



**UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
DEPARTAMENTO DE ODONTOLOGIA**

**AVALIAÇÃO DA PERFORMANCE DE DOIS SOFTWARES COM
INTELIGÊNCIA ARTIFICIAL POR MEIO DAS MEDIDAS
GERADAS PELA ANÁLISE DE MCNAMARA EM
TELERRADIOGRAFIA CEFALOMÉTRICA LATERAL**

Aracaju
Outubro/ 2022

LAURA LUIZA TRINDADE DE SOUZA

**AVALIAÇÃO DA PERFORMANCE DE DOIS SOFTWARES COM
INTELIGÊNCIA ARTIFICIAL POR MEIO DAS MEDIDAS
GERADAS PELA ANÁLISE DE MCNAMARA EM
TELERRADIOGRAFIA CEFALOMÉTRICA LATERAL**

Monografia apresentada ao
Departamento de Odontologia como
requisito para a conclusão do curso de
Odontologia da Universidade Federal de
Sergipe para obtenção do grau de
Cirurgiã-Dentista.

Orientador: Prof. Dr. Wilton Mitsunari
Takeshita

Coorientadora: Mestranda Thaísa
Pinheiro Silva

Aracaju

Outubro/ 2022



**UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
DEPARTAMENTO DE ODONTOLOGIA**

A comissão julgadora dos Trabalhos de Conclusão de Curso, em sessão pública realizada em 17 de outubro de 2022, considerou o candidato (a) Laura Luiza Trindade de Souza aprovado (a).

Prof. Dr. Wilton Mitsunari Takeshita – Orientador
Universidade Federal de Sergipe

Prof.^(a) Dr(a). – 1º Examinador

Prof.^(a) Dr(a). – 2º Examinador

AGRADECIMENTOS

Hoje, entendo que chegar até aqui só foi possível por estar cercada de pessoas que me deram suporte ao longo de toda essa jornada. Quero agradecer em especial aos meus pais, Aldenice e José Luiz, por serem meu maior alicerce e os maiores incentivadores dos meus sonhos. Também agradeço ao meu irmão, Candinho, por ser sempre presente e comemorar cada conquista.

Agradeço aos meus mestres da UFS por toda aprendizagem que me proporcionaram. Em especial, ao Prof. Dr. Wilton Takeshita pelo apoio dedicado ao meu projeto de pesquisa, assim como por todos os ensinamentos durante a graduação, e pela confiança depositada em mim. E à Prof. Thaísa Pinheiro por ser uma inspiração desde o começo da graduação e pela imensa contribuição para a concretização desse trabalho científico.

Agradeço também aos meus amigos por estarem sempre ao meu lado; vocês me inspiram. E a todos os pacientes e funcionários do Departamento de Odontologia que compartilharam suas histórias e ensinamentos comigo.

Obrigada a todos que contribuíram direta ou indiretamente para essa conquista!

RESUMO

Introdução: Na Odontologia, a Ortodontia tornou-se uma das mais importantes áreas de aplicabilidade dos softwares com inteligência artificial (IA). A análise cefalométrica é uma ferramenta essencial no diagnóstico e planejamento ortodôntico e/ou cirúrgico. Atualmente, com intuito de otimizar o tempo do profissional e reduzir erros intrínsecos à condição humana, a identificação automática dos pontos cefalométricos por meio da IA vem sendo cada vez mais difundida.

Objetivo: O objetivo desta pesquisa foi comparar a performance de dois softwares com IA em telerradiografia cefalométrica lateral, por meio da avaliação da reprodutibilidade e confiabilidade das medidas lineares e angulares da análise de McNamara.

Materiais e Métodos: 30 telerradiografias cefalométricas foram marcadas por meio do método digital pelo examinador calibrado utilizando a plataforma do software Radioceph (RadioMemory). Posteriormente, as mesmas telerradiografias cefalométricas laterais foram marcadas por meio da IA dos softwares CEFBOT (RadioMemory) e WebCeph™ (AssembleCircle), para avaliação de duas vertentes de marcação: reprodutibilidade e confiabilidade em relação ao examinador e os softwares em questão. Foi realizada a análise estatística por meio do Coeficiente de Correlação Intraclasse (ICC) para calibração do examinador e avaliação da confiabilidade das marcações do examinador calibrado, do CEFBOT (RadioMemory), e WebCeph™ (AssembleCircle) software (Korea), bem como o teste ANOVA e pós teste de Tukey para avaliação da reprodutibilidade dos softwares com IA, por meio dos pontos cefalométricos que compõem a análise de McNamara.

Resultados: o ICC médio gerado pelas marcações cefalométricas do examinador calibrado, do CEFBOT e do WebCeph foram 0.960, 0.940 e 0.954, respectivamente, indicando concordância quase perfeita. Quando comparado o CEFBOT (RadioMemory) com o examinador calibrado, observou-se diferença estatística ($p < 0.01$) apenas na medida A-N perpendicular. Quanto ao WebCeph™ (AssembleCircle), ao comparar com o examinador calibrado houve diferença significativa entre os fatores dois ao seis e o dez. E comparado ao CEFBOT (RadioMemory), houve divergência nos mesmos fatores somado ao fator onze. Além disso, o WebCeph™ (AssembleCircle) não conseguiu realizar as medidas Nfa-Nfp e Bfa-Bfp.

Conclusão: O software CEFBOT apresentou reprodutibilidade e confiabilidade na identificação dos pontos cefalométricos determinados pela análise de McNamara, porém necessitando supervisão humana. O WebCeph apresentou concordância quase perfeita nas marcações, porém seis medidas apresentaram-se diferentes do examinador calibrado e duas não foram realizadas pela aplicação.

Palavras-chave: Inteligência artificial. Cefalometria. Reprodutibilidade. Confiabilidade.

ABSTRACT

Introduction: In dentistry, orthodontics has become one of the most important areas of applicability of artificial intelligence (AI) software. Cephalometric analysis is an essential tool in orthodontic and/or surgical diagnosis and planning. Currently, to optimize professional time and reduce errors intrinsic to the human condition, the automatic identification of cephalometric landmarks using AI is becoming increasingly widespread.

Aim: The aim of this study was to compare the performance of two software with AI in lateral cephalometric telerradiograph by assessing the reproducibility and reliability of linear and angular measurements of McNamara's analysis.

Materials and Methods: 30 cephalometric telerradiographs were marked by the digital method by the calibrated examiner using the Radiocef (RadioMemory) software platform. Subsequently, the same lateral cephalometric radiographs were marked using the CEFBOT (RadioMemory) and WebCephTM (AssembleCircle) software AI to evaluate two aspects of marking: reproducibility and reliability in relation to the examiner and the software in question. Statistical analysis was performed using the Intraclass Correlation Coefficient (ICC) for examiner calibration and assessment of the reliability of the calibrated examiner, CEFBOT (RadioMemory), and WebCephTM (AssembleCircle) software (Korea), as well as the ANOVA test and Tukey's post-test to evaluate the reproducibility of the software with AI, using the cephalometric landmarks that comprise McNamara's analysis.

Results: The mean ICC generated by the cephalometric markings of the calibrated examiner, CEFBOT and WebCeph were 0.960, 0.940, and 0.954, respectively, indicating almost perfect agreement. When comparing the CEFBOT (RadioMemory) with the calibrated examiner, a statistical difference ($p < 0.01$) was observed only in the perpendicular A-N measurement. As for WebCephTM (AssembleCircle), when compared to the calibrated examiner there was a significant difference between factors two through six and ten. And compared to CEFBOT (RadioMemory), there was divergence in the same factors plus factor eleven. In addition, WebCephTM (AssembleCircle) failed to perform the measurements Nfa-Nfp and Bfa-Bfp.

Conclusion: CEFBOT software showed reproducibility and reliability in identifying cephalometric landmarks determined by McNamara's analysis but required human supervision. WebCeph showed almost perfect agreement in the markings, but six measurements were different from the calibrated examiner and two were not performed by the application.

Keywords: Artificial intelligence. Cephalometry. Reproducibility of Results. Data Accuracy.

LISTA DE FIGURAS

| | |
|---|----|
| FIGURA 1– Interface do software do CEFBOT. | 24 |
| FIGURA 2 – Seleção da análise de McNamara para a marcação pelo examinador. | 25 |
| FIGURA 3 – Seleção da análise de McNamara e o botão vermelho "CEFBOT" para a ativação da marcação da I.A. | 26 |
| FIGURA 4 – Interface do software WebCeph™ (AssembleCircle)..... | 27 |
| FIGURA 5 – Seleção do botão "Digitalização com I.A" e no botão "Salvar" para ativação da marcação da I.A. | 28 |
| FIGURA 6 – Seleção da análise de McNamara. | 29 |

LISTA DE ABREVIATURAS

ANN's: Artificial Neural Networks

C: CEFBOT

CAD: Computer Aided Detection

CNN's: Convolutional Neural Networks

DP: Desvio Padrão

E: Examinador

IA: Inteligência Artificial

ICC: Índice de Correlação Inter-classe

ML: Machine Learning

NA: Násio

NI: Não Identificado

NS: Diferença não Significativa

Or: Orbital

PNI: Ponto Não Identificado

PR: Porion

SSD: Single Shot Multibox Detector

ST: Sela Túrcica

TB: Twin Block

TCFC: Tomografia Computadorizada de Feixe Cônico

UFS: Universidade Federal de Sergipe

W: WebCeph

YOLOv3: YOLO version 3

SUMÁRIO

| | |
|--|-----------|
| 1 INTRODUÇÃO..... | 10 |
| 2 JUSTIFICATIVA | 12 |
| 3 OBJETIVOS..... | 13 |
| 3.1 OBJETIVO GERAL..... | 13 |
| 3.2 OBJETIVOS ESPECÍFICOS..... | 13 |
| 4 REVISÃO DE LITERATURA | 14 |
| 4.1 ANÁLISES CEFALOMÉTRICAS..... | 14 |
| 4.1.1 Aplicação da Análise de McNamara | 16 |
| 4.2 INTELIGÊNCIA ARTIFICIAL..... | 18 |
| 4.2.1 Inteligência Artificial na Odontologia..... | 19 |
| 5 METODOLOGIA..... | 23 |
| 5.1 PRINCÍPIOS ÉTICOS E CÁLCULO AMOSTRAL..... | 23 |
| 5.2 CARACTERÍSTICAS DA AMOSTRA..... | 23 |
| 5.3 ESTUDO EM DIFERENTES SOFTWARES DE INTELIGÊNCIA ARTIFICIAL (CEFBOT e WEBCEPH)..... | 23 |
| 6 RESULTADOS | 29 |
| 7 DISCUSSÃO..... | 33 |
| 8 CONCLUSÃO..... | 36 |
| REFERÊNCIAS..... | 37 |
| ANEXO A – APROVAÇÃO COMITÊ DE ÉTICA..... | 44 |

1 INTRODUÇÃO

A telerradiografia cefalométrica lateral é um exame complementar utilizado para o diagnóstico e tratamento de discrepâncias esqueléticas, relações dentárias e de tecidos moles das estruturas faciais (ALBARAKATI, KULA e GHONEIMA, 2012; DEBELMAS *et al.*, 2019). A partir da introdução do cefalostato, por Broadbent, em 1931, foi possível a padronização do posicionamento da cabeça do paciente durante a execução da técnica radiográfica, em consequência disso, o exame tornou-se uma ferramenta essencial no planejamento ortodôntico e cirúrgico (CHIEN *et al.*, 2009; OLMEZ *et al.*, 2011; LIVAS *et al.*, 2019; KHAN *et al.*, 2020; MERIC e NAOUMOVA, 2020).

