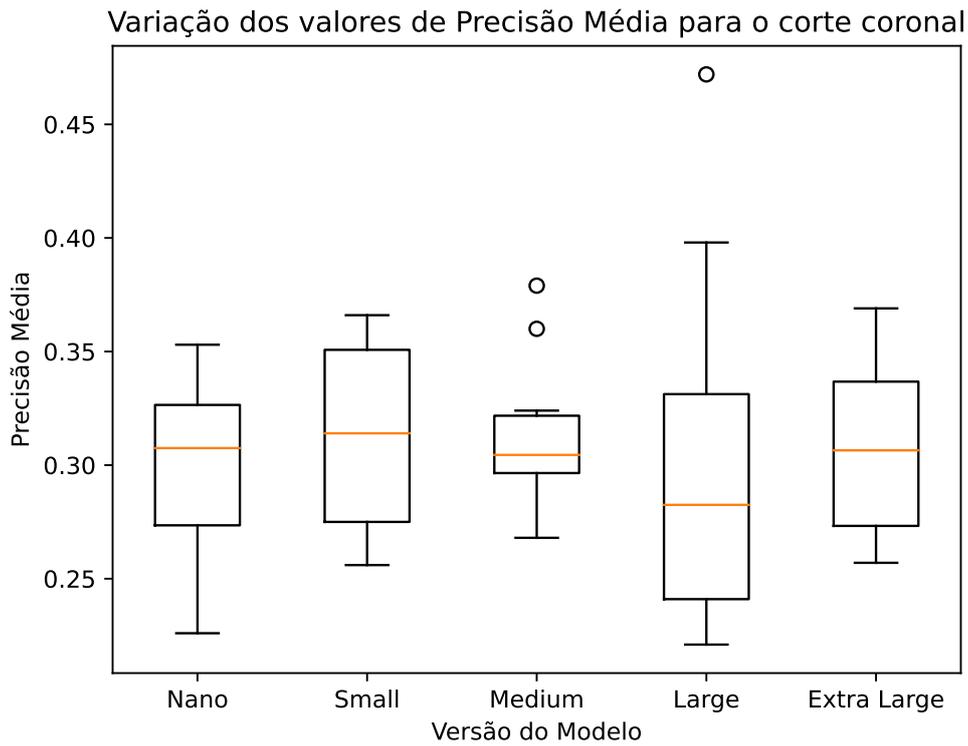


Figura 5.34: *Boxplot* apresentando a variação dos valores de Precisão Média de cada modelo a partir das 10 sementes aleatórias para o corte coronal.

Na Figura 5.35, há os *boxplots* de cada versão do modelo, apresentando a variação dos valores de *F-Measure* Médio de seus 10 experimentos. Adicionalmente, é possível observar que todos os modelos têm variações semelhantes, com resultados variando de 0,24 até 0,36.

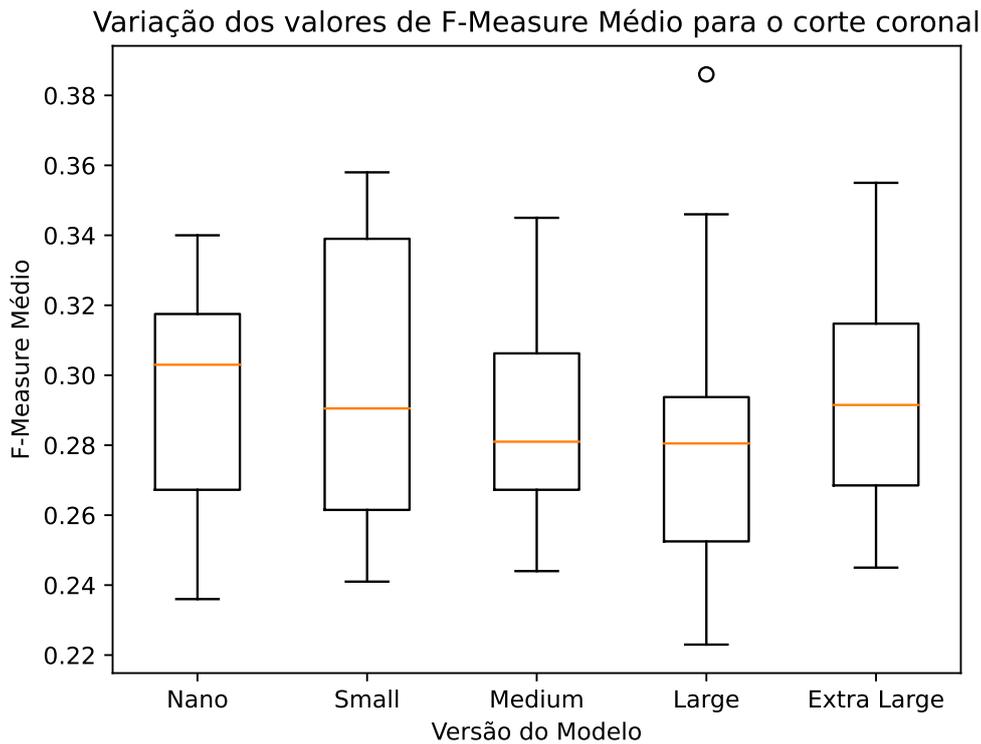


Figura 5.35: *Boxplot* apresentando a variação dos valores de *F-Measure* Médio de cada modelo a partir das 10 sementes aleatórias para o corte coronal.

### 5.6.2 Corte Sagital

Na Tabela 5.2, é possível realizar uma comparação entre as 5 versões de modelo para o corte sagital. Ela apresenta os resultados de acurácia, *Recall* Médio, *Precisão* Média e *F-Measure* Médio. Além disso, são apresentadas as médias e desvios padrão destes critérios, calculados a partir de 10 experimentos distintos, cada um conduzido com sementes aleatórias para cada uma das cinco versões.

O modelo *Small* obteve a maior média para a Acurácia e *Recall* Médio, com os valores de 0,455 e 0,425, respectivamente. Além disso, também atingiu a maior média para a *Precisão* Média e *F-Measure* Médio, com os valores de 0,448 e 0,418, respectivamente. Ademais, seus valores de desvio padrão também foram os menores em todos os critérios, exceto para a acurácia, em que seu valor foi de 0,035, sendo o menor 0,03 para o modelo *Large*.

Na Figura 5.36, há os *boxplots* de cada versão do modelo, apresentando a variação dos valores de suas acurácias nos 10 experimentos.

Versão do Modelo	Acurácia		Recall Médio		Precisão Média		F-Measure Médio	
	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão
<i>Nano</i>	0,417	0,054	0,394	0,057	0,421	0,058	0,392	0,058
<b><i>Small</i></b>	<b>0,455</b>	0,035	<b>0,425</b>	<b>0,034</b>	<b>0,448</b>	<b>0,044</b>	<b>0,418</b>	<b>0,039</b>
<i>Medium</i>	0,343	0,046	0,309	0,051	0,328	0,057	0,299	0,051
<i>Large</i>	0,342	<b>0,030</b>	0,310	<b>0,034</b>	0,291	0,051	0,287	0,041
<i>Extra Large</i>	0,359	0,042	0,332	0,041	0,347	0,065	0,325	0,047

Tabela 5.2: Comparação entre as versões do modelo para o Corte Sagital

Além disso, é possível observar como os modelos *Nano* e *Small* possuem os maiores valores, variando cerca de 0,35 até 0,5. Os modelos restantes tendem a apresentar valores reduzidos, chegando no máximo por volta de 0,4, sendo que seus resultados tendem a concentrar em 0,35, além de possuírem pouca variação.

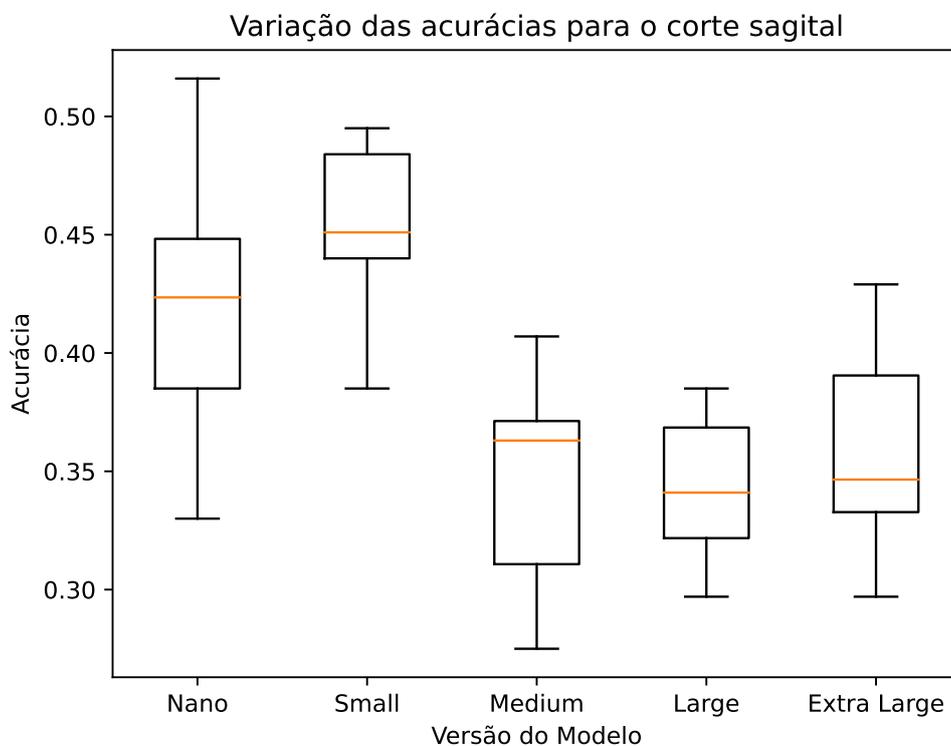


Figura 5.36: *Boxplot* apresentando a variação das Acurácias de cada modelo a partir das 10 sementes aleatórias para o corte sagital.

Na Figura 5.37, há os *boxplots* de cada versão do modelo, apresentando a variação dos valores de *Recall* Médio em seus 10 experimentos. Além disso, é possível observar como os modelos *Nano* e *Small* possuem os maiores valores, variando cerca de 0,3 até 0,5. Neste caso, a variação do modelo *Nano* é maior, enquanto o modelo *Small* concentra seus resultados em cerca de 0,43 com pouca variação.

Os modelos restantes tendem a apresentar valores reduzidos, chegando no máximo

por volta de 0,4, sendo que seus resultados tendem a concentrar em cerca de 0,33, além de possuírem pouca variação.