Antigamente, os traçados cefalométricos somente eram realizados manualmente, sendo cada ponto cefalométrico identificado de forma manual em um papel acetato. Contudo, esse método demanda tempo e há uma tendência à subjetividade do profissional que irá realizar os traçados, o que pode levar a grandes discrepâncias no diagnóstico e planejamento ortodôntico e/ou cirúrgico. Com isso, o método digital foi inserido e as telerradiografias cefalométricas laterais passaram a ser digitalizadas e o traçado realizado em softwares de telerradiografia cefalométrica lateral (CHEN *et al.*, 2004; ALBARAKATI, KULA e GHONEIMA, 2012).

Com o crescimento exponencial da tecnologia e ciência, a realização da análise cefalométrica de forma digital passou a substituir o método convencional (ALBARAKATI, KULA e GHONEIMA, 2012). Com isso, surgiram programas computadorizados de análise cefalométrica, que quando comparados ao método manual, possuem benefícios como possibilidade de compartilhamento e aprimoramento de imagens, identificação de pontos cefalométricos em menor tempo e automatização de etapas como desenho do cefalograma e geração dos valores lineares e angulares (MERIC e NAOUMOVA, 2020; KHANAGAR *et al.*, 2021). No entanto, esses programas ainda necessitam que o profissional localize manualmente os pontos cefalométricos, e, portanto, a subjetividade na identificação dos pontos permanece inerente ao método (HUNG *et al.*, 2019).

Assim, na tentativa de otimizar o tempo do profissional e reduzir a subjetividade inerente a essa tarefa, há um crescente desenvolvimento de softwares com inteligência artificial com a finalidade de identificação dos pontos cefalométricos. E a inteligência artificial (IA) pode ser definida como uma constelação de itens (algoritmos, robótica,

redes neurais) que permitem a um software propriedades de inteligência comparáveis as de um ser humano, dentre elas, o aprendizado de bancos de dados com mínima interferência humana (FORSTING, 2017), sendo cada vez mais difundida e imperativa na resolução de problemas complexos e ainda em usos mais triviais que ultrapassam fronteiras como um simples software de tradução ou assistente virtual (OBERMEYER e EMANUEL, 2016; CHEN, STANLEY e ATT, 2020).

Dentre as possibilidades de uso das análises, a análise de McNamara é bastante referenciada em ser a primeira análise que não é apenas sensível à posição dos dentes, mas também às estruturas da base do crânio (MCNAMARA, 1984; RAVIKUMAR *et al.*, 2019). Devido aos riscos que os erros na avaliação cefalométrica podem implicar, um deles é o tratamento inadequado. Diversos estudos surgiram com o intuito de analisar a reprodutibilidade das marcações automatizadas com IA, e assim, aumentar a qualidade e precisão do tratamento ortodôntico (FAROOQ, 2016; KHANAGAR *et al.*, 2021).

Na Odontologia, a IA vem sendo aplicada em diferentes vertentes, e uma delas é na análise de exames digitais para diagnóstico e planejamento. Para isso, foram desenvolvidos softwares que, por meio da análise de um grande volume de dados, equipados com algoritmos aperfeiçoados por especialistas, conseguem identificar e reproduzir padrões (HUNG *et al.*, 2019; KUNZ *et al.*, 2019). Para avaliar a performance de um software, estuda-se a sua reprodutibilidade, isto é, a concordância entre a mesma medida marcadas por dois métodos distintos e a confiabilidade, ou seja, a concordância entre uma mesma medida marcada pelo mesmo método em dois momentos diferentes (ZAMRIK e ISERI, 2021).

A Ortodontia tornou-se uma das mais importantes áreas de aplicação da IA (Chen, Stanley e Att, 2020). Diversos softwares com IA já estão disponíveis no mercado, todavia, a maioria deles possuem licença paga, o que torna menos acessível para os profissionais utilizarem na prática clínica. Diante do exposto, a pesquisa tem como objetivo comparar a performance de dois softwares com IA — WebCephTM (AssembleCircle), disponível gratuitamente, e CEFBOT (RadioMemory Ltd., Belo Horizonte, Brasil), disponível de forma paga— por meio da avaliação da confiabilidade e reprodutibilidade da marcação cefalométrica automatizada na análise de McNamara (1984).

2 JUSTIFICATIVA

Considerando a importância da análise de McNamara no diagnóstico das discrepâncias dento esqueléticas, para o planejamento de procedimentos cirúrgicos e de instalação de aparelhos funcionais (MCNAMARA, 1984; RAVIKUMAR *et al.*, 2019), e diante do crescimento do uso da Inteligência Artificial (IA) como ferramenta na Odontologia, notou-se a necessidade de comparar a performance de dois softwares com IA, já disponíveis no mercado, e assim atestar a viabilidade ou não a aplicação destas plataformas na clínica odontológica.

3 OBJETIVOS

3.1 OBJETIVO GERAL

- Avaliar a reprodutibilidade e confiabilidade das medidas lineares e angulares da análise cefalométrica de McNamara nos softwares CEFBOT (RadioMemory) e WebCeph™ (AssembleCircle) com inteligência artificial.

3.2 OBJETIVOS ESPECÍFICOS

- Descrever historicamente a evolução da inteligência artificial voltada para marcação automatizada.
- Comparar a performance dos softwares CEFBOT (RadioMemory) e WebCeph™ (AssembleCircle) quanto às medidas lineares e angulares na análise de McNamara.

4 REVISÃO DE LITERATURA

Para um melhor entendimento, o presente capítulo foi didaticamente dividido nos seguintes subcapítulos: Análises Cefalométricas e sua subdivisão Aplicação da Análise de McNamara; Inteligência Artificial. Este último, por sua vez, foi ainda subdividido em Inteligência Artificial na Odontologia.

4.1 ANÁLISES CEFALOMÉTRICAS

A região craniofacial é uma estrutura do corpo humano considerada complexa, e, portanto, de difícil análise. Dentre os métodos de estudo, a telerradiografia cefalométrica lateral é usada para descrever a morfologia e o crescimento do esqueleto craniofacial, como também planejar e avaliar os resultados de tratamento (OLMEZ *et al.*, 2011). Assim, a telerradiografia cefalométrica lateral é considerada a arte de mensurar as estruturas do indivíduo (MOSLEH *et al.*, 2016).

Introduzido em 1931, por Broadbent, o cefalostato permitiu a padronização do posicionamento da cabeça do paciente durante a obtenção do exame radiográfico e, desde então, a telerradiografia cefalométrica lateral tornou-se uma ferramenta clínica e de pesquisa essencial na Odontologia (KHAN *et al.*, 2020; MERIC e NAOUMOVA, 2020). Os traçados cefalométricos convencionais são realizados em papel de acetato, faz-se as identificações dos pontos anatômicos de referências e determina as medidas lineares e angulares (SHETTIGAR *et al.*, 2019).

A introdução do cefalostato trouxe não somente a padronização do posicionamento da cabeça para a aquisição radiográfica, mas também o fim da análise subjetiva que assombrava os ortodontistas, permitindo um posicionamento ósseo-esquelético reproduzível, uma vez que as metas cefalométricas eram obtidas acompanhadas, muitas vezes, do comprometimento da estética facial. Nesse contexto, surgiu a grande necessidade de análise de parâmetros faciais para o diagnóstico e planejamento. Ao longo da história, várias análises foram desenvolvidas na tentativa de tornar o diagnóstico e o tratamento mais precisos e adequados às necessidades individuais (LUNDSTROM *et al.*, 1995).

Downs (1948) descreveu os pontos A e B, que representavam a região dos limites anteriores da maxila e mandíbula, respectivamente (JACOBSON, 2003). Em outro artigo,

o pesquisador explicou que o perfil facial refletia as características individuais e que esse fato deveria ser considerado ao realizar um tratamento ortodôntico. Além disso, classificou em quatro tipos faciais: retrognático, mesognático, prognático e prognático verdadeiro (DOWNS, 1952). Posteriormente, Enlow e colaboradores desenvolveram uma análise cefalométrica baseada nos padrões faciais individuais com o objetivo as variações anatômicas entre os indivíduos (ENLOW *et al.*, 1969).

Em 1953, Riedel teve como objetivo estudar a posição anteroposterior da maxila e da mandíbula e introduziu os ângulos SNA, SNB e o ANB, esse último foi mais tarde popularizado por Steiner (AHMED, SHAIKH e FIDA, 2018). No mesmo ano, Steiner (1953) propôs a primeira análise cefalométrica considerada moderna, uma vez que já levava em consideração o perfil facial do indivíduo e teve como referência as grandezas lineares e angulares e diversas medidas de autores como Downs, Wylie, Riedel, entre outros.

Ricketts (1960) analisou 1000 indivíduos para identificar e conhecer os problemas mais comuns em Ortodontia e utilizou planos de referências dentárias e a posição do incisivo inferior.

Em um dos seus estudos, Jacobson (1975) demonstrou um método de avaliar denominada de “Wits” que consistiu em mensurar as desarmonias esqueléticas no sentido anteroposterior. E isso, foi baseado na execução de projeções perpendiculares ao plano oclusal dos pontos A e B, localizados na maxila e mandíbula respectivamente e foram denominados de AO e BO e não sofriam interferência das variações da base do crânio, quando comparada com a medida ANB (JACOBSON, 2003).

McNamara (1984) desenvolveu uma análise cefalométrica justificando que a maioria das análises, que estavam disponíveis no período de 1940 a 1970, quando ainda as alterações significativas nas relações estruturais craniofaciais eram consideradas impossíveis. Sua análise utilizou medidas lineares que relacionavam a maxila com mandíbula, e estas com a base do crânio e os dentes com suas bases apicais. Desse modo, foi baseada em alguns princípios das análises de Ricketts e Harvold, mas outros aspectos, como a construção da linha perpendicular do násio e o ponto A vertical são considerados originais. E, com isso, a análise tinha como propósito facilitar o diagnóstico nos casos cirúrgicos e nos casos em que os aparelhos funcionais fossem utilizados.

Arnett e Bergman (1993) em sua análise afirmaram que um planejamento ortodôntico-cirúrgico inapropriado pode comprometer muito mais a harmonia facial do que as próprias compensações dentais. Através desta análise, pode se observar que faces

esteticamente agradáveis possuíam o ângulo naso-labial variando de 85° a 105°, por exemplo. Os autores observaram, ainda, que a distância bizigomática equivale a dimensão mais larga da face (30% maior que a bigoníaca).

Dessa forma, a telerradiografia cefalométrica lateral passou a ser a grande aliada na busca de informações para o diagnóstico ortodôntico, cirúrgico e para as análises de crescimento e desenvolvimento craniofacial (COLLINS *et al.*, 2007). Sendo assim, é um exame de suma importância do ponto de vista clínico, pois fornece orientações fundamentais para os profissionais (SANTOS *et al.*, 2017).

4.1.1 Aplicação da Análise de McNamara

A análise de McNamara (1984) possui características consideradas superiores a outros métodos (NOURI *et al.*, 2015), dentre algumas vantagens, pode-se citar: maior precisão nas análises das deformidades faciais (TAMADA *et al.*, 2010), avaliação das medidas lineares de discrepância da base apical e da base dental em relação a apical. Desse modo, essa análise é uma importante ferramenta no diagnóstico e planejamento do tratamento de casos ortopédicos ou cirúrgicos (STORNILO-SOUZA *et al.*, 2021).

A linha McNamara é formada pelos ângulos retos entre os pontos cefalométricos orbital (Or) e o porion (Po) com o ponto násio (N). Desse modo, a posição da maxila e da mandíbula pode ser facilmente avaliada pela distância entre o ponto A e o pogônio e a linha projetada. Ademais, é considerada bastante útil na cirurgia ortognática, pois consegue mensurar a extensão da deficiência maxilo-mandibular quantitativamente, e assim, auxiliando no planejamento cirúrgico (TAMADA *et al.*, 2010).