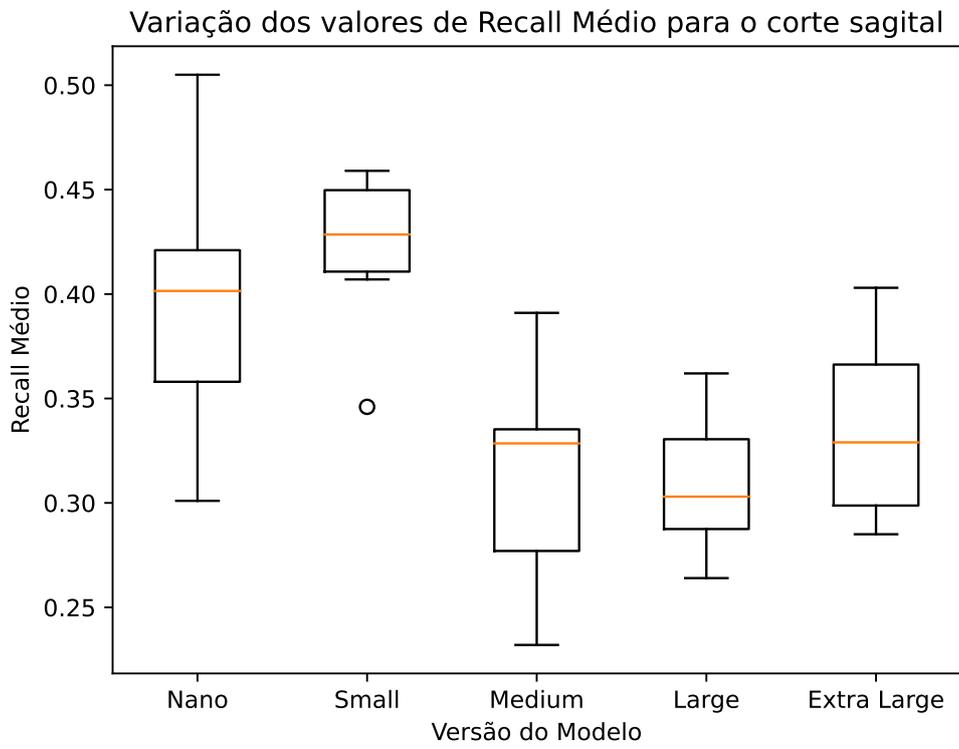


Figura 5.37: *Boxplot* apresentando a variação dos valores de *Recall Médio* de cada modelo a partir das 10 sementes aleatórias para o corte sagital.

Na Figura 5.38, há os *boxplots* de cada versão do modelo, apresentando a variação dos valores de *Precisão Média* em seus 10 experimentos. Além disso, é possível observar como os modelos *Nano* e *Small* possuem os maiores valores, variando cerca de 0,34 até 0,51. Neste caso, a variação do modelo *Nano* é maior, enquanto o modelo *Small* concentra seus resultados em cerca de 0,46 com pouca variação.

Os modelos restantes tendem a apresentar valores reduzidos, chegando no máximo por volta de 0,4, sendo que seus resultados tendem a concentrar em cerca de 0,32, além de possuírem variação parecida com a do modelo *Nano*.

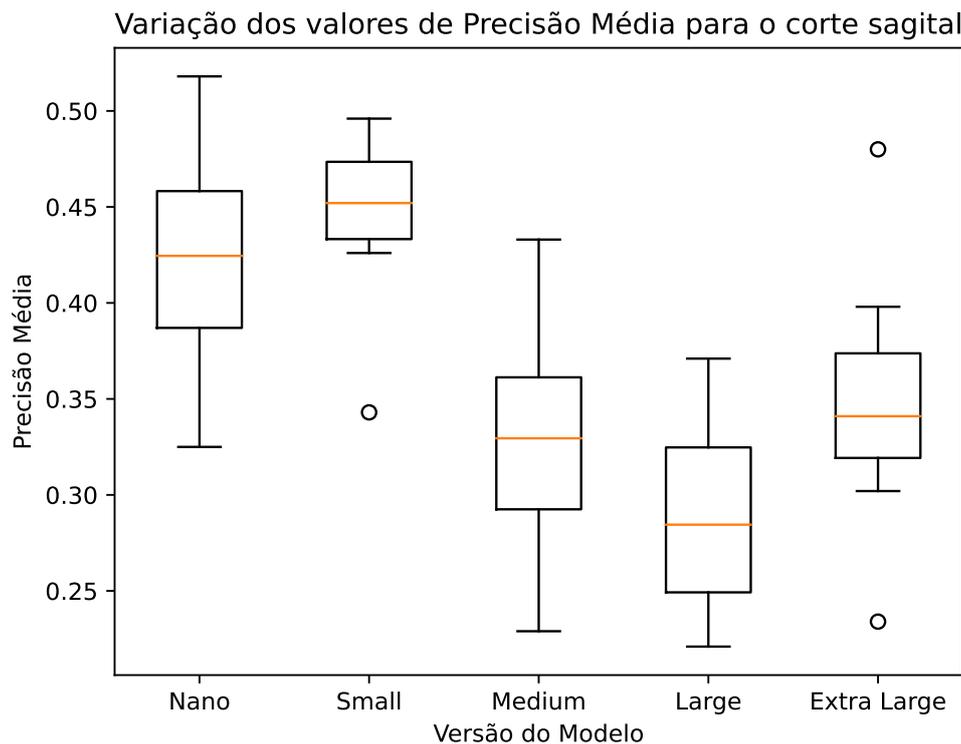


Figura 5.38: *Boxplot* apresentando a variação dos valores de Precisão Média de cada modelo a partir das 10 sementes aleatórias para o corte sagital.

Na Figura 5.39, há os *boxplots* de cada versão do modelo, apresentando a variação dos valores de *F-Measure* Médio em seus 10 experimentos. Além disso, é possível observar como os modelos *Nano* e *Small* possuem os maiores valores, variando cerca de 0,3 até 0,5. Neste caso, a variação do modelo *Nano* é maior, enquanto o modelo *Small* concentra seus resultados em cerca de 0,44 com pouca variação.

Os modelos restantes tendem a apresentar valores reduzidos, abaixo de 0,4, sendo que seus resultados tendem a concentrar em cerca de 0,3, além de possuírem variação de valores reduzida, próximas à variação do modelo *Small*.

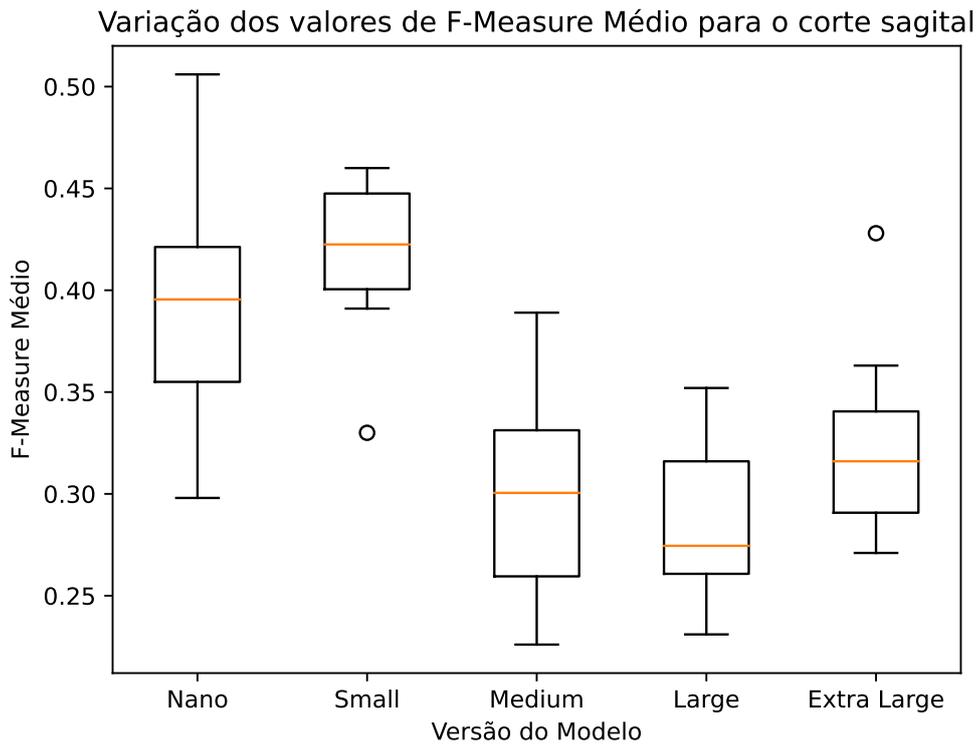


Figura 5.39: *Boxplot* apresentando a variação dos valores de *F-Measure* Médio de cada modelo a partir das 10 sementes aleatórias para o corte sagital.

### 5.6.3 Corte Concatenado

Na Tabela 5.3, é possível realizar uma comparação entre as 5 versões de modelo para o corte concatenado. Ela apresenta os resultados de acurácia, *Recall* Médio, Precisão Média e *F-Measure* Médio. Além disso, são apresentadas as médias e desvios padrão destes critérios, calculados a partir de 10 experimentos distintos, cada um conduzido com sementes aleatórias para cada uma das cinco versões.

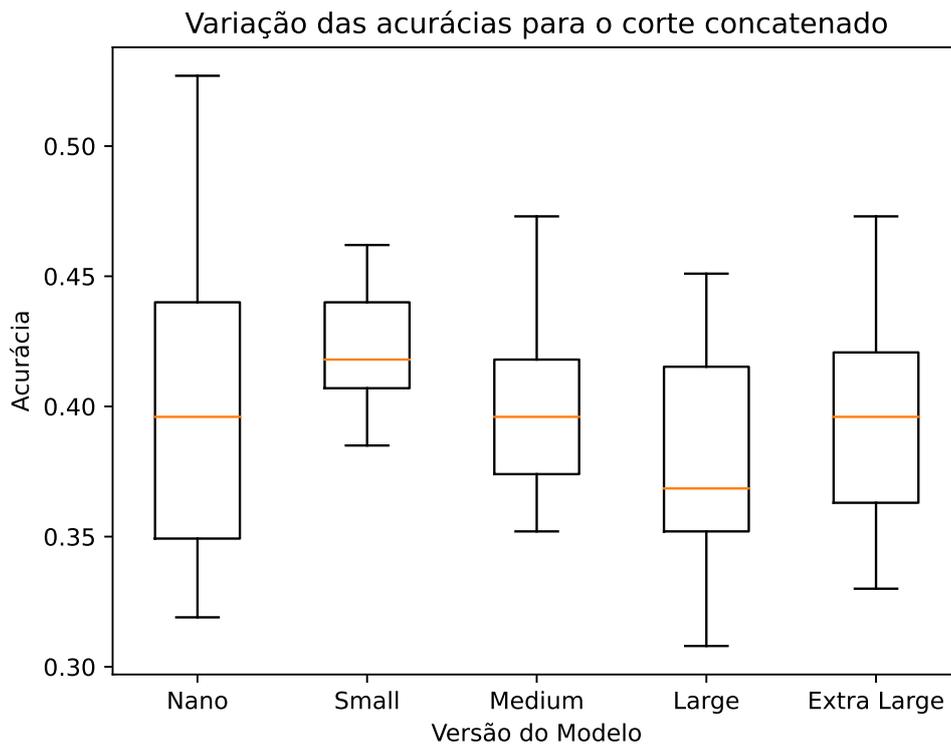
O modelo *Small* obteve a maior média para a Acurácia e *Recall* Médio, com os valores de 0,421 e 0,394, respectivamente. Além disso, também atingiu a maior média para a Precisão Média e *F-Measure* Médio, com os valores de 0,404 e 0,384, respectivamente. Ademais, seus valores de desvio padrão também foram os menores em todos os critérios, exceto para a Precisão Média, em que seu valor foi de 0,035. Porém, este valor foi muito próximo ao menor 0,034 para o modelo *Medium*.

Na Figura 5.40, há os *boxplots* de cada versão de modelo, apresentando a variação dos valores de suas acurácias nos 10 experimentos.

Versão do Modelo	Acurácia		Recall Médio		Precisão Média		F-Measure Médio	
	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão
<i>Nano</i>	0,403	0,066	0,367	0,067	0,376	0,082	0,359	0,069
<b><i>Small</i></b>	<b>0,421</b>	<b>0,024</b>	<b>0,394</b>	<b>0,023</b>	<b>0,404</b>	0,035	<b>0,384</b>	<b>0,027</b>
<i>Medium</i>	0,399	0,039	0,371	0,032	0,398	<b>0,034</b>	0,365	0,032
<i>Large</i>	0,381	0,046	0,354	0,050	0,361	0,065	0,339	0,060
<i>Extra Large</i>	0,392	0,044	0,361	0,046	0,362	0,056	0,348	0,051

Tabela 5.3: Comparação entre as versões do modelo para o Corte Concatenado

Ademais, é possível observar como o modelo *Nano* varia consideravelmente em seus valores, oscilando de 0,33 até 0,53. Enquanto isso, o modelo *Small* é o que tem a tendência de apresentar os melhores resultados, com seus valores concentrados em cerca de 0,43 e possuindo a menor variação de todas. Os modelos restantes tendem a apresentar valores reduzidos, com a tendência a se concentrar em 0,39, oscilando em cerca de 0,32 até 0,45.

Figura 5.40: *Boxplot* apresentando a variação das Acurácias de cada modelo a partir das 10 sementes aleatórias para o corte concatenado.

Na Figura 5.41, há os *boxplots* de cada versão de modelo, apresentando a variação dos valores de *Recall* Médio nos 10 experimentos.

Ademais, é possível observar como o modelo *Nano* varia consideravelmente em seus valores, oscilando de 0,30 até 0,49. Enquanto isso, o modelo *Small* é o que tem a

tendência de apresentar os melhores resultados, com seus valores concentrados em cerca de 0,4 e possuindo a menor variação de todas. Os modelos restantes tendem a apresentar valores reduzidos, com a tendência a se concentrar em 0,36, oscilando em cerca de 0,28 até 0,45.

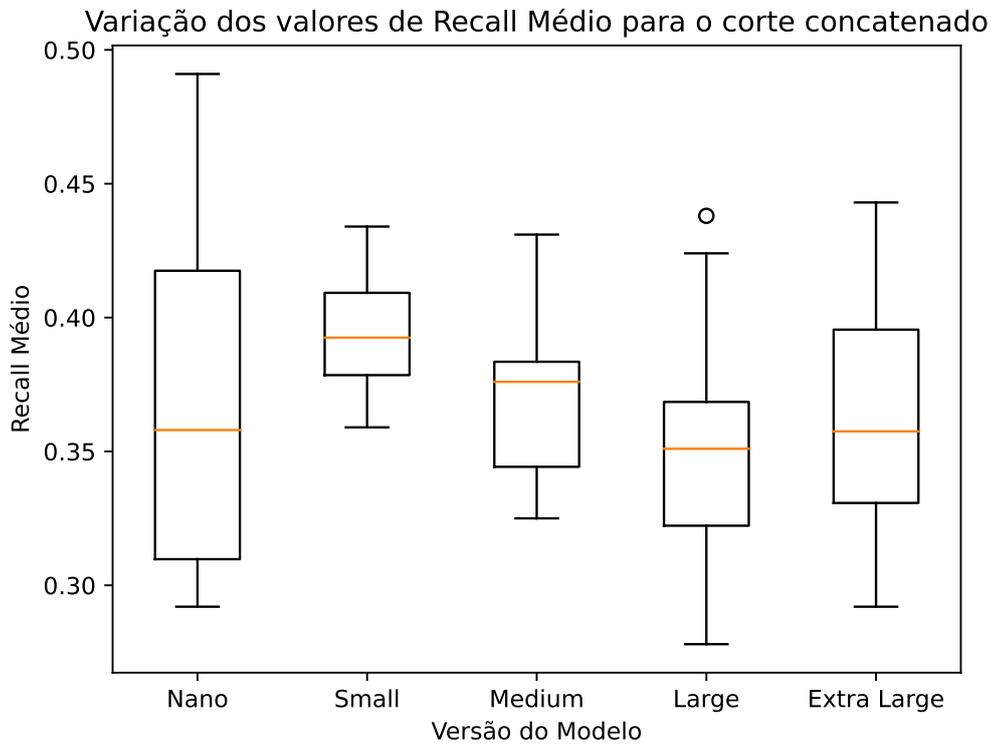


Figura 5.41: *Boxplot* apresentando a variação dos valores de *Recall* Médio de cada modelo a partir das 10 sementes aleatórias para o corte concatenado.

Na Figura 5.42, há os *boxplots* de cada versão de modelo, apresentando a variação dos valores de *Precisão* Média nos 10 experimentos.

Ademais, é possível observar como o modelo *Nano* varia consideravelmente em seus valores, oscilando de 0,29 até 0,52. Enquanto isso, os modelos *Small* e *Medium* são os que tem a tendência de apresentar os melhores resultados, com seus valores concentrados em cerca de 0,4 e possuindo as menores variações. Os modelos restantes tendem a apresentar valores reduzidos, com a tendência a se concentrar em 0,36, oscilando em cerca de 0,25 até 0,45.

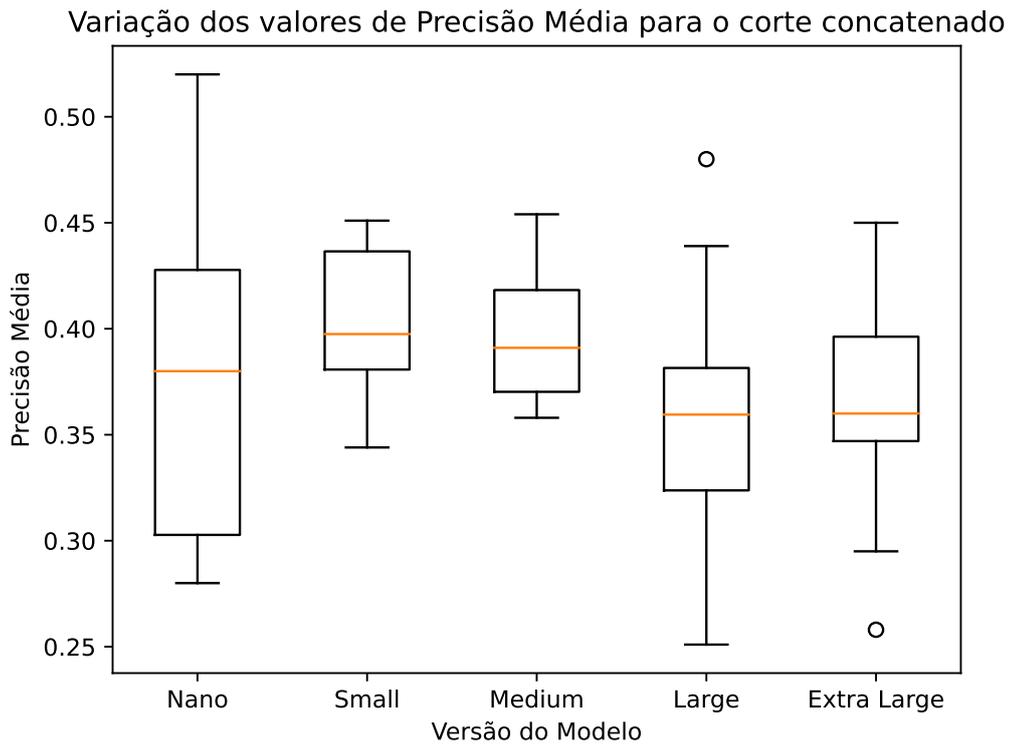


Figura 5.42: *Boxplot* apresentando a variação dos valores de Precisão Média de cada modelo a partir das 10 sementes aleatórias para o corte concatenado.

Na Figura 5.43, há os *boxplots* de cada versão de modelo, apresentando a variação dos valores de *F-Measure* Médio nos 10 experimentos.

Ademais, é possível observar como o modelo *Nano* varia consideravelmente em seus valores, oscilando de 0,3 até 0,5. Enquanto isso, os modelos *Small* e *Medium* são os que tem a tendência de apresentar os melhores resultados, com seus valores concentrados em cerca de 0,37 e possuindo as menores variações. Os modelos restantes tendem a apresentar valores reduzidos, com a tendência a se concentrar em 0,35, oscilando em cerca de 0,25 até 0,43.