Gu *et al.* (2011) utilizaram em seu estudo os pontos cefalométricos e a grandeza linear da análise de McNamara para definir as normas da população jovem adulta chinesa e compararam com uma amostra de caucasianos. E assim, concluíram que há diferenças significativas nas características do perfil ósseo e tegumentar em diferentes etnias e gêneros, e esse fato precisa ser levado em consideração para o diagnóstico e planejamento do tratamento de cada paciente.

Storniolo-Souza *et al.* (2021), seguindo a mesma linha de pesquisa do estudo citado anteriormente, objetivaram determinar a média dos valores de normalidade relacionadas à análise de McNamara (1984) em jovens mestiços nipo-brasileiros, com oclusão normal, descendentes de japoneses e brasileiros, e comparar os resultados das variáveis com amostras de jovens brasileiros leucodermas e xantodermas que apresentem

as mesmas características. Com isso, concluíram que os grupos de leucodermas, xantodermas e nipo-brasileiros apresentaram características semelhantes cefalométricas da análise de McNamara, visto que, apenas os xantodermas do sexo masculino tiveram o ângulo nasolabial significativamente mais agudo, quando comparados com os leucodermas.

Borba *et al.* (2016), realizaram um estudo de coorte retrospectivo para desenvolver uma metodologia reproduzível que permitissem aos Cirurgiões Bucomaxilofaciais uma melhor compreensão do planejamento ortognático virtual desde o momento transoperatório até o acompanhamento pós operatório. Desse modo, incluíram na sua amostra casos de cirurgia ortognática bimaxilar (maxilar e mandibular) realizados entre abril de 2012 e agosto de 2015 e para avaliar a variabilidade entre planejamento e a cirurgia usou-se pontos craniofaciais por meio da avaliação tridimensional com a análise de McNamara. Como conclusão, foi possível ter uma confiável avaliação na prática clínica dos resultados trans e pós operatórios por meio do planejamento virtual, pois apenas houve uma variação clinicamente aceitável entre os resultados planejados e obtidos, sugerindo que o planejamento virtual para cirurgia ortognática é uma ferramenta viável com alto grau de acurácia e precisão.

Santos *et al.* (2017) buscaram em seu estudo padronizar normas cefalométricas de adultos brasileiros, incluindo as diferenças anatômicas entre homens e mulheres, por meio da tomografia computadorizada de feixe cônico (TCFC) e utilizaram as medidas da análise de McNamara. Os resultados mostraram que a padronização cefalométrica baseada em TCFC foi confiável para aplicabilidade clínica e em pesquisas. Além disso, os valores possuíram diferenças insignificantes para ambos os sexos, ou seja, possuem morfologia craniofacial semelhante, à exceção de que os homens possuem a região mandibular mais proeminente do que as mulheres.

Entrenas *et al.* (2019) realizaram um estudo de caso-controle prospectivo para descrever o desenvolvimento de crianças entre 8-12 anos com hipoplasia mandibular e a influência do tratamento com aparelho Twin Block (TB). A telerradiografia cefalométrica lateral foi o método de escolha para o acompanhamento de cada criança e para avaliar as vias áreas superiores à análise de McNamara. Nesse estudo, concluiu-se que houve um desenvolvimento significativo no tamanho das vias aéreas superiores após o tratamento precoce com o aparelho TB em pacientes com má oclusão de Classe II mandibular, sendo eficaz em pacientes com manifestações clínicas sugestivas de distúrbios respiratórios do

sono, respiração oral e/ou ronco, pois todos as crianças apresentaram melhora da qualidade respiratória.

4.2 INTELIGÊNCIA ARTIFICIAL

A Inteligência Artificial (IA) é uma área da ciência da computação com objetivo de desenvolver algoritmos de computador para realizar tarefas tradicionalmente associadas com a inteligência humana, como a capacidade de aprender e resolver problemas (YAJI, PRASAD e PAI, 2019). Há alguns anos era utópico pensar em máquinas que pudessem ter raciocínio como pessoas e serem autônomas como o homem, mas é fato que a corrida tecnológica vem causando essa grande transformação (CORVALÁN, 2018).

Foi em 1950 que, o matemático inglês, Alan Turing deu o primeiro passo acerca do assunto com o seu artigo “Computing Machinery and Intelligence” no qual foi proposto um teste comparando o desempenho de um computador e de uma pessoa na resolução de um problema (TURING, 1950). A partir daí, o termo Inteligência Artificial (IA) veio a ser empregado por Mc Carthy *et al.*, em 1955 (MCCARTHY *et al.*, 1955).

A IA envolve várias etapas ou competências: como reconhecer padrões e imagens, entender linguagem aberta escrita e falada, perceber relações e nexos, seguir algoritmos de decisão propostos por especialistas, ser capaz de entender conceitos, adquirir a capacidade de raciocinar pela integração de novas experiências se auto aperfeiçoando (“self-learning”), resolvendo problemas ou realizando tarefas (LOBO, 2018).

É importante saber que a Inteligência Artificial abrange termos como Machine Learning (ML), Artificial Neural Networks (ANN’s), Deep Learning, Convolutional Neural Networks (CNN’s), Computer Aided Detection (CAD), dentre outros, que funcionam como as células deste grande sistema de inteligência. Logo, a IA pode ser definida como a constelação de itens (algoritmos, robótica, redes neurais) que permitem que um software tenha propriedades de inteligência comparáveis as de um ser humano, dentre elas o aprendizado de bancos de dados com mínima interferência humana (FORSTING, 2017), sendo cada vez mais difundida e imperativa na resolução de problemas complexos e ainda em usos mais triviais que ultrapassam fronteiras como um simples software de tradução (OBERMEYER E EMANUEL, 2016).

O Machine Learning (ML), por exemplo, consiste na capacidade de identificar padrões vistos em casos anteriores e a partir destas informações ser possível classificar

uma imagem dentro de grupos pré-estabelecidos, como por exemplo o normal versus patológico (ERICKSON *et al.*, 2017). Para que isso seja possível, cria-se uma função de treinamento para um conjunto de informações e inferência lógica. E realizar esse treinamento requer um teste de referência, isto é, oferecer um modelo "padrão ouro", por exemplo, avaliação histológica para lesões de cárie (SCHWENDICKE *et al.*, 2019). Desse modo, o algoritmo visa a formação de grupos que apresentam padrões semelhantes (PARMAR *et al.*, 2015; LEGER *et al.*, 2017).

As ANN's são elementos inspirados em sistemas nervosos biológicos. Como na natureza, a função de rede é determinada em grande parte pelas conexões entre os elementos. Pode-se treinar uma rede neural para executar uma função particular ajustando os valores das conexões (pesos) entre elementos (BAHAA, NOOR e YOUSIF, 2011). Além do diagnóstico de cárie dentária visualmente confirmada e dentes impactados, estudos aplicando aprendizado de máquina baseado em redes neurais artificiais para tratamento odontológico através de análise de ressonância magnética dental, tomografia computadorizada e telerradiografias cefalométricas laterais estão em andamento, e alguns resultados visíveis estão surgindo num ritmo rápido de comercialização (CHEN *et al.*, 2019).

O Deep Learning se desenvolveu nos últimos anos e é capaz de extrair automaticamente recursos de imagem usando as informações de pixel originais como entrada, principalmente através das CNN's. Esses novos algoritmos reduzem significativamente a carga de trabalho de especialistas humanos e podem extrair certos recursos que são difíceis de serem reconhecidos pelos humanos (SOFFER *et al.*, 2019). É considerada a linha da IA mais promissora por conseguir evoluir e realizar tarefas para às quais não foi treinada através do aumento da sua experiência. Por outro lado, é preciso uma quantidade de imagens (centenas a milhares) muito maior, além da necessidade de que a imagem tenha boa qualidade, associado ainda aos dados clínicos, e apresenta a dificuldade em identificar a lógica utilizada, o que pode afetar o seu desempenho (LEE *et al.*, 2018; MAZUROWSKI *et al.*, 2019).

4.2.1 Inteligência Artificial na Odontologia

A aplicabilidade da IA na Odontologia vem sendo estudada e empregada em diferentes vertentes. As primeiras tentativas promissoras foram feitas na detecção automatizada de cárie em radiografias intraoral (LEE *et al.*, 2018). Sendo assim, Lee *et*

al.(2018) avaliaram a acurácia na detecção e diagnóstico de cárie utilizando o software que envolve CNN's. Os resultados mostraram a evidência do uso desses sistemas, visto que obtiveram um resultado de até 95% de acurácia, ratificando que esse sistema tem potencial para se tornar um método efetivo e eficiente no dia a dia clínico.

O estudo realizado por Ekert *et al.* (2019) mostrou uma discriminação satisfatória na detecção de lesões apicais em panorâmicas por sistema com inteligência artificial envolvendo CNN's. A aplicação desses algoritmos também foi utilizada na detecção de perda óssea por doença periodontal, como apresentado no estudo de Krois *et al.* (2019), onde mostrou que o diagnóstico automatizado foi tão eficaz quanto o diagnóstico pelo profissional, tendo uma acurácia na habilidade de discriminação muito semelhante um do outro.

A Ortodontia é uma grande vertente alvo de software com IA. A aplicação de algoritmos para diagnóstico e marcação cefalométrica automatizados é crescente (LEE *et al.*, 2018; MONTÚFAR, ROMERO e SCUGALL-VILCHIS, 2018; ROSSATO *et al.*, 2018). Kunz *et al.* (2019), desenvolveram um algoritmo para marcação cefalométrica automatizada com base no CNN de forma personalizada e compararam a precisão dessa análise com o padrão ouro atual (análises realizadas por 12 profissionais especialistas) para avaliar a viabilidade de tal sistema na rotina diária ortodôntica. Como resultado, apenas a inclinação dos incisivos superiores mostrou uma diferença estatisticamente maior de 2,18° entre os valores do IA e examinador. Sendo assim, puderam concluir que a precisão das marcações de IA são comparáveis às medições dos 12 profissionais especialistas.

No ano seguinte, Kim *et al.* (2020) desenvolveram um algoritmo para identificação cefalométrica automatizada com base no deep learning. Através da marcação de 23 pontos cefalométricos, os autores avaliaram a acurácia da marcação automatizada, através de um sistema de coordenadas dos pontos, que apresentou uma média de erro de $1,37 \pm 1,79$ mm e sucesso na marcação correta 88,43% da amostra. Concluíram que o algoritmo desenvolvido apresentava alta acurácia com grande potencial de aplicação clínica.

Shahidi *et al.* (2013) realizaram estudos para a avaliação da acurácia em software desenvolvido com as linguagens de programa Delphi e Matlab. Para tal, 40 telerradiografias cefalométricas laterais foram selecionadas e 16 marcações foram identificadas. Ao comparar as marcações automatizadas com as marcações manuais realizadas pelos dois examinadores selecionados e devidamente calibrados, foi

encontrado erro de 2,59 mm, dentre as quais, 12,5% das marcações apresentaram erro menor que 1 mm e 43,75% menor que 2 mm.

A pesquisa recente e revolucionária de Park *et al.* (2019) propôs o estudo da identificação automatizada de marcações cefalométricas, e para isso, o trabalho foi dividido em duas partes. A primeira parte trouxe a comparação de dois algoritmos de marcações diferentes, baseados no deep learning utilizando os algoritmos YOLO version 3 (YOLOv3) e Single Shot Multibox Detector (SSD) para avaliar a acurácia e a performance de cada um. Na primeira parte do estudo, conclui-se que o YOLOv3 teve melhor performance no que tange a acurácia em 38 das 80 marcações realizadas, e as outras 42 não mostraram diferenças estatisticamente significantes. Além disso, o tempo para realização das marcações foi de 0,05 segundos e 2,89 segundos respectivamente para os algoritmos YOLOv3 e SSD. Quando comparados com o que há de mais recente na literatura, esta pesquisa mostrou que o YOLOv3 apresenta os melhores resultados quanto à acurácia (ARIK, IBRAGIMOV e XING, 2017; WANG *et al.*, 2015, 2016).