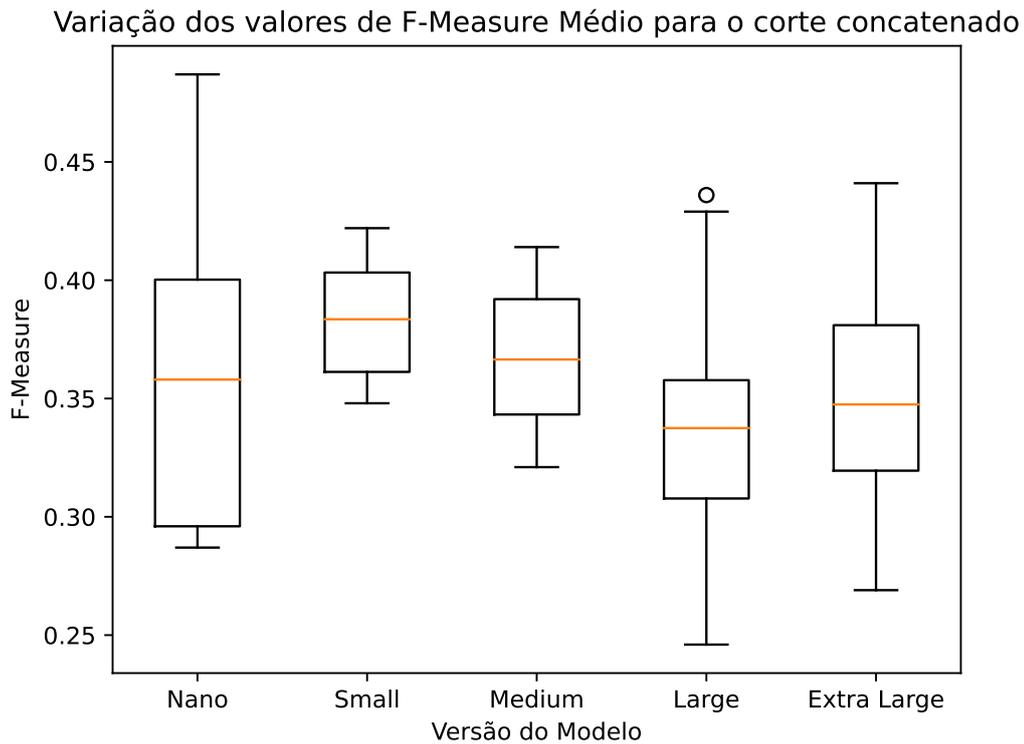


Figura 5.43: *Boxplot* apresentando a variação dos valores de *F-Measure* Médio de cada modelo a partir das 10 sementes aleatórias para o corte concatenado.

#### 5.6.4 Melhores Modelos de cada Corte

Para cada um dos três tipos de corte e dentre todos os experimentos realizados, houve modelos que apresentaram os maiores valores de acurácia. Esses modelos foram selecionados, e na Tabela 5.4, é possível observar uma comparação entre os resultados de seus critérios de avaliação.

O melhor modelo no geral foi a versão *Nano* para o corte concatenado, com uma acurácia de 0,527 e o melhor valor de Precisão Média de 0,52. O modelo para o corte sagital também teve acurácia acima de 0,5, com o resultado de 0,516, além de apresentar os melhores valores de *Recall* Médio e *F-Measure* Médio, sendo 0,505 e 0,506, respectivamente. Além disso, o melhor modelo para o corte coronal teve acurácia muito reduzida em relação aos demais, com o valor de 0,429, e o restante dos critérios abaixo de 0,39.

Melhores Modelos					
Tipo de Corte	Versão do Modelo	Acurácia	<i>Recall</i> Médio	Precisão Média	<i>F-Measure</i> Médio
Coronal	<i>Small</i>	0,429	0,386	0,364	0,342
Sagital	<i>Nano</i>	0,516	<b>0,505</b>	0,518	<b>0,506</b>
Concatenado	<b><i>Nano</i></b>	<b>0,527</b>	0,491	<b>0,520</b>	0,487

Tabela 5.4: Comparação entre os melhores modelos gerados para cada tipo de corte

## 5.7 Análise Visual

A fim de explicitar visualmente como os modelos tomam suas decisões de classificação das faixas de idade para cada um dos três cortes, a técnica EigenCAM foi utilizada. Esta é uma técnica muito utilizada em CNNs, criando mapas de destaque nas imagens e indicando quais são as áreas que mais influenciam na previsão do modelo (MUHAMMAD; YEASIN, 2020; MUHAMMAD; YEASIN, 2021). Além disso, essa abordagem busca superar limitações de outros métodos, como Grad-CAM, por exemplo, oferecendo vantagens como a independência em relação à classe, o que a torna capaz de fornecer explicações consistentes mesmo em situações de erro de predição (MUHAMMAD; YEASIN, 2021).

O código para a implementação da técnica EigenCAM na arquitetura do YOLOv8, utilizado nesta monografia, foi obtido a partir de um repositório de código aberto disponível no GitHub (rigvedrs, 2023). Os mapas de destaque a serem apresentados são de imagens pertencentes ao conjunto de teste e foram produzidos utilizando os melhores modelos para cada um dos três cortes.

### 5.7.1 Corte Coronal

A Figura 5.44 apresenta o mapa de destaque usando EigenCAM para o corte coronal. Em relação ao conjunto de teste do corte coronal, a maioria das imagens teve destaque nas regiões entre os dentes, conforme observado nas imagens 5.44c, 5.44d e 5.44f. Além disso, as estruturas dentárias adjacentes também obtiveram maior atenção pelo modelo em sua maioria: imagens 5.44b, 5.44d, 5.44e e 5.44f.

Por outro lado, a câmara pulpar do dente para o corte coronal não foi a área de maior foco pelo modelo. A quantidade de imagens que apresentaram destaque nessa região foi menor em comparação com as outras. Nas Figuras 5.44a e 5.44e, é possível observar certo destaque nessa região.

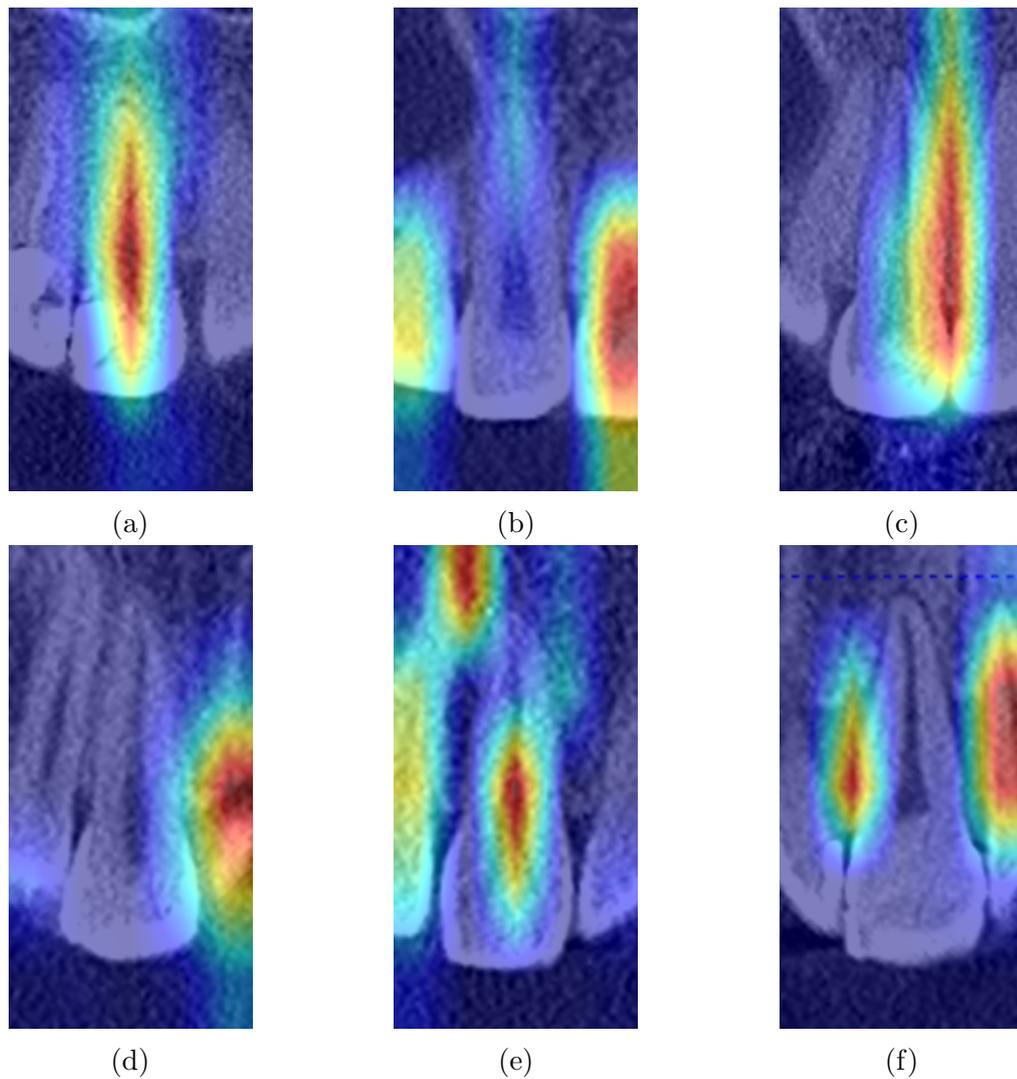


Figura 5.44: Mapa de destaque utilizando EigenCAM para imagens do corte coronal.

### 5.7.2 Corte Sagital

A Figura 5.45 apresenta o mapa de destaque usando EigenCAM para o corte sagital. Em relação ao conjunto de teste do corte sagital, a maioria das imagens teve destaque nas regiões da câmara pulpar e da coroa dentária. É possível observar como a câmara pulpar apresentou maior destaque nas imagens 5.45b e 5.45f, além de um destaque mais reduzido na imagem 5.45a. Além disso, estruturas dentárias próximas ao dente também receberam certa atenção do modelo em algumas imagens. Essas estruturas podem ser notadas nas imagens 5.45a, 5.45c e um pouco na imagem 5.45d.

Ademais, a câmara pulpar do dente para o corte sagital, mesmo nas imagens em que ela possui certo destaque, não foi a área em que o modelo focou unicamente. Nesses casos, observam-se também destaques em regiões próximas.