Na segunda parte do estudo, os autores utilizam o algoritmo YOLOv3, anteriormente citado, comparando-o, agora, com a marcação manual feita por dois examinadores especialistas. Através deste estudo concluiu-se que a marcação realizada com IA mostrou maior acurácia em 14, das 46 telerradiografias cefalométricas laterais. Em 14, de 46, os examinadores mostraram maior acurácia. Nas 18 telerradiografias cefalométricas laterais remanescentes, das 46, não houve diferenças estatisticamente significantes. No que tange a confiabilidade, a IA mostrou realizar as marcações em posições idênticas, enquanto os examinadores mostraram uma variação de $0,97 \pm 1,03$ mm. As variações estatisticamente significantes, quando comparados IA e examinadores, não excederam 0,9 mm, porém, como a IA apresentou reprodutibilidade de excelência, fato não ocorrido entre os examinadores humanos, a IA foi considerada com melhor desempenho. Em sua conclusão, os autores trazem que a IA comportou-se como humano, até mesmo nas dificuldades, onde os examinadores apresentaram alguma dificuldade pela qualidade da imagem, por exemplo, a IA mostrou semelhante comportamento (HWANG *et al.*, 2020).

Hung *et al.* (2019) realizaram uma revisão sistemática com o objetivo de investigar as aplicações clínicas da IA na radiologia oral e fornecer uma atualização acerca do desempenho da IA no diagnóstico na radiologia e imagiologia. Para isso, foram incluídos 50 estudos na revisão entre o período de novembro de 1992 a janeiro de 2019 e concluíram que há um grande potencial e aplicabilidade clínica em vários ramos

da radiologia, desde a detecção de lesão cariiosa e marcação automatizada cefalométrica ao diagnóstico de osteoporose e cistos e tumores odontogênicos. E ressaltou a importância de mais estudos sistemáticos para verificar a reprodutibilidade dos algoritmos antes de implementar na clínica.

Alam e Alfawzan (2020) avaliaram a morfologia do ponto cefalométrico sela túrcica (ST) – é um ponto importante na análise cefalométrica e auxilia a identificar patologias, especificamente síndromes que afetam a região craniofacial – em indivíduos com quatro tipos diferentes de fissuras, e compararam com indivíduos sem fissuras, utilizando a análise cefalométrica automatizada de IA no software WebCephTM (AssembleCircle) . E assim, concluíram que ponto ST, má oclusão e as anomalias dentárias associadas são comuns em indivíduos com fissuras quando comparados ao grupo sem fissuras e que não houve diferença significativa de gênero encontrada nos quatro tipos diferentes de indivíduos com e sem fissura.

Silva *et al.* (2021) realizaram o primeiro estudo que avaliou a confiabilidade do software CEFBOT através das medidas geradas pela análise de Arnett. E dentre os resultados, nenhum dos fatores apresentou $p < 0,05$, sendo possível afirmar que o software em questão possui a capacidade de reproduzir a sua marcação equiparável à capacidade do examinador humano, e, apresenta reprodutibilidade nos dados que fornece. E assim, concluíram que o software com IA pode ser uma excelente ferramenta auxiliar para o Radiologista.

Mahto *et al.* (2022) propuseram avaliar a confiabilidade das medidas lineares e angulares cefalométricas a partir do software WebCephTM (AssembleCircle) e comparar com a marcação manual. Dentre os resultados obtidos, o software avaliado obteve um Índice de Correlação Inter-classe (ICC) considerado alto em todas as medidas avaliadas. Com isso, pode-se concluir que a marcação automatizada utilizando o software WebCephTM (AssembleCircle) possui uma acurácia comparável com a marcação manual.

5 METODOLOGIA

5.1 PRINCÍPIOS ÉTICOS E CÁLCULO AMOSTRAL

Trata-se de um estudo observacional, retrospectivo e comparativo. Este estudo foi cadastrado com (CAAE: 47835221.5.0000.5546) e aprovado pelo Comitê de Ética e Pesquisa em seres humanos sob o parecer 4.919.920, na Universidade Federal de Sergipe (UFS), Hospital Universitário (ANEXO A), estando de acordo com a declaração de Helsinki e realizado conforme a iniciativa STROBE. Devido à natureza retrospectiva deste estudo, o consentimento informado assinado não foi exigido pelo Comitê.

Foi realizado cálculo amostral em que o coeficiente de correlação intraclasse (ICC) foi de 0.70, com poder de teste de 99% e nível de significância de 5%, para tanto, foram necessárias 28 telerradiografias cefalométricas laterais, baseado nos estudos de Durão *et al.* (2015) e Silva *et al.* (2021).

5.2 CARACTERÍSTICAS DA AMOSTRA

Foram selecionadas trinta telerradiografias cefalométricas laterais digitais do arquivo da disciplina de Radiologia, no ambulatório de Radiologia, localizado no Hospital Universitário da Universidade Federal de Sergipe. O critério de inclusão contou com telerradiografias cefalométricas em norma lateral de pacientes sem distinção de sexo ou idade. E como critérios de exclusão foram considerados: (1) mau posicionamento da cabeça no cefalostato; (2) pacientes com deformidades craniofaciais graves e assimetrias faciais; (3) pacientes com pinos, placas ou parafusos sobrepondo estruturas anatômicas de interesse.

5.3 ESTUDO EM DIFERENTES SOFTWARES DE INTELIGÊNCIA ARTIFICIAL (CEFBOT e WEBCEPH)

Após seleção da amostra, as telerradiografias cefalométricas laterais foram salvas nos formatos de imagem JPEG. Sendo então armazenadas e analisadas em um computador de uso pessoal Inspiron 7472 14 polegadas (Dell, Intel Core i5, 8 GB RAM, armazenamento 1TB (5.4K), tela Full HD). Utilizou-se a resolução de 600dpi Bissoli *et al.* (2007).

Para início das análises cefalométricas foi necessária a calibração do examinador, e para tanto, 30% da amostra foi utilizada. O examinador foi treinado previamente por radiologista com mais 20 anos de experiência em traçados cefalométricos computadorizados, e, somente quando o Coeficiente de Correlação Intraclasse (ICC) foi superior a 0.90, o traçado pelo examinador calibrado foi iniciado.

Para que as marcações pudessem ser realizadas, solicitou-se uma licença de uso do software CEFBOT, da empresa RadioMemory e foi realizado cadastro na plataforma WebCeph™, da empresa Assemble Cicle. Assim, foram selecionadas, de acordo com os critérios de inclusão e exclusão, 30 telerradiografias cefalométricas laterais que foram marcadas através do método digital pelo examinador calibrado utilizando a própria plataforma do software Radiocef (RadioMemory). Posteriormente, as mesmas telerradiografias cefalométricas laterais foram marcadas através da IA dos softwares CEFBOT (RadioMemory) e WebCeph™ (AssembleCircle), para avaliação de duas vertentes de marcação: reprodutibilidade e confiabilidade em relação ao examinador e os softwares em questão.

A marcação foi realizada por um examinador devidamente calibrado. O examinador calibrado foi responsável pela identificação dos pontos cefalométricos do grupo controle, com base na análise McNamara (MCNAMARA, 1984) de tal forma a serem geradas grandezas lineares e angulares na plataforma do Radiocef (RadioMemory) (Figuras 1 e 2).

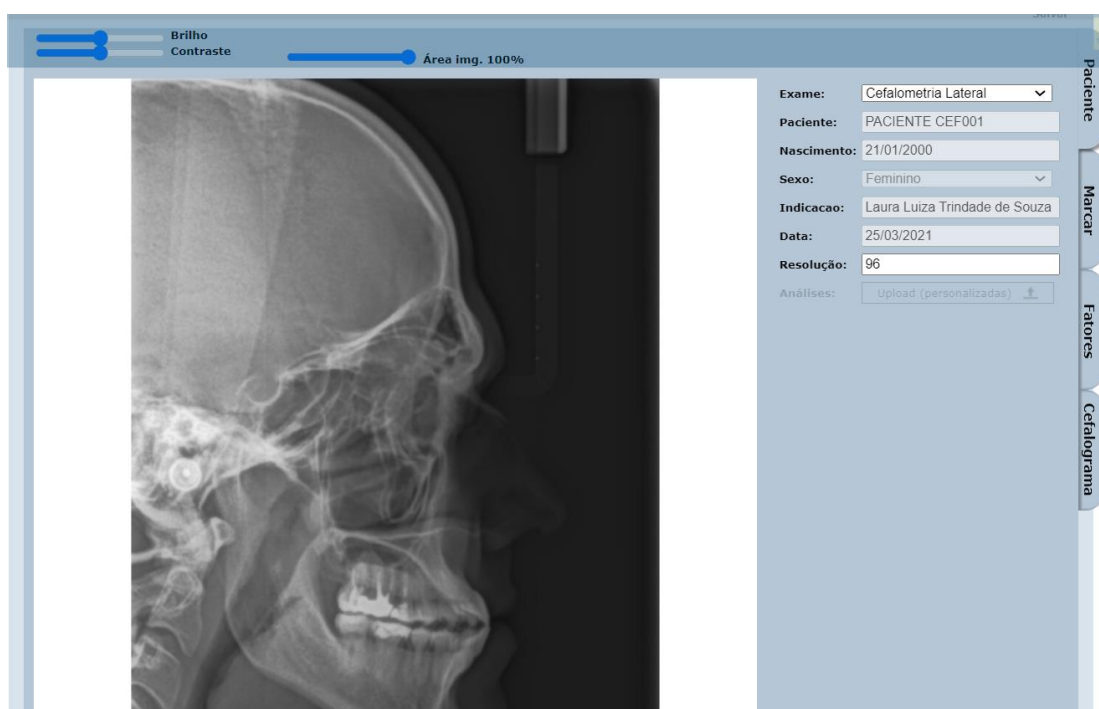


FIGURA 1– Interface do software do CEFBOT.

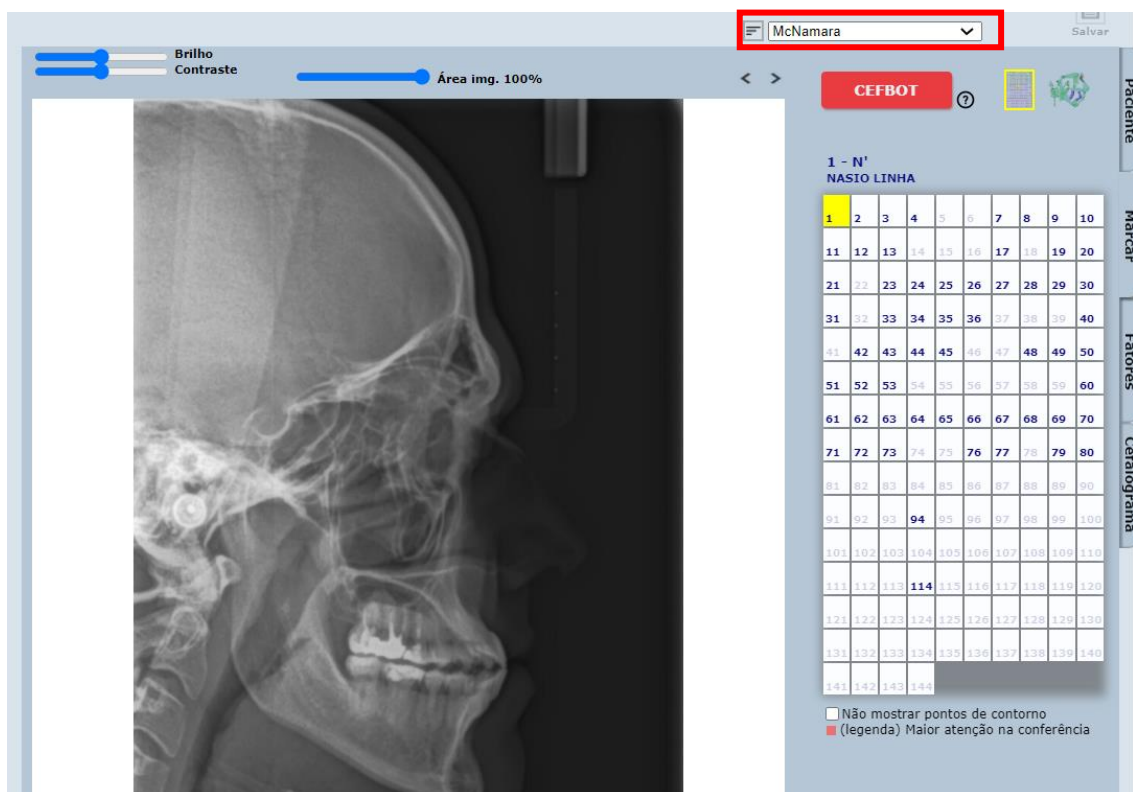


FIGURA 2 – Seleção da análise de McNamara para a marcação pelo examinador.

A mesma amostra foi submetida à marcação automatizada por meio da IA do software CEFBOT (RadioMemory) (Figura 3), e por meio da IA do software WebCephTM (AssembleCircle) (Figura 4, 5 e 6). As grandezas lineares e angulares geradas pelos sistemas através da IA foram armazenadas em planilha no Office Excel 2019 para posterior análise.

Para a avaliação da confiabilidade, após 15 dias do primeiro registro, foi realizada nova marcação pelos examinadores humanos e pelos dois softwares de IA seguindo o mesmo padrão metodológico. Para os dois softwares, os registros anteriormente realizados foram deletados da memória do sistema e, em seguida, as radiografias foram novamente “upadas” e submetidas à marcação automática pela IA do software. Os valores obtidos na primeira e segunda marcação de todos os grupos foram armazenados em planilha no Office Excel 2019 para posterior análise.

5.4 ANÁLISE ESTATÍSTICA

Os procedimentos estatísticos foram realizados no programa estatístico BioEstat 5.3 (Instituto Mamirauá, Belém, Brasil). A calibração do examinador foi realizada pelo

Coeficiente de Correlação Intraclasse (ICC). A confiabilidade da medição executada pelos examinadores e os programas CEFBOT (RadioMemory) e WebCeph™ (AssembleCircle), foram calculadas pelo coeficiente de correlações intraclases (ICC) entre a primeira marcação e a segunda marcação. Com a finalidade de investigar a reprodutibilidade entre os grupos utilizou o teste ANOVA e o teste de Tukey para variáveis independentes. O nível de significância foi estabelecido em $p < 0.05$, para todos os testes realizados.

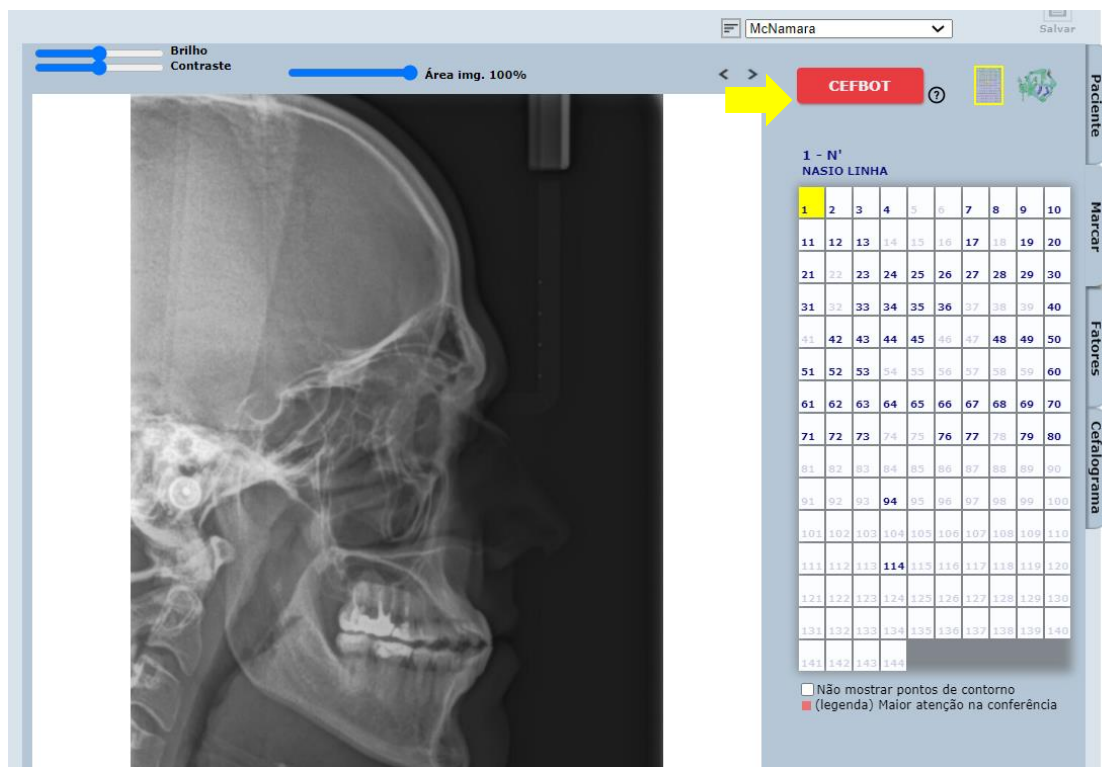


FIGURA 3 – Seleção da análise de McNamara e o botão vermelho "CEFBOT" para a ativação da marcação da I.A.

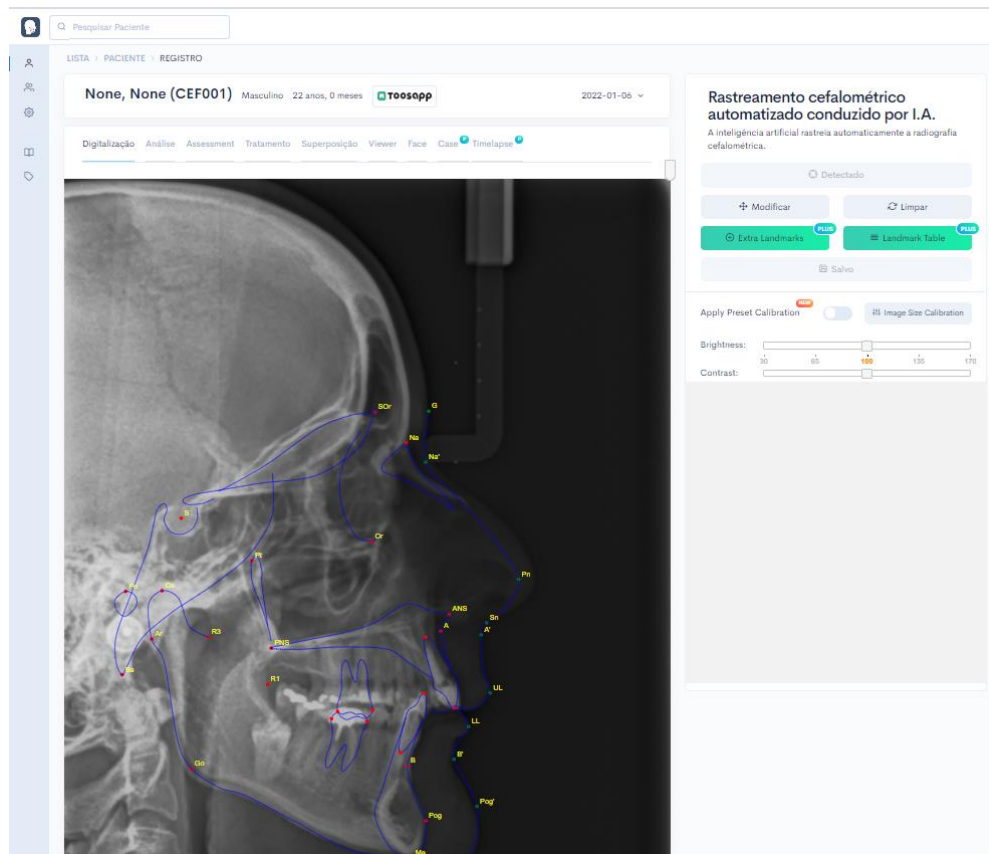


FIGURA 4 – Interface do software WebCeph™ (AssembleCircle).

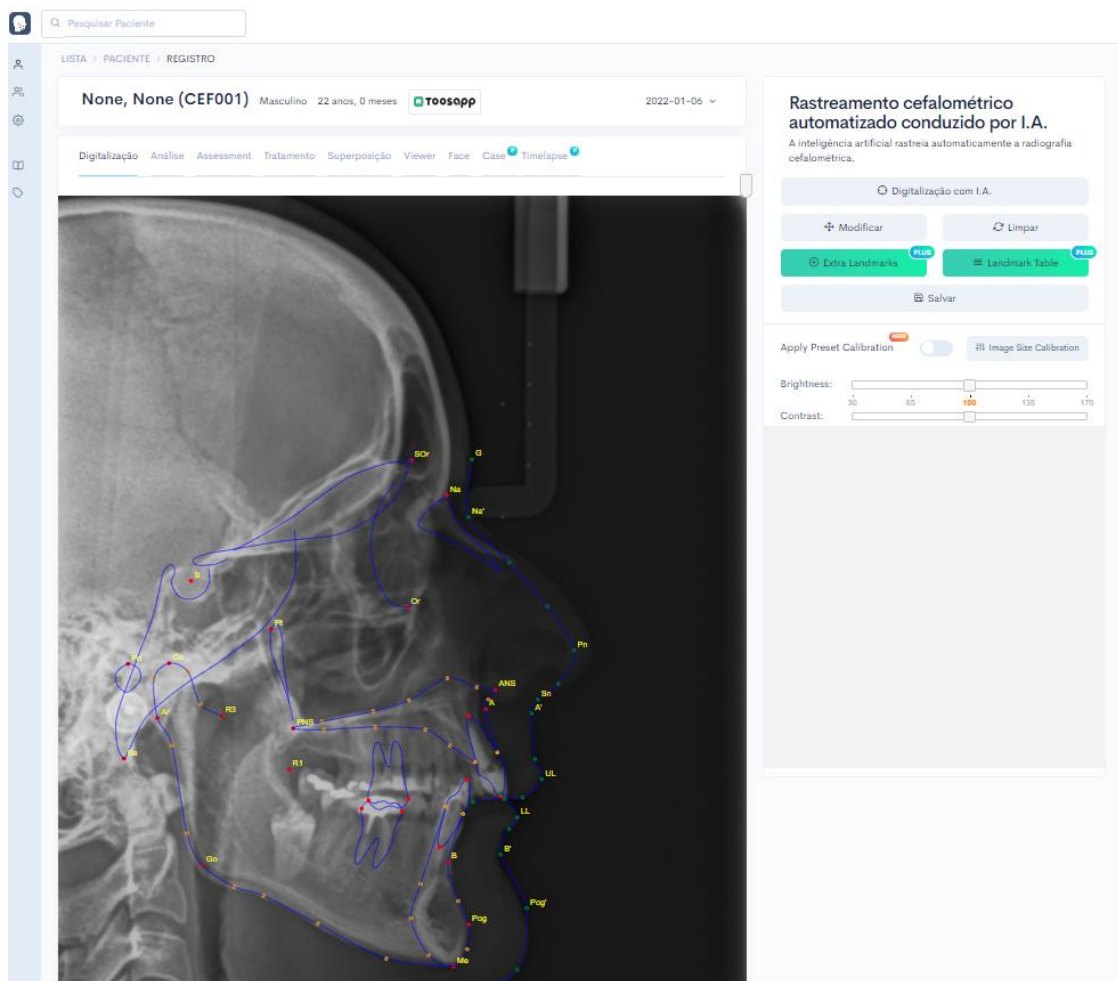


FIGURA 5 – Seleção do botão "Digitalização com I.A" e no botão "Salvar" para ativação da marcação da I.A.