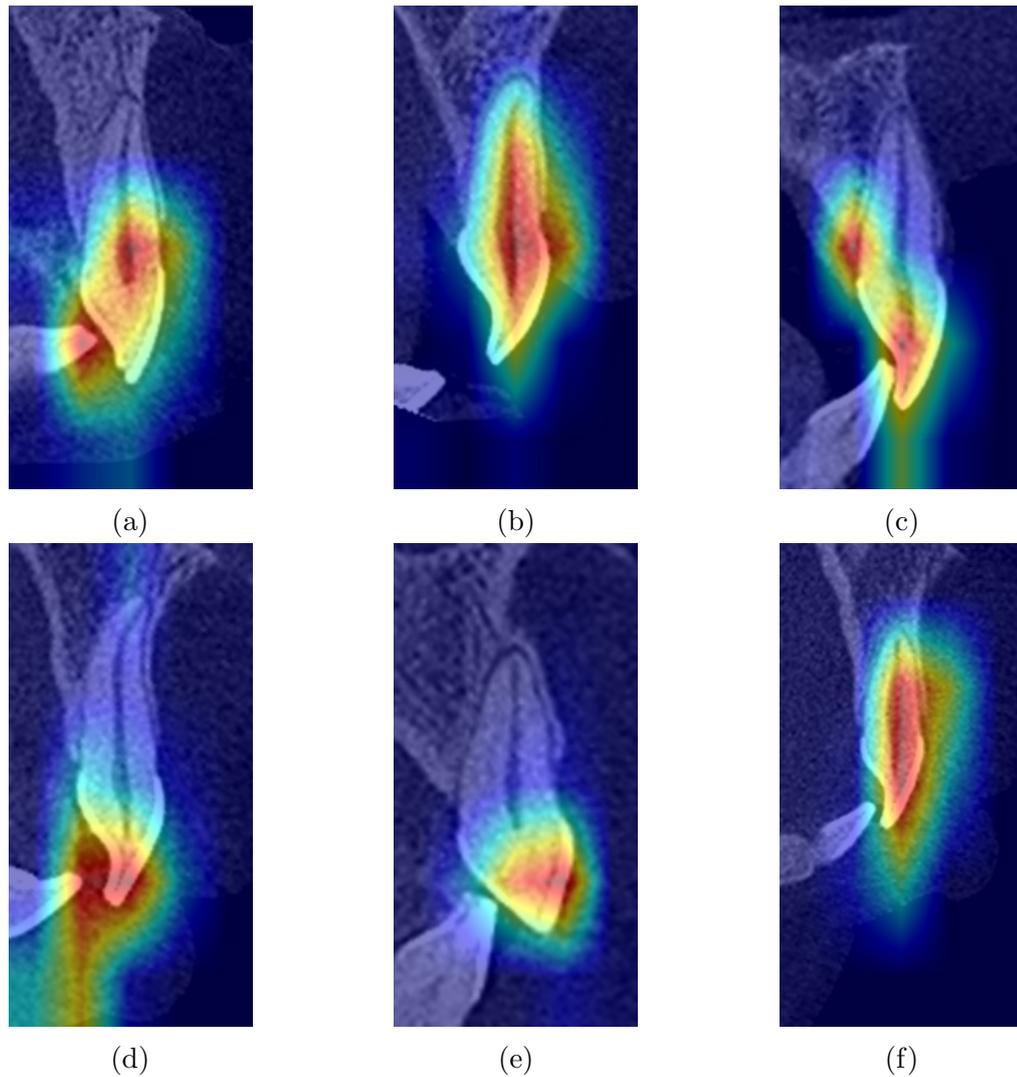


Figura 5.45: Mapa de destaque utilizando EigenCAM para imagens do corte sagital.

### 5.7.3 Corte Concatenado

A Figura 5.46 apresenta o mapa de destaque usando EigenCAM para o corte concatenado. Em relação ao conjunto de teste do corte concatenado, quase todas as imagens teve destaque maior na estrutura dentária correspondente ao corte sagital, sem dar ênfase à parte relacionada ao corte coronal. Ademais, de todas as imagens presentes no conjunto de teste, as Figuras 5.46d e 5.46f foram as únicas em que houve um certo destaque nas imagens correspondentes aos dois tipos de corte simultaneamente.

Além disso, nas imagens em que apenas a parte do corte sagital recebe destaque, é possível observar uma redução de foco na região da câmara pulpar. Uma exemplificação dessa redução, comparada às imagens de corte sagital da última subseção, pode ser observada nas Figuras 5.46a, 5.46b, 5.46e e 5.46g. Adicionalmente, estruturas dentárias

próximas ao dente também receberam certa atenção do modelo.

Desse modo, pode-se entender que o modelo considerou as imagens correspondentes ao corte sagital como as mais adequadas para apresentar características que auxiliam na previsão da faixa de idade.

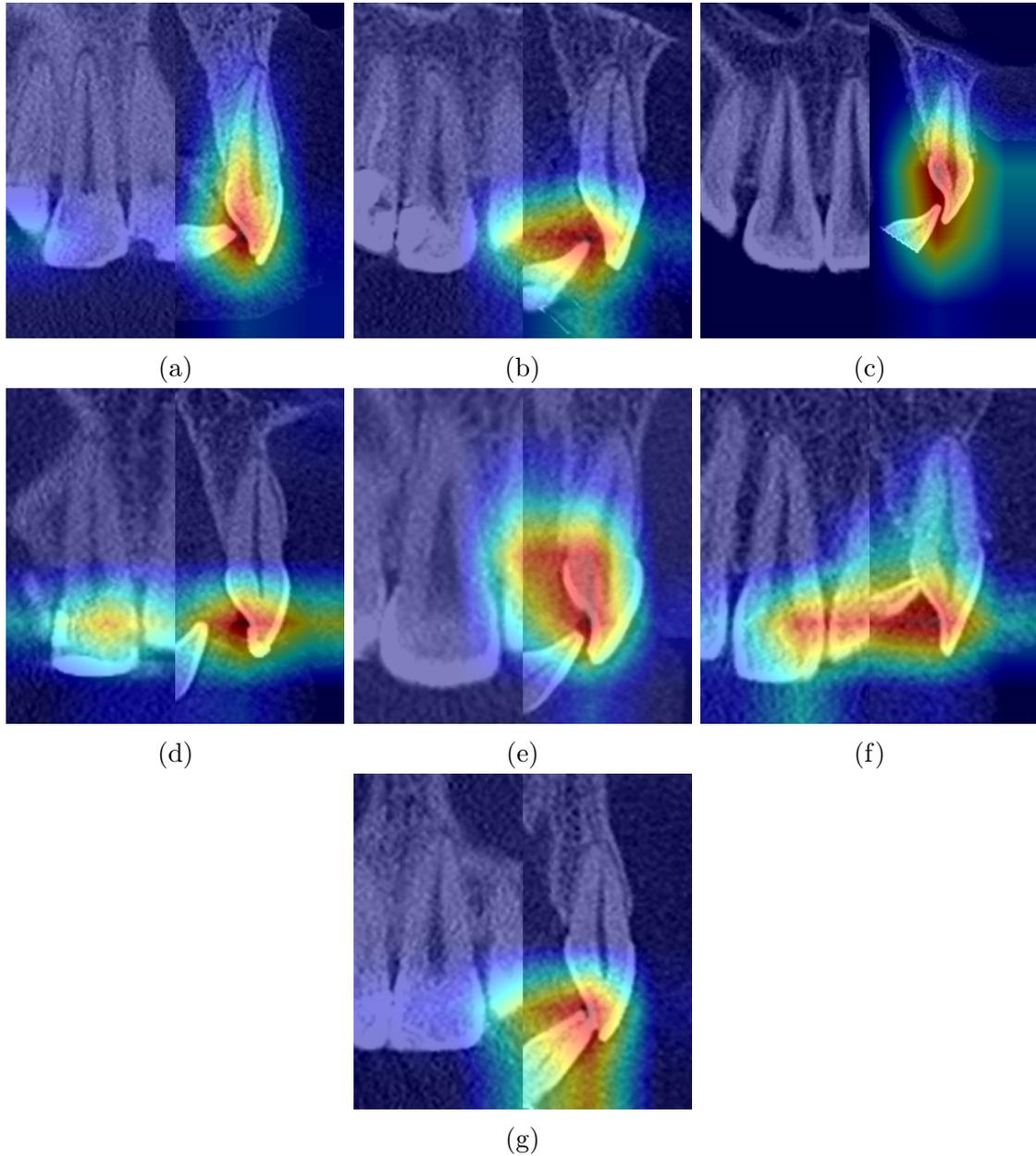


Figura 5.46: Mapa de destaque utilizando EigenCAM para imagens do corte concatenado.

## 5.8 Considerações Finais

Com os resultados apresentados, pode-se observar que o corte coronal é o que possui as menores acurácias para todas as versões de modelo treinadas. Uma das razões para

isso são imagens com uma menor nitidez da câmara pulpar, além de muitas estruturas dentárias próximas, o que faz com que o modelo não consiga focar na região de interesse para a determinação da faixa etária, que é a região da polpa.

Em contraste às imagens do corte coronal, as imagens do corte sagital possuem uma nitidez maior, além de o dente ser praticamente a única estrutura da imagem. Desse modo, a região de interesse, que é a câmara pulpar, é capaz de receber maior destaque pelos modelos. Na Figura 5.47, é possível observar essas diferenças entre as imagens de cada corte para um mesmo indivíduo. Na Figura 5.47b, nota-se como a região da câmara pulpar está mais explícita, além de haver menos estruturas adjacentes ao dente.



Figura 5.47: Imagem ilustrativa das diferenças entre os cortes coronais e sagitais.

Assim, o corte coronal foi o que obteve menor acurácia dentre todos, sendo o seu melhor valor 0,429. Para o corte sagital, a acurácia foi maior, com o valor de 0,516. Ademais, pode-se notar que uma junção dos dois cortes em uma única imagem, formando o corte concatenado, ainda fez com que os modelos buscassem informações para a classificação da faixa de idade na parte correspondente ao corte sagital.

Desse modo, não houve uma combinação simultânea de características das imagens dos cortes coronais e sagitais pelos modelos. Apesar disso, um experimento com a versão do modelo *Nano* utilizando o corte concatenado foi o que obteve o maior valor de acurácia dentre todos: 0,527.

Uma outra característica observada ao analisar os modelos dos cortes, de forma geral, é que versões mais simplificadas da arquitetura tiveram resultados melhores. Por exemplo, tanto para o corte sagital quanto para o corte concatenado, a versão *Small* apresentou os melhores resultados na média para os quatro critérios.

Além disso, para o corte coronal, a versão *Nano* obteve os melhores valores de média para Acurácia e *Recall* Médio. Ademais, em relação aos valores de Precisão Média e *F-Measure* Médio, a versão *Small* foi aquela que alcançou os melhores resultados.

Desse modo, as versões *Nano* e *Small* dos modelos se sobressaíram em relação às versões *Medium*, *Large* e *Extra Large*. Vale ressaltar que, em testes preliminares, redes mais complexas como a CoAtNet (DAI et al., 2021) também encontrou dificuldades e obteve acurácias menores do que as apresentadas até então.

## 6 Conclusões

Nesta monografia, foi explorado o uso de CNNs para a estimativa da idade de indivíduos com base em imagens da câmara pulpar dos dentes obtidas por meio de imagens de TCFC. São abordadas imagens que compreendem três tipos de cortes dos dentes incisivos centrais superiores: corte coronal, sagital e o corte concatenado, que realiza uma junção dos dois últimos cortes em uma única imagem.

Além disso, para atingir os objetivos de determinar as faixas etárias dos indivíduos, explorou-se diferentes versões do modelo da arquitetura YOLOv8, assim como diversos hiperparâmetros em treinamentos preliminares.

Desse modo, a partir dos experimentos realizados, o corte coronal teve seu melhor valor de acurácia em 0,429, o corte sagital atingiu seu valor máximo de acurácia em 0,516, e o corte concatenado obteve o valor de 0,527, sendo o maior alcançado.