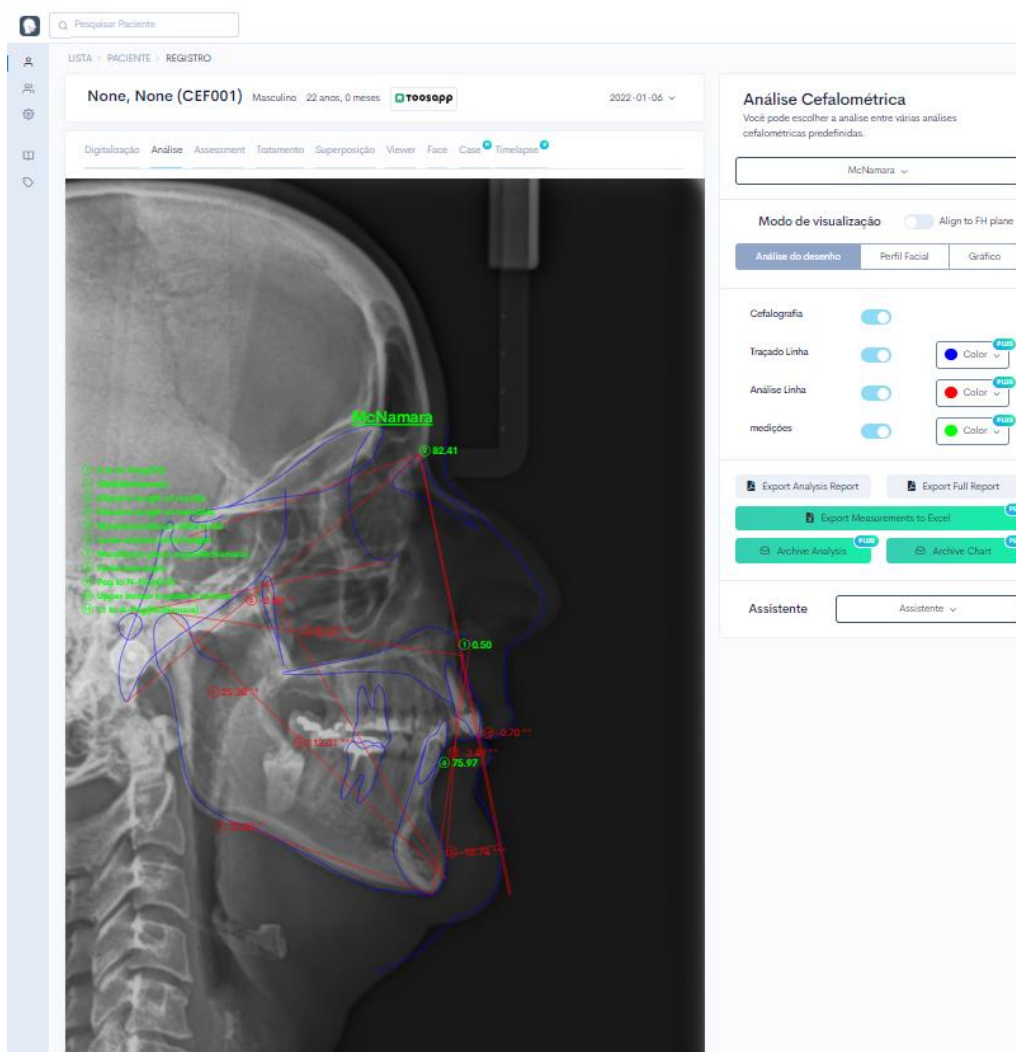


FIGURA 6 – Seleção da análise de McNamara.

6 RESULTADOS

O teste de calibração avaliado pelo coeficiente de correlação Intraclass (ICC) entre o examinador após treinamento e o examinador experiente obteve valor maior que 0.9, correspondente à concordância excelente, de acordo com Landis (1997).

A tabela 1 apresenta os resultados da avaliação da reprodutibilidade dos softwares com o examinador calibrado. Quando comparado o CEFBOT (RadioMemory) com o examinador calibrado, observou-se diferença estatística ($p < 0.01$) apenas no fator um (A-N perpendicular).

Quanto ao WebCeph™ (AssembleCircle), ao comparar com o examinador houve diferença significativa entre os fatores dois ao seis e o dez. E comparado ao CEFBOT (RadioMemory), houve divergência nos mesmos fatores somado ao fator onze. Além

disso, o WebCeph™ (AssembleCircle) não conseguiu marcar os fatores doze (Nfa-Nfp) e treze (Bfa-Bfp) (Tabela 1).

O ICC também foi aplicado para avaliar a confiabilidade das medições nos três grupos (Tabela 2). Ao avaliar a confiabilidade dos diferentes métodos de avaliação, o maior valor de ICC foi identificado no CEFBOT (RadioMemory), apresentando 1.000 para o fator *Prn.(Sn-Ls)*, o que é "quase perfeito" (LANDIS e KOCH, 1977). Por outro lado, o valor de ICC mais baixo foi 0.611, relacionado ao terceiro fator da análise (Co-Gn), todavia, ainda considerado de concordância substancial (LANDIS; KOCH, 1977). Ainda a respeito do CEFBOT (RadioMemory), com exceção do terceiro fator, todos os outros fatores, apresentaram valores considerados "quase perfeito" (LANDIS; KOCH, 1977). No que tange a confiabilidade do WebCeph™ (AssembleCircle), pôde-se observar o valor de ICC mais alto para os fatores nove (Pog-N perpendicular) e dez (SF1/-A perpendicular) da análise, apresentando 0.999, e o menor valor para o primeiro fator (A-N perpendicular), apresentando 0.781. Os fatores doze (Nfa-Nfp) e treze (Bfa-Bfp) não puderam ser avaliados uma vez que não foram gerados pelo WebCeph™ (AssembleCircle) (Tabela 2).

Em relação à confiabilidade do examinador calibrado, o maior valor de ICC foi de 0.991 para o terceiro e o sexto fator, e o menor valor foi de 0.894 para o quinto fator (Tabela 2). Por fim, o ICC médio gerado pelas marcações cefalométrica do examinador calibrado, do CEFBOT e do WebCeph foram 0.960, 0.940 e 0.954, respectivamente, indicando concordância quase perfeita.

TABELA 1 – Valores de média, Desvio Padrão (DP) e teste ANOVA com pós teste de Tukey para variáveis independentes entre o examinador, o programa CEFBOT e WebCeph para análise Mcnamara.

| | Examinador | | CEFBOT | | WebCeph | | E x C | E x W | C x W |
|---------------------------|------------|-------|--------|-------|---------|------|------------|------------|------------|
| | Média | DP | Média | DP | Média | DP | Valor de p | Valor de p | Valor de p |
| 1. A-N perpendicular | 1.09 | 8.36 | 0.27 | 9.06 | -0.05 | 3.46 | NS | NS | NS |
| 2. <i>Prn.(Sn-Ls)</i> | 102.35 | 12.77 | 114.06 | 11.16 | 81.31 | 4.42 | <0.01* | <0.01* | <0.01* |
| 3. Co-Gn | 227.11 | 18.19 | 223.53 | 25.51 | 112.74 | 8.30 | NS | <0.01* | <0.01* |
| 4. Co-A | 170.94 | 9.24 | 169.76 | 9.90 | 82.50 | 4.66 | NS | <0.01* | <0.01* |
| 5. Diferença Mx - Md | 55.36 | 14.63 | 56.78 | 14.55 | 30.24 | 5.88 | NS | <0.01* | <0.01* |
| 6. Ena-Me | 129.40 | 15.04 | 130.53 | 14.86 | 67.98 | 6.33 | NS | <0.01* | <0.01* |
| 7. <i>(Po-Or).(Go-Me)</i> | 24.94 | 5.73 | 25.77 | 5.78 | 24.41 | 5.61 | NS | NS | NS |
| 8. <i>(Ba-N).(Ptm-Gn)</i> | -1.42 | 4.52 | -2.12 | 4.01 | -2.49 | 3.84 | NS | NS | NS |
| 9. Pog-N perpendicular | 0.39 | 16.82 | -1.55 | 15.43 | -5.07 | 5.97 | NS | NS | NS |
| 10. SF1/-A perpendicular | 16.06 | 8.28 | 15.24 | 7.87 | 4.93 | 3.21 | NS | <0.01* | <0.01* |
| 11. <i>lii-(A-Pog)</i> | 6.36 | 7.83 | 7.12 | 7.21 | 2.62 | 3.29 | NS | NS | <0.01* |
| 12. Nfa-Nfp | 27.44 | 7.72 | 28.71 | 8.54 | NI | NI | NS | NI | NI |
| 13. Bfa-Bfp | 19.64 | 6.97 | 22.72 | 6.35 | NI | NI | NS | NI | NI |

Fontes regulares, medições lineares (milímetros). Itálico, medições angulares (graus).

DP, desvio padrão. NI, Não Identificado. NS, diferença não significativa.

*Diferença estatisticamente significativa (p<0.01)

E, examinador. C, CEFBOT. W, WebCeph

TABELA 2- Medidas cefalométricas com grandeza linear e angular aplicando o Índice de Correlação Inter-classe (ICC) para análise de McNamara.

| | Examinador | | CEFBOT | | WebCeph | |
|---------------------------|------------|------------|--------|------------|---------|------------|
| | ICC | Valor de p | ICC | Valor de p | ICC | Valor de p |
| 1. A-N perpendicular | 0.989 | <0.001* | 0.937 | <0.001* | 0.781 | <0.001* |
| 2. <i>Prn.(Sn-Ls)</i> | 0.957 | <0.001* | 1.000 | <0.001* | 0.996 | <0.001* |
| 3. Co-Gn | 0.991 | <0.001* | 0.611 | <0.001* | 0.950 | <0.001* |
| 4. Co-A | 0.945 | <0.001* | 0.961 | <0.001* | 0.916 | <0.001* |
| 5. Diferença Mx - Md | 0.894 | <0.001* | 0.985 | <0.001* | 0.983 | <0.001* |
| 6. Ena-Me | 0.991 | <0.001* | 0.999 | <0.001* | 0.970 | <0.001* |
| 7. <i>(Po-Or).(Go-Me)</i> | 0.983 | <0.001* | 0.987 | <0.001* | 0.997 | <0.001* |
| 8. <i>(Ba-N).(Ptm-Gn)</i> | 0.963 | <0.001* | 0.896 | <0.001* | 0.999 | <0.001* |
| 9. Pog-N perpendicular | 0.967 | <0.001* | 0.954 | <0.001* | 0.999 | <0.001* |
| 10. SF1/-A perpendicular | 0.990 | <0.001* | 0.956 | <0.001* | 0.926 | <0.001* |
| 11. <i>lii-(A-Pog)</i> | 0.960 | <0.001* | 0.971 | <0.001* | 0.981 | <0.001* |
| 12. Nfa-Nfp | 0.930 | <0.001* | 0.974 | <0.001* | PNI | NI |
| 13. Bfa-Bfp | 0.921 | <0.001* | 0.994 | <0.001* | PNI | NI |

Fontes regulares, medições lineares (milímetros). Itálico, medições angulares (graus).

*Diferença estatisticamente significativa (p<0.05)

PNI, Ponto Não Identificado. NI, Não Identificado

7 DISCUSSÃO

A presente pesquisa teve como objetivo comparar a performance dos softwares com inteligência artificial, CEFBOT (RadioMemory) e WebCeph™ (AssembleCircle), de acordo com a avaliação da reprodutibilidade e confiabilidade, das medidas lineares e angulares da análise de McNamara.

Na Ortodontia, as principais considerações para um tratamento ortodôntico eficaz são: um diagnóstico preciso, um excepcional plano de tratamento e um bom prognóstico (SUBRAMANIAN *et al.*, 2022). Dado que muitos tratamentos são irreversíveis ou causam efeitos secundários permanentes, tais como: reabsorção radicular e recessão gengival. Por essa razão, é importante ter ferramentas que aumentem a precisão do tratamento proposto (YU *et al.*, 2020).

Em nosso estudo foram avaliadas as medidas lineares e angulares ao invés da localização dos pontos cefalométricos, uma vez que, as medições são o produto do processo da identificação cefalométrica e fornecem dados para o planejamento e tratamento ortodôntico (ONGKOSUWITO *et al.*, 2002; SANTORO; JARJOURA; CANGIALOSI, 2006; SHAHIDI *et al.*, 2013).

Ao comparar a performance dos dois softwares com o examinador devidamente calibrado foi avaliado a performance de cada um. Assim, dos 13 fatores presentes na análise de McNamara, quando comparado o CEFBOT (RadioMemory) com o examinador, observou-se diferença estatística ($p < 0.01$) apenas na medida Prn.(Sn-Ls). O que demonstra excelente capacidade do software marcar corretamente os pontos cefalométricos da análise em questão.

É importante mencionar o recente estudo de Silva *et al.* (2021), de metodologia semelhante ao nosso, o qual buscou avaliar o CEFBOT (RadioMemory). Em relação à reprodutibilidade, embora o software não tenha gerado um fator da análise de Arnett, os outros fatores não tiveram diferença significativa comparado com o examinador calibrado. Baseado nisso, pode-se perceber que o software é reprodutível, além disso, está em constante evolução, dado que no nosso estudo todos os fatores foram gerados e com excelente resultado.

Em relação à reprodutibilidade do WebCeph™ (AssembleCircle), quando comparado com o examinador houve diferença significativa entre os fatores 2, 3, 4, 5, 6 e o 10. E ao comparar com o CEFBOT (RadioMemory), houve divergência nos mesmos fatores e no 11. Além disso, o WebCeph™ (AssembleCircle) não conseguiu marcar os

fatores 12 e 13 (Tabela 1). E essas diferenças são consideradas significativas para a precisão do diagnóstico.

O estudo de Moon *et al.* (2020) investigou qual o número necessário de dados para a aprendizagem de um software baseado em aprendizado profundo da máquina. Desse modo, concluíram que a acurácia da IA é diretamente proporcional à quantidade de dados disponível no banco de aprendizagem e ao número de pontos que o software consegue identificar. Baseado nisso, podemos afirmar que os fatores não gerados pelo WebCeph™, assim como, os quais não foram considerados confiáveis comparados ao examinador e ao CEFBOT podem ser aprimorados com a ampliação do banco de dados e aprendizagem de máquina.

E para avaliar a capacidade dos softwares marcaram diversas vezes no mesmo local foi avaliada a confiabilidade. Sendo assim, cada software marcou duas vezes a mesma telerradiografia cefalométrica lateral, e mostraram uma concordância quase perfeita em 12 medições e uma concordância substancial em 1 medida (Tabela 2).

Para avaliar a confiabilidade dos dois softwares analisou-se a marcação realizada por um examinador calibrado mostrando uma concordância quase perfeita em todas as medidas (Tabela 2). E nos três casos (examinador, CEFBOT e WebCeph™) foi possível identificar pontos em locais idênticos aos identificados anteriormente, formando-se grandezas lineares e angulares sem diferença significativa como visto na Tabela 2. Logo, é possível afirmar que, em termos de confiabilidade, os dois softwares são equiparáveis ao examinador.

Os nossos resultados em relação à confiabilidade do WebCeph™ foram semelhantes à recente pesquisa de Mahto *et al.* (2022) . Sendo necessário destacar que, apesar do objetivo ter sido avaliar a reprodutibilidade e confiabilidade do software, na verdade, o estudo apenas utilizou o teste estatístico de Índice de Correlação Inter-classe (ICC). Isto é, comparou a capacidade do software identificar diversas vezes o mesmo local. E ao avaliar somente com esse teste não é possível afirmar se os pontos identificados pelo programa estão corretos. E esse fato é importante ser avaliado, dado que, em nosso estudo houve discordância em relação à sua reprodutibilidade, em razão do software ter tido divergência em vários fatores comparado ao CEFBOT (RadioMemory) e ao examinador.

Ao passo que, estudos mais antigos como os de Leonardi, Giordano e Maiorana (2009) e Shahidi *et al.* (2013) demonstraram que a aplicação clínica de softwares em IA ainda não seria viável, visto que os resultados não mostraram muita reprodutibilidade na

marcação dos pontos. Nossos resultados também diferem dos de Hwang *et al.* (2020) no qual o software de IA alcançou uma reprodutibilidade considerada superior à marcação humana. O que podemos afirmar, a partir do nosso estudo, é que os softwares podem complementar na rotina do profissional, mas sem substituí-lo.

Outras pesquisas como a de Kunz *et al.* (2019) e Meriç e Naoumova (2020) puderam concluir que uso de software baseado em IA é uma importante ferramenta auxiliar para otimizar a identificação cefalométrica; ainda que houve diferenças estatísticas em algumas medidas e ambos apresentaram no ângulo SN.GoGn. Curiosamente, dentre os pontos que formam esse ângulo é o SN, esse mesmo ponto faz parte do fator que apresentou o menor ICC do CEFBOT (RadioMemory) e pode-se levantar a hipótese de uma maior dificuldade de o software marcar esse ponto por ser em tecido mole.

A Ortodontia é uma especialidade que continuará a evoluir, especialmente com os avanços da IA (MASSE, 2019). E dentre as vantagens dos softwares de IA, pode-se destacar o fato que uma marcação cefalométrica, realizada manualmente, demanda um tempo médio de 15 a 20 minutos (DREYER; RAYMOND GEIS, 2017). Por consequência, as utilizações de softwares baseados em IA, como o CEFBOT e WebCephTM, são uma excelente ferramenta para otimizar o tempo (Silva *et al.*, 2021). Neste estudo, a identificação e marcação de 59 pontos cefalométricos demorou menos de um minuto. E esse marco não poderia ser alcançado através de um examinador humano, o que torna o uso de softwares baseado em IA relevante para prática clínica.

Outro fato positivo é que o CEFBOT (RadioMemory) é comparável a um examinador humano em termos de reprodutibilidade e confiabilidade. De fato, o software mostrou uma concordância quase perfeita em 12 dos 13 fatores em termo de confiabilidade e apenas 1 fator apresentou reprodutibilidade com diferença estatística comparado ao examinador padrão ouro.

No entanto, a reprodutibilidade do WebCephTM pode ser considerada insuficiente, pelo menos na sua versão atual. O software não conseguiu calcular as duas medidas que avaliam as vias aéreas da análise de McNamara (os espaços nasofaríngeo e bucofaríngeo). Além disso, apresentou diferença estatística em seis fatores comparado ao examinador e ao CEFBOT (RadioMemory).

Como limitações do nosso estudo, podemos apontar que a pesquisa não avaliou a possibilidade de interferência na precisão de marcação na presença de fatores metálicos como aparelhos ortodônticos, grades cirúrgicas, dentre outros, apesar de não ser ter sido

o objetivo do estudo. Assim, novos estudos são sugeridos para elucidar realmente essas questões referentes à presença de fatores, que podem contribuir para diminuição da confiabilidade dos softwares em questão.

8 CONCLUSÃO

A avaliação da reprodutibilidade e confiabilidade das marcações realizadas pelo CEFBOT (RadioMemory) por meio da análise de McNamara mostrou que o software é reprodutível visto que não houve diferenças significativas entre a marcação do software e a marcação do examinador calibrado. A confiabilidade do software, assim como a do examinador calibrado, mostrou que a segunda marcação foi realizada em posição idêntica à marcação anterior, e, portanto, de excelente confiabilidade.

Em relação ao WebCeph™ (AssembleCircle), o software também apresentou uma excelente confiabilidade. No entanto, no que diz respeito à reprodutibilidade o programa apresentou diferença significava e, assim, não conseguiu marcar alguns fatores corretamente nem identificar dois fatores da análise de McNamara. Como alternativa, o profissional pode utilizar a opção de correção manual de pontos de referência fornecida pelo "WebCeph"™ e, assim, pode melhorar a precisão das medidas cefalométricas.

Da perspectiva clínica, o objetivo desse estudo não é substituir o profissional experiente, mas sim, suplementar, colaborar, otimizar e ampliar a performance do profissional, e, portanto, a partir desta pesquisa foram avaliadas a reprodutibilidade e confiabilidade dos softwares CEFBOT (RadioMemory) e WebCeph™ (AssembleCircle), demonstrando, por fim, a sua capacidade de somar no ambiente clínico odontológico.

REFERÊNCIAS

AHMED, M.; SHAIKH, A.; FIDA, M. Diagnostic validity of different cephalometric analyses for assessment of the sagittal skeletal pattern. **Dental Press Journal of Orthodontics**, v. 23, n. 5, p. 75–81, out. 2018.

ALAM, M. K.; ALFAWZAN, A. A. Evaluation of Sella Turcica Bridging and Morphology in Different Types of Cleft Patients. **Frontiers in Cell and Developmental Biology**, v. 8, p. 656, 22 jul. 2020.

ALBARAKATI, S.; KULA, K.; GHONEIMA, A. The reliability and reproducibility of cephalometric measurements: a comparison of conventional and digital methods. **Dentomaxillofacial Radiology**, v. 41, n. 1, p. 11–17, jan. 2012.

ARIK, S. Ö.; IBRAGIMOV, B.; XING, L. Fully automated quantitative cephalometry using convolutional neural networks. **Journal of Medical Imaging**, v. 4, n. 1, p. 1- 11, 6 jan. 2017.

ARNETT, G. W.; BERGMAN, R. T. Facial keys to orthodontic diagnosis and treatment planning. Part I. **American Journal of Orthodontics and Dentofacial Orthopedics**, v. 103, n. 4, p. 299–312, abr. 1993.

BAHAA, K.; NOOR, G.; YOUSIF, Y. The Artificial Intelligence Approach for Diagnosis, Treatment and Modelling in Orthodontic. *In*: NARETTO, S. **Principles in Contemporary Orthodontics**, Dr. Silvano Naretto (Ed.), 2011, p. 451- 492.

BISSOLI, C. F.; TAKESHITA, W.M.; CASTILHO, J.C.M.; MÉDICI-FILHO, E.M.. Digitalização de imagens em radiologia: uma nova visão de futuro. **Revista Odonto**, São Bernardo do Campo, v. 30, n. 15, p. 34-39, jul. 2007.

BORBA, A. M.; HAUPT, D.; ALMEIDA ROMUALDO, L. T. DE; SILVA, A. L. F. DA; GRAÇA NACLÉRIO-HOMEM, M. DA; MILORO, M. How Many Oral and Maxillofacial Surgeons Does It Take to Perform Virtual Orthognathic Surgical Planning? **Journal of Oral and Maxillofacial Surgery**, v. 74, n. 9, p. 1807–1826, set. 2016.