Adicionalmente, nota-se que as versões de modelo menos complexas (*Nano* e *Small*), foram as que apresentaram os melhores resultados dos critérios de avaliação. Ademais, foi possível observar uma tendência do corte sagital ser o que melhor oferece características aos modelos para que sejam capazes de prever a idade dentária. Devido a isso, uma opção que pode ser interessante em futuros testes é a de concentrar mais esforços nas versões de modelos mais simples e explorar mais as imagens pertencentes ao corte sagital.

Logo, ao longo do trabalho foi possível o alcance dos objetivos propostos. Sendo iniciada a criação do conjunto de dados de imagens coletadas na Faculdade de Odontologia da UFJF, totalizando 1110 imagens até a data desta monografia. Além disso, os modelos de aprendizado profundo foram desenvolvidos, assim como os códigos disponibilizados na URL <https://www.kaggle.com/code/alexvpp/odonto-train-yolo>.

## 6.1 Desafios

Neste estudo, a tentativa de realizar a classificação das imagens radiográficas em faixas de idade se deparou com alguns desafios. Consequentemente, afetando a acurácia e previsão correta dos modelos. Assim, os seguintes desafios podem ser citados:

- Desbalanceamento do Conjunto de Dados

A diferença no número de amostras entre as diferentes faixas etárias apresentou um desafio considerável. Como pode ser observado na Figura 4.3, os grupos iniciais, em específico os grupos 1 e 2 caracterizados por idades mais jovens, possuem um número maior de imagens. Em contraste, as faixas intermediárias e superiores possuem uma quantidade mais reduzida de dados.

Esse desbalanceamento impacta negativamente o treinamento e desempenho dos modelos, tornando-os mais propensos a favorecer as classes com maior quantidade de dados. Além disso, esse fato pôde ser observado em muitas matrizes de confusão, as quais os valores de *Recall* para o grupo2 tendeu a ser o maior. Enquanto isso, grupos intermediários (3 ao 5) frequentemente tiveram valores de *Recall* menores, e muitas das vezes, o modelo previa que essas imagens eram pertencentes a grupos com idades mais baixas.

- Número reduzido de imagens no Conjunto de Dados

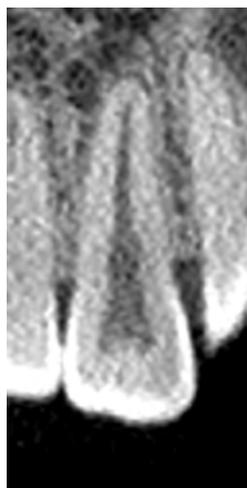
A limitação na quantidade de imagens disponíveis representa outro desafio. Nesse caso, o fato de haver menos dados, em comparação com trabalhos relacionados que utilizam milhares de dados, pode resultar em modelos menos robustos, incapazes de generalizar eficientemente para situações não observadas durante o treinamento. Os conjuntos de dados correspondentes a cada um dos cortes (coronal, sagital e concatenado) possuíam, ao todo, 555 imagens cada, enquanto trabalhos como os de Han et al. (2022) e Wang et al. (2023) utilizaram cerca de 10.000 imagens no total.

- Similaridade da estrutura dentária em faixas de idade distintas

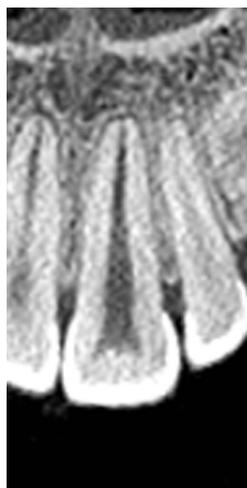
Há uma tendência clara, como já apresentado no presente trabalho, do encurtamento da região da câmara pulpar com o avanço da idade. Porém, em alguns casos,

é possível observar na base de dados a presença de dentes com câmaras pulpares visualmente semelhantes, apesar de pertencerem a faixas etárias (grupos) diferentes. Afinal, cada indivíduo irá apresentar um desenvolvimento diferente ao longo da vida. Esse fato adiciona uma maior complexidade à tarefa de classificação. Já que essa variabilidade desafia a capacidade do modelo em diferenciar características distintivas que permitam uma classificação precisa, principalmente se a base não apresentar uma maior variedade e quantidade de dados como mencionado no desafio anterior. Apenas para ilustração, a Figura 6.1 ilustra imagens pertencentes ao conjunto de treinamento da base de dados para o corte Coronal, onde indivíduos de faixas etárias distintas apresentam estruturas dentárias semelhantes, principalmente levando em consideração a região da câmara pulpar. Por fim, pode ser observado na Figura 6.2 como em alguns casos a câmara pulpar em idades mais jovens podem estar reduzidas. Estas imagens também se encontram no conjunto de treinamento.

Uma outra limitação que pode ser apresentada nesta monografia é o fato de não terem sido encontradas referências detalhando a arquitetura YOLOv8 para problemas de classificação, além de descrever cada uma das diferenças entre as versões de modelo (*Nano*, *Small*, *Medium*, *Large* e *Extra Large*).



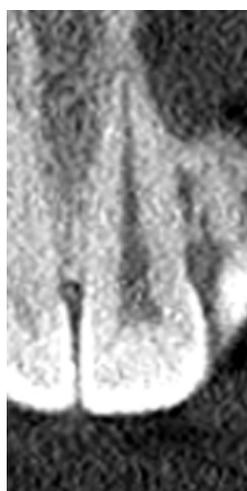
(a) Grupo 1 - 21 anos



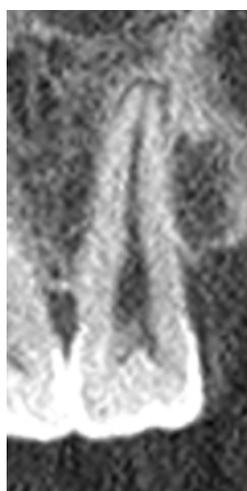
(b) Grupo 2 - 23 anos



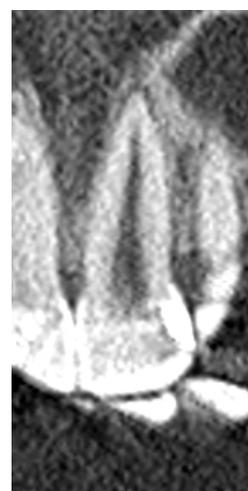
(c) Grupo 3 - 29 anos



(d) Grupo 4 - 36 anos



(e) Grupo 5 - 39 anos



(f) Grupo 8 - 55 anos

Figura 6.1: Comparação entre estruturas dentárias semelhantes de indivíduos de faixas etárias diferentes, com foco na câmara pulpar

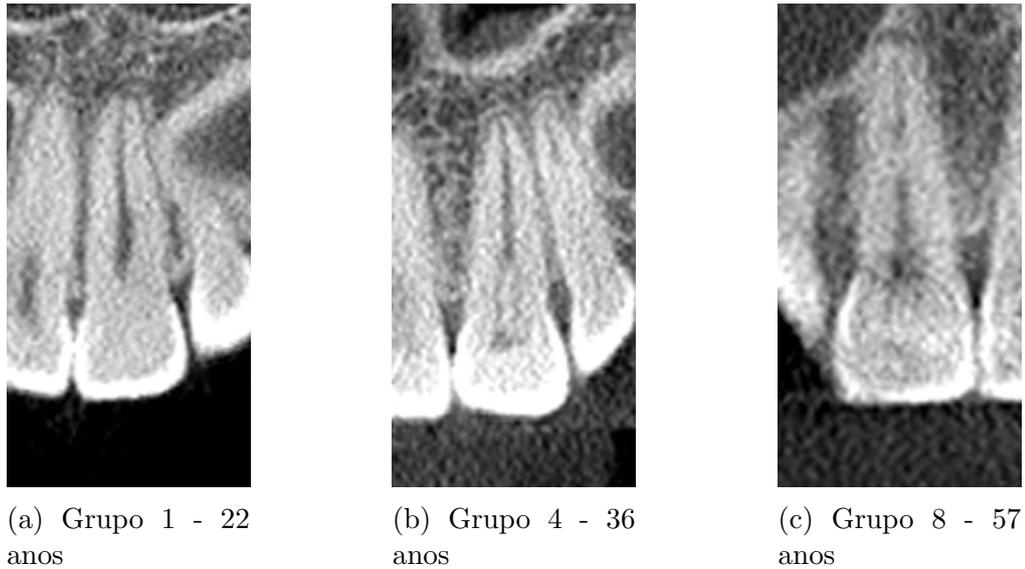


Figura 6.2: Câmara pulpar reduzida mesmo em idades mais jovens

## 6.2 Trabalhos Futuros

A partir desta monografia é possível realizar novas pesquisas e buscar melhorias na determinação da idade de indivíduos a partir das imagens tomográficas de suas câmaras pulpares, utilizando de aprendizado profundo para esse fim.

A expansão do conjunto de dados é um dos trabalhos que poderão auxiliar ainda mais nos aprendizados dos modelos. É esperado que, com um maior número de imagens, o modelo possa estar apto a reconhecer os mais diversos padrões e características das estruturas dentárias, sendo capaz de atingir acurácias maiores.

Além disso, auxiliando essa expansão da base de dados, pôde-se notar a relevância de garantir, futuramente, que grupos com menores quantidades de imagens tenham um balanceamento maior.

Outras técnicas também podem ser exploradas de modo a garantir um maior foco na região de interesse para a estimativa da idade. Por exemplo, a segmentação de partes importantes como a polpa e a coroa dentária, o que permite uma extração de características de forma mais detalhada dessas regiões, reduzindo ruídos e interferências de outras partes da imagem.

Por fim, é possível buscar explorar a região da polpa dentária de modo 3D, graças a tecnologia da TCFC, assim como foi explorado no trabalho de Zheng et al. (2020).

## Bibliografia

- ABOSHI, H.; TAKAHASHI, T.; KOMURO, T. Age estimation using microfocus x-ray computed tomography of lower premolars. *Forensic science international*, v. 200 1-3, p. 35–40, 2010.
- ARAI, Y.; TAMMISALO, E.; IWAI, K.; HASHIMOTO, K.; SHINODA, K. Development of a compact computed tomographic apparatus for dental use. *Dento maxillo facial radiology*, v. 28 4, p. 245–8, 1999.
- BRITANNICA, T. E. o. E. *Tooth Anatomy*. 2023. Disponível em: <https://www.britannica.com/science/tooth-anatomy>. Acessado em: 09/05/2023.
- CHAN, T. F.; VESE, L. A. Active contours without edges. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, v. 10 2, p. 266–77, 2001.
- DAI, Z.; LIU, H.; LE, Q. V.; TAN, M. Coatnet: Marrying convolution and attention for all data sizes. *ArXiv*, abs/2106.04803, 2021. Disponível em: <https://api.semanticscholar.org/CorpusID:235376986>.
- DUMOULIN, V.; VISIN, F. A guide to convolution arithmetic for deep learning. *ArXiv*, abs/1603.07285, 2016.
- FANG, W.; LOVE, P. E. D.; LUO, H.; DING, L. Computer vision for behaviour-based safety in construction: A review and future directions. *Adv. Eng. Informatics*, v. 43, p. 100980, 2020.
- FELDKAMP, L. A.; DAVIS, L. C.; KRESS, J. W. Practical cone-beam algorithm. *Journal of The Optical Society of America A-optics Image Science and Vision*, v. 1, p. 612–619, 1984.
- GOLDMAN, L. W. Principles of ct and ct technology\*. *Journal of Nuclear Medicine Technology*, v. 35, p. 115 – 128, 2007.
- GONZALEZ, R.; WOODS, R. *Digital Image Processing*. Pearson, 2018. ISBN 9780133356724. Disponível em: <https://books.google.com.br/books?id=0F05vgAACAAJ>.
- GOODFELLOW, I. J.; BENGIO, Y.; COURVILLE, A. Deep learning. In: . Cambridge, MA, USA: MIT Press, 2016. <http://www.deeplearningbook.org>.
- GUO, Y.-C.; HAN, M.; CHI, Y.; LONG, H.; ZHANG, D.; YANG, J.; YANG, Y.; CHEN, T.; DU, S. Accurate age classification using manual method and deep convolutional neural network based on orthopantomogram images. *International Journal of Legal Medicine*, v. 135, p. 1589 – 1597, 2021.
- GUPTA, S.; AGNIHOTRI, A.; CHANDRA, A.; GUPTA, O. P. Contemporary practice in forensic odontology. *Journal of Oral and Maxillofacial Pathology : JOMFP*, v. 18, p. 244 – 250, 2014.

HAN, M.; DU, S.; GE, Y.; ZHANG, D.; CHI, Y.; LONG, H.; YANG, J.; YANG, Y.; XIN, J.; CHEN, T.; ZHENG, N.; GUO, Y.-C. With or without human interference for precise age estimation based on machine learning? *International Journal of Legal Medicine*, v. 136, p. 821 – 831, 2022.

HARRIS, C. R.; MILLMAN, K. J.; WALT, S. J. van der; GOMMERS, R.; VIRTANEN, P.; COURNAPEAU, D.; WIESER, E.; TAYLOR, J.; BERG, S.; SMITH, N. J.; KERN, R.; PICUS, M.; HOYER, S.; KERKWIJK, M. H. van; BRETT, M.; HALDANE, A.; RÍO, J. F. del; WIEBE, M.; PETERSON, P.; GÉRARD-MARCHANT, P.; SHEPPARD, K.; REDDY, T.; WECKESSER, W.; ABBASI, H.; GOHLKE, C.; OLIPHANT, T. E. Array programming with NumPy. *Nature*, Springer Science and Business Media LLC, v. 585, n. 7825, p. 357–362, set. 2020. Disponível em: <https://doi.org/10.1038/s41586-020-2649-2>.

HAYKIN, S. Redes neurais: princípios e prática. In: . [S.l.]: Bookman Editora, 2001.

HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 770–778, 2015. Disponível em: <https://api.semanticscholar.org/CorpusID:206594692>.

HUNTER, J. D. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, IEEE COMPUTER SOC, v. 9, n. 3, p. 90–95, 2007.

JAIN, A. K.; MAO, J.; MOHIUDDIN, K. M. Artificial neural networks: A tutorial. *Computer*, v. 29, p. 31–44, 1996. Disponível em: <https://api.semanticscholar.org/CorpusID:5132449>.

JHONG, S.-Y.; TSENG, P.-Y.; SIRIPHOCKPIROM, N.; HSIA, C.-H.; HUANG, M.-S.; HUA, K.-L.; CHEN, Y.-Y. An automated biometric identification system using cnn-based palm vein recognition. *2020 International Conference on Advanced Robotics and Intelligent Systems (ARIS)*, p. 1–6, 2020.

JOCHER, G.; CHAURASIA, A.; QIU, J. *Ultralytics YOLOv8*. 2023. Disponível em: <https://github.com/ultralytics/ultralytics>.

KAGGLE. 2023. Disponível em: <https://www.kaggle.com/>.

KAZMI, S.; MÂNICA, S.; REVIE, G. F.; SHEPHERD, S.; HECTOR, M. P. Age estimation using canine pulp volumes in adults: a cbct image analysis. *International Journal of Legal Medicine*, v. 133, p. 1967 – 1976, 2019.

KIM, S.; LEE, Y.-H.; NOH, Y.-K.; PARK, F. C.; AUH, Q.-S. Age-group determination of living individuals using first molar images based on artificial intelligence. *Scientific Reports*, v. 11, 2021.

KUMAR, M.; SHANAVAS, M.; SIDAPPA, A.; KIRAN, M. Cone beam computed tomography - know its secrets. *Journal of International Oral Health : JIOH*, v. 7, p. 64 – 68, 2015.

LI, C.; XU, C.; GUI, C.; FOX, M. D. Level set evolution without re-initialization: a new variational formulation. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, v. 1, p. 430–436 vol. 1, 2005.

LI, G.; ZHANG, M.; LI, J.; LV, F.; TONG, G. Efficient densely connected convolutional neural networks. *Pattern Recognit.*, v. 109, p. 107610, 2021.

- LI, H.-C.; DENG, Z.-Y.; CHIANG, H.-H. Lightweight and resource-constrained learning network for face recognition with performance optimization. *Sensors (Basel, Switzerland)*, v. 20, 2020.
- MOZZO, P.; PROCACCI, C.; TACCONI, A.; MARTINI, P. T.; ANDREIS, I. A. B. A new volumetric ct machine for dental imaging based on the cone-beam technique: preliminary results. *European Radiology*, v. 8, p. 1558–1564, 1998.
- MUHAMMAD, M. B.; YEASIN, M. Eigen-cam: Class activation map using principal components. *2020 International Joint Conference on Neural Networks (IJCNN)*, p. 1–7, 2020. Disponível em: <https://api.semanticscholar.org/CorpusID:220936532>.
- MUHAMMAD, M. B.; YEASIN, M. Eigen-cam: Visual explanations for deep convolutional neural networks. *SN Computer Science*, v. 2, 2021. Disponível em: <https://api.semanticscholar.org/CorpusID:231681793>.
- PALAZ, D.; MAGIMAI-DOSS, M.; COLLOBERT, R. End-to-end acoustic modeling using convolutional neural networks for hmm-based automatic speech recognition. *Speech Commun.*, v. 108, p. 15–32, 2019.
- PALOMO, L.; PALOMO, J. M. Cone beam ct for diagnosis and treatment planning in trauma cases. *Dental clinics of North America*, v. 53 4, p. 717–27, vi–vii, 2009.
- PHAROAH, M. J.; WHITE, S. C. Oral radiology: principles and interpretation. In: . [S.l.]: Mosby/Elsevier, 2009.
- PINCHI, V.; PRADELLA, F.; BUTI, J.; BALDINOTTI, C.; FOCARDI, M.; NORELLI, G. A new age estimation procedure based on the 3d cbct study of the pulp cavity and hard tissues of the teeth for forensic purposes: A pilot study. *Journal of forensic and legal medicine*, v. 36, p. 150–7, 2015.
- RAZAVI, T.; PALMER, R. M.; DAVIES, J.; WILSON, R. F.; PALMER, P. J. Accuracy of measuring the cortical bone thickness adjacent to dental implants using cone beam computed tomography. *Clinical oral implants research*, v. 21 7, p. 718–25, 2010.
- REZENDE, S. *Sistemas inteligentes: fundamentos e aplicações*. Manole, 2003. ISBN 9788520416839. Disponível em: [https://books.google.com.br/books?id=UsJe\\_PlbnWcC](https://books.google.com.br/books?id=UsJe_PlbnWcC).
- rigvedrs. *YOLO-V8-CAM Repository*. 2023. <https://github.com/rigvedrs/YOLO-V8-CAM>.
- RUSSAKOVSKY, O.; DENG, J.; SU, H.; KRAUSE, J.; SATHEESH, S.; MA, S.; HUANG, Z.; KARPATY, A.; KHOSLA, A.; BERNSTEIN, M.; BERG, A. C.; FEI-FEI, L. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision (IJCV)*, v. 115, n. 3, p. 211–252, 2015.
- SCARFE, W. C.; FARMAN, A. G.; SUKOVIC, P. Clinical applications of cone-beam computed tomography in dental practice. *Journal*, v. 72 1, p. 75–80, 2006.
- SCHULZE, D.; HEILAND, M.; THURMANN, H.; ADAM, G. Radiation exposure during midfacial imaging using 4- and 16-slice computed tomography, cone beam computed tomography systems and conventional radiography. *Dento maxillo facial radiology*, v. 33 2, p. 83–6, 2004.

- SHAH, J. S.; RANGHANI, A. F.; LIMDIWALA, P. G. Age estimation by assessment of dentin translucency in permanent teeth. *Indian Journal of Dental Research*, v. 31, p. 31 – 36, 2020.
- STAR, H.; THEVISSSEN, P. W.; JACOBS, R.; FIEUWS, S.; SOLHEIM, T.; WILLEMS, G. Human dental age estimation by calculation of pulp–tooth volume ratios yielded on clinically acquired cone beam computed tomography images of monoradicular teeth \*. *Journal of Forensic Sciences*, v. 56, 2011. Disponível em: [⟨https://api.semanticscholar.org/CorpusID:7267754⟩](https://api.semanticscholar.org/CorpusID:7267754).
- TEAM, T. pandas development. *pandas-dev/pandas: Pandas*. Zenodo, 2020. Disponível em: [⟨https://doi.org/10.5281/zenodo.3509134⟩](https://doi.org/10.5281/zenodo.3509134).
- VENKATESH, E.; ELLURU, S. V. Cone beam computed tomography: basics and applications in dentistry. *Journal of Istanbul University Faculty of Dentistry*, v. 51, p. S102 – S121, 2017.
- VILA-BLANCO, N.; CARREIRA, M. J.; VARAS-QUINTANA, P.; BALSACASTRO, C.; TOMÁS, I. Deep neural networks for chronological age estimation from opg images. *IEEE Transactions on Medical Imaging*, v. 39, p. 2374–2384, 2020.
- VOSSOUGH, M.; MOVAHHEDIAN, N.; GHAFORI, A. The impact of age mimicry bias on the accuracy of methods for age estimation based on kvaal’s pulp/tooth ratios: a bootstrap study. *International Journal of Legal Medicine*, v. 136, p. 269–278, 2021. Disponível em: [⟨https://api.semanticscholar.org/CorpusID:236149511⟩](https://api.semanticscholar.org/CorpusID:236149511).
- WANG, J.; DOU, J.; HAN, J.; LI, G.; TAO, J. A population-based study to assess two convolutional neural networks for dental age estimation. *BMC Oral Health*, v. 23, 2023.
- WASKOM, M. L. seaborn: statistical data visualization. *Journal of Open Source Software*, The Open Journal, v. 6, n. 60, p. 3021, 2021. Disponível em: [⟨https://doi.org/10.21105/joss.03021⟩](https://doi.org/10.21105/joss.03021).
- WILLEMS, G. A review of the most commonly used dental age estimation techniques. *The Journal of forensic odonto-stomatology*, v. 19 1, p. 9–17, 2001.
- WILLEMS, G.; MOULIN-ROMSEE, C.; SOLHEIM, T. Non-destructive dental-age calculation methods in adults: intra- and inter-observer effects. *Forensic science international*, v. 126 3, p. 221–6, 2002.
- YAMASHITA, R.; NISHIO, M.; DO, R. K. G.; TOGASHI, K. Convolutional neural networks: an overview and application in radiology. *Insights into Imaging*, v. 9, p. 611 – 629, 2018.
- YANG, F.; JACOBS, R.; WILLEMS, G. Dental age estimation through volume matching of teeth imaged by cone-beam ct. *Forensic science international*, v. 159 Suppl 1, p. S78–83, 2006.
- ZHENG, Q.; GE, Z. pu; DU, H.; LI, G. Age estimation based on 3d pulp chamber segmentation of first molars from cone-beam–computed tomography by integrated deep learning and level set. *International Journal of Legal Medicine*, v. 135, p. 365–373, 2020.
- ZHOU, D.-X. Theory of deep convolutional neural networks: Downsampling. *Neural networks : the official journal of the International Neural Network Society*, v. 124, p. 319–327, 2020.