CHEN, H.; ZHANG, K.; LYU, P.; LI, H.; ZHANG, L.; WU, J.; LEE, C.-H. A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films. **Scientific Reports**, v. 9, n. 1, p. 38-40, dez. 2019.

CHEN, S.-K.; CHEN, Y.-J.; YAO, C.-C. J.; CHANG, H.-F. Enhanced Speed and Precision of Measurement in a Computer-Assisted Digital Cephalometric Analysis System. **Angle Orthodontist**, v. 74, n. 4, p. 1-11, 2004.

CHEN, Y.; STANLEY, K.; ATT, W. Artificial intelligence in dentistry: current applications and future perspectives. **Quintessence International**, v. 51, n. 3, p. 248–257, 14 fev. 2020.

CHIEN, P.; PARKS, E.; ERASO, F.; HARTSFIELD, J.; ROBERTS, W.; OFNER, S. Comparison of reliability in anatomical landmark identification using two-dimensional digital cephalometrics and three-dimensional cone beam computed tomography *in vivo*. **Dentomaxillofacial Radiology**, v. 38, n. 5, p. 262–273, jul. 2009.

COLLINS, J.; SHAH, A.; MCCARTHY, C.; SANDLER, J. Comparison of measurements from photographed lateral cephalograms and scanned cephalograms. **American Journal of Orthodontics and Dentofacial Orthopedics**, v. 132, n. 6, p. 830–833, dez. 2007.

CORVALÁN, J. G. Inteligencia artificial: retos, desafíos y oportunidades – Prometea: la primera inteligencia artificial de Latinoamérica al servicio de la Justicia. **Revista de Investigações Constitucionais**, v. 5, n. 1, p. 295, 1 jan. 2018.

DEBELMAS, A.; KETOFF, S.; LANCIAUX, S.; CORRE, P.; FRIESS, M.; KHONSARI, R. H. Reproducibility assessment of Delaire cephalometric analysis using reconstructions from computed tomography. **Journal of Stomatology, Oral and Maxillofacial Surgery**, v. 121, n. 1, p. 35–39, abr. 2019.

DOWNS, W. B. Variations in facial relationships: Their significance in treatment and prognosis. **American Journal of Orthodontics**, v. 34, n. 10, p. 812–840, out. 1948.

DOWNS, W. B. The role of cephalometrics in orthodontic case analysis and diagnosis. **American Journal of Orthodontics**, v. 38, n. 3, p. 162–182, mar. 1952.

DREYER, K. J.; RAYMOND GEIS, J. When machines think: Radiology's next frontier. **Radiology**, v. 285, n. 3, p. 713–718, 1 dez. 2017.

DURÃO, A. P. R.; MOROSOLLI, A.; PITTAYAPAT, P.; BOLSTAD, N.; FERREIRA, A. P.; JACOBS, R. Cephalometric landmark variability among orthodontists and dentomaxillofacial radiologists: a comparative study. **Imaging Science in Dentistry**, v. 45, n. 4, p. 213–20, 2015.

EKERT, T.; KROIS, J.; MEINHOLD, L.; ELHENNAWY, K.; EMARA, R.; GOLLA, T.; SCHWENDICKE, F. Deep Learning for the Radiographic Detection of Apical Lesions. **Journal of Endodontics**, v. 45, n. 7, p. 917- 922, jul. 2019.

ENLOW, D. H.; MOYERS, R. E.; HUNTER, W. S.; MCNAMARA, J. A. A procedure for the analysis of intrinsic facial form and growth an equivalent-balance concept. **American Journal of Orthodontics**, v. 56, n. 1, p. 6–23, jul. 1969.

ENTRENAS, I.; GONZÁLEZ-CHAMORRO, E.; ÁLVAREZ-ABAD, C.; MURIEL, J.; MENÉNDEZ-DÍAZ, I.; COBO, T. Evaluation of changes in the upper airway after Twin Block treatment in patients with Class II malocclusion. **Clinical and Experimental Dental Research**, v. 5, n. 3, p. 259–268, jun. 2019.

ERICKSON, B. J.; KORFIATIS, P.; AKKUS, Z.; KLINE, T. L. Machine Learning for Medical Imaging. **RadioGraphics**, v. 37, n. 2, p. 505–515, mar. 2017.

FAROOQ, M. U.; KHAN, MOHD. A.; IMRAN, S.; SAMEERA, A.; QURESHI, A.; AHMED, S. A.; KUMAR, S.; RAHMAN, MOHD. A. U. Assessing the Reliability of Digitalized Cephalometric Analysis in Comparison with Manual Cephalometric Analysis. **Journal of Clinical and Diagnostic Research**, v. 10, n. 10, p. 20–23, 2016.

FORSTING, M. Machine Learning Will Change Medicine. **Journal of Nuclear Medicine**, v. 58, n. 3, p. 357–358, mar. 2017.

GU, Y.; MCNAMARA, J. A.; SIGLER, L. M.; BACCETTI, T. Comparison of craniofacial characteristics of typical Chinese and Caucasian young adults. **The European Journal of Orthodontics**, v. 33, n. 2, p. 205–211, 1 abr. 2011.

HUNG, K.; MONTALVAO, C.; TANAKA, R.; KAWAI, T.; BORNSTEIN, M. M. The use and performance of artificial intelligence applications in dental and maxillofacial radiology: A systematic review. **Dentomaxillofacial Radiology**, v. 48, p. 22, jul. 2019.

HWANG, H.-W.; PARK, J.-H.; MOON, J.-H.; YU, Y.; KIM, H.; HER, S.-B.; SRINIVASAN, G.; ALJANABI, M. N. A.; DONATELLI, R. E.; LEE, S.-J. Automated identification of cephalometric landmarks: Part 2- Might it be better than human? **The Angle Orthodontist**, v. 90, n. 1, p. 69–76, 14 jan. 2020.

JACOBSON, A. The “Wits” appraisal of jaw disharmony. **American Journal of Orthodontics**, v. 67, n. 2, p. 125–138, fev. 1975.

JACOBSON, A. The “Wits” appraisal of jaw disharmony. **American Journal of Orthodontics and Dentofacial Orthopedics**, v. 124, n. 5, p. 470-479, 2003.

KHAN, A.; JAVED, M. Q.; BILAL, R.; GAIKWAD, R. N. Retrospective quality assurance audit of Lateral Cephalometric Radiographs at postgraduate teaching hospital. **Pakistan Journal of Medical Sciences**, v. 36, n. 7, p. 1601-1606, 28 set. 2020.

KHANAGAR, S. B.; AL-EHAIDEB, A.; MAGANUR, P. C.; VISHWANATHAIAH, S.; PATIL, S.; BAESHEN, H. A.; SARODE, S. C.; BHANDI, S. Developments, application, and performance of artificial intelligence in dentistry – A systematic review. **Journal of Dental Sciences**, v. 16, n. 1, p. 508–522, jan. 2021.

KIM, H.; SHIM, E.; PARK, J.; KIM, Y.-J.; LEE, U.; KIM, Y. Web-based fully automated cephalometric analysis by deep learning. **Computer Methods and Programs in Biomedicine**, v. 194, p. 105-513, out. 2020.

KROIS, J.; EKERT, T.; MEINHOLD, L.; GOLLA, T.; KHARBOT, B.; WITTEMEIER, A.; DÖRFER, C.; SCHWENDICKE, F. Deep Learning for the Radiographic Detection of Periodontal Bone Loss. **Scientific Reports**, v. 9, n. 1, p. 84-95, dez. 2019.

KUNZ, F.; STELLZIG-EISENHAUER, A.; ZEMAN, F.; BOLDT, J. Artificial intelligence in orthodontics: Evaluation of a fully automated cephalometric analysis using a customized convolutional neural network. **Journal of Orofacial Orthopedics / Fortschritte der Kieferorthopädie**, v. 81, n. 1, p. 52–68, dez. 2019.

LEE, J.-H.; KIM, D.-H.; JEONG, S.-N.; CHOI, S.-H. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. **Journal of Dentistry**, v. 77, p. 106–111, out. 2018.

LEGER, S. *et al.* A comparative study of machine learning methods for time-to-event survival data for radiomics risk modelling. **Scientific Reports**, v. 7, n. 1, p. 1-11, dez. 2017.

LEONARDI, R.; GIORDANO, D.; MAIORANA, F. An evaluation of cellular neural networks for the automatic identification of cephalometric landmarks on digital images. **Journal of Biomedicine and Biotechnology**, v. 2009, 2009.

LIVAS, C.; DELLI, K.; SPIJKERVET, F. K. L.; VISSINK, A.; DIJKSTRA, P. U. Concurrent validity and reliability of cephalometric analysis using smartphone apps and computer software. **The Angle Orthodontist**, v. 89, n. 6, p. 889–896, 1 nov. 2019.

LOBO, L. C. Inteligência artificial, o Futuro da Medicina e a Educação Médica. **Revista Brasileira de Educação Médica**, v. 42, n. 3, p. 3–8, set. 2018.

LUNDSTROM, A.; LUNDSTROM, F.; LEBRET, L. M. L.; MOORREES, C. F. A. Natural head position and natural head orientation: basic considerations in cephalometric analysis and research. **The European Journal of Orthodontics**, v. 17, n. 2, p. 111–120, 1 abr. 1995.

MAHTO, R. K.; KAFLE, D.; GIRI, A.; LUNTEL, S.; KARKI, A. Evaluation of fully automated cephalometric measurements obtained from web-based artificial intelligence driven platform. **BMC Oral Health**, v. 22, n. 1, 1 dez. 2022.

MASSE, J.-F. Will the orthodontic profession disappear? **Journal of Dental Sleep Medicine**, v. 6, n. 2, 10 abr. 2019.

MAZUROWSKI, M. A.; BUDA, M.; SAHA, A.; BASHIR, M. R. Deep Learning In Radiology: An Overview Of The Concepts And A Survey Of The State Of The Art With Focus On MRI. **Journal of Magnetic Resonance Imaging**, v. 49, n. 4, p. 939–954, abr. 2019.

MCCARTHY, J.; MINSKY, M. L.; ROCHESTER, N.; CORPORATION, I. B. M.; SHANNON, C. E. **A Proposal For The Dartmouth Summer Research Project On Artificial Intelligence**. p. 12-14, ago. 1955.

MCNAMARA, A. A method of cephalometric evaluation. **American Journal of Orthodontics**, n. 6, p. 449-469, dez. 1984.

MERICİ, P.; NAOUMOVA, J. Web-based Fully Automated Cephalometric Analysis: Comparisons between App-aided, Computerized, and Manual Tracings. **Turkish Journal of Orthodontics**, v. 33, n. 3, p. 142–149, 1 set. 2020.

MONTÚFAR, J.; ROMERO, M.; SCOUGALL-VILCHIS, R. J. Automatic 3-dimensional cephalometric landmarking based on active shape models in related

projections. **American Journal of Orthodontics and Dentofacial Orthopedics**, v. 153, n. 3, p. 449–458, mar. 2018.

MOON, J. H.; HWANG, H. W.; YU, Y.; KIM, M. G.; DONATELLI, R. E.; LEE, S. J. How much deep learning is enough for automatic identification to be reliable? A cephalometric example. **Angle Orthodontist**, v. 90, n. 6, p. 823–830, 1 nov. 2020.

MOSLEH, M. A. A.; BABA, M. S.; MALEK, S.; ALMAKTARI, R. A. Ceph-X: development and evaluation of 2D cephalometric system. **BMC Bioinformatics**, v. 17, n.19, p. 194-295, dez. 2016.

NOURI, M.; HAMIDIAVAL, S.; ALIREZA BAGHBAN; BASAFA, M.; FAHIM, M. Efficacy of a Newly Designed Cephalometric Analysis Software for McNamara Analysis in Comparison with Dolphin Software