## A Anexo

### A.1 Parecer do Comitê de Ética em Pesquisa

**PARECER CONSUBSTANCIADO DO CEP****DADOS DO PROJETO DE PESQUISA**

**Título da Pesquisa:** O uso da inteligência artificial na estimativa da idade cronológica por meio da avaliação de imagens tomográficas da câmara pulpar

**Pesquisador:** Karina Lopes Devito

**Área Temática:**

**Versão:** 1

**CAAE:** 59247922.1.0000.5147

**Instituição Proponente:** FACULDADE DE ODONTOLOGIA

**Patrocinador Principal:** Financiamento Próprio

**DADOS DO PARECER**

**Número do Parecer:** 5.485.995

**Apresentação do Projeto:**

Trata-se de um estudo retrospectivo observacional transversal, em que serão utilizados exames de tomografia computadorizada de feixe cônico (TCFC) pertencentes ao arquivo (banco de dados) do Serviço de Radiologia Odontológica da Faculdade de Odontologia da Universidade Federal de Juiz de Fora (FO/UFJF). Serão analisadas 1200 amostras de imagens de dentes que serão previamente divididos em grupos de acordo com a faixa etária dos participantes.

**Objetivo da Pesquisa:**

- Avaliar as dimensões da câmara pulpar em imagens de tomográficas de participantes de uma subpopulação brasileira como uma possível ferramenta para estimativa da idade cronológica.
- Realizar medidas lineares, de área e volumétricas para aferir as dimensões da câmara pulpar em imagens de TCFC;
- Correlacionar essas medidas com a idade cronológica para definir os melhores indicadores de estimativa de idade para ambos os sexos;
- Desenvolver redes neurais convolucionais para verificar a otimização da estimativa de idade usando as imagens dos dentes/câmaras pulpares.

**Avaliação dos Riscos e Benefícios:**

RISCOS:

Como se trata de um estudo com imagens tomográficas de arquivo (banco de dados), esse

**Endereço:** JOSE LOURENCO KELMER S/N

**Bairro:** SAO PEDRO

**CEP:** 36.036-900

**UF:** MG

**Município:** JUIZ DE FORA

**Telefone:** (32)2102-3788

**E-mail:** cep.propp@ufjf.edu.br

Continuação do Parecer: 5.485.995

projeto de pesquisa envolve riscos mínimos. Os riscos estariam associados à Identificação dos participantes. No entanto, enfatiza-se que não haverá qualquer identificação pessoal dos envolvidos, uma vez que serão atribuídos códigos numéricos aos exames selecionados para o estudo, sendo salvas para aplicação da metodologia apenas as imagens isoladas dos dentes a serem avaliados.

#### BENEFÍCIOS:

Uma vez que a aposição de dentina secundária tem se mostrado um preditor significativo da idade cronológica e que a TCFC tem sido cada vez mais utilizada na prática odontológica, os resultados do presente projeto fornecerão informações importantes para auxiliar em processos de identificação de indivíduos brasileiros da região da Zona da Mata Mineira, indicando a medida dentária mais acurada para estimativa de idade.

#### Comentários e Considerações sobre a Pesquisa:

O projeto está bem estruturado, delineado e fundamentado, sustenta os objetivos do estudo em sua metodologia de forma clara e objetiva. O projeto apresenta o tipo de estudo, o "n", os critérios de inclusão e exclusão. As referencias bibliográficas

são atuais, sustentam os objetivos do estudo e seguem uma normatização. O cronograma mostra as diversas etapas da pesquisa, além de mostrar que a coleta de dados ocorrerá após aprovação do projeto pelo CEP. O orçamento lista a relação dos custos da pesquisa que serão financiados com recursos próprios conforme consta no campo apoio financeiro. A pesquisa proposta está de acordo com as atribuições definidas na Resolução CNS 466 de 2012, itens IV.6, II.11 e XI.2; com a Norma Operacional CNS 001 de 2013. Itens: 3.4.1-6, 8, 9, 10 e 11; 3.3 - f; com o Manual Operacional para CEPS Item: VI - c.

#### Considerações sobre os Termos de apresentação obrigatória:

O protocolo de pesquisa está em configuração adequada, apresenta FOLHA DE ROSTO devidamente preenchida, com o título em português, identifica o patrocinador da pesquisa, estando de acordo com as atribuições definidas na Norma Operacional CNS 001 de 2013 item 3.3 letra a; e 3.4.1 item 16. Pede dispensa do TERMO DE CONSENTIMENTO LIVRE ESCLARECIDO com a seguinte justificativa: "pelo fato de que se trata de um estudo observacional retrospectivo com uso de informações de um banco de dados (exames de TCFC) institucional, em que, pelo tamanho da amostra (n = 1200 dentes) e pelos exames terem sido realizados em datas anteriores (entre 2010 a 2019), não é possível a localização de cada participante para a obtenção do TCLE individual (sem endereço residencial e/ou contato telefônico válido). Além disso, apenas as imagens

tomográficas de dentes anteriores hígidos serão utilizadas para as medidas, o que não permite a

**Endereço:** JOSE LOURENCO KELMER S/N

**Bairro:** SAO PEDRO

**CEP:** 36.036-900

**UF:** MG

**Município:** JUIZ DE FORA

**Telefone:** (32)2102-3788

**E-mail:** cep.propp@ufjf.edu.br

Continuação do Parecer: 5.485.995

identificação pessoal dos participantes. Vale destacar que os dados serão analisados de forma anônima e os resultados serão apresentados de forma agregada, não permitindo a identificação dos participantes da pesquisa". A pesquisadora garante a confidencialidade e sigilo, explicita os riscos de forma adequada, e está de acordo com as atribuições definidas na Resolução CNS 466 de 2012. O Pesquisador apresenta titulação e experiência compatível com o projeto de pesquisa, estando de acordo com as atribuições definidas no Manual Operacional para CPEs. Apresenta DECLARAÇÃO de infraestrutura e de concordância com a realização da pesquisa da Instituição proponente, de acordo com as atribuições definidas na Norma Operacional CNS 001 de 2013 item 3.3 letra h.

**Recomendações:**

-

**Conclusões ou Pendências e Lista de Inadequações:**

Diante do exposto, o projeto está aprovado, pois está de acordo com os princípios éticos norteadores da ética em pesquisa estabelecido na Res. 466/12 CNS e com a Norma Operacional Nº 001/2013 CNS. Data prevista para o término da pesquisa: dezembro de 2023.

**Considerações Finais a critério do CEP:**

Diante do exposto, o Comitê de Ética em Pesquisa CEP/UFJF, de acordo com as atribuições definidas na Res. CNS 466/12 e com a Norma Operacional Nº001/2013 CNS, manifesta-se pela APROVAÇÃO do protocolo de pesquisa proposto. Vale lembrar ao pesquisador responsável pelo projeto, o compromisso de envio ao CEP de relatórios parciais e/ou total de sua pesquisa informando o andamento da mesma, comunicando também eventos adversos e eventuais modificações no protocolo.

**Este parecer foi elaborado baseado nos documentos abaixo relacionados:**

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1958524.pdf	01/06/2022 18:32:20		Aceito
Outros	Curriculo_Lattes.pdf	01/06/2022 18:30:42	Karina Lopes Devito	Aceito
Outros	Termo_Confidencialidade_Sigilo.pdf	01/06/2022 18:29:40	Karina Lopes Devito	Aceito
TCLE / Termos de Assentimento /	Dispensa_TCLE.pdf	01/06/2022 18:29:08	Karina Lopes Devito	Aceito

**Endereço:** JOSE LOURENCO KELMER S/N**Bairro:** SAO PEDRO**CEP:** 36.036-900**UF:** MG**Município:** JUIZ DE FORA**Telefone:** (32)2102-3788**E-mail:** cep.propp@ufjf.edu.br

Continuação do Parecer: 5.485.995

Justificativa de Ausência	Dispensa_TCLE.pdf	01/06/2022 18:29:08	Karina Lopes Devito	Aceito
Declaração de Instituição e Infraestrutura	Declaracao_Infraestrutura.pdf	01/06/2022 18:28:45	Karina Lopes Devito	Aceito
Projeto Detalhado / Brochura Investigador	Projeto_CEP.pdf	01/06/2022 18:28:36	Karina Lopes Devito	Aceito
Folha de Rosto	folhaDeRosto.pdf	01/06/2022 18:28:13	Karina Lopes Devito	Aceito

**Situação do Parecer:**

Aprovado

**Necessita Apreciação da CONEP:**

Não

JUIZ DE FORA, 23 de Junho de 2022

---

**Assinado por:**

**Patrícia Aparecida Baumgratz de Paula  
(Coordenador(a))**

**Endereço:** JOSE LOURENCO KELMER S/N

**Bairro:** SAO PEDRO

**CEP:** 36.036-900

**UF:** MG

**Município:** JUIZ DE FORA

**Telefone:** (32)2102-3788

**E-mail:** cep.propp@ufjf.edu.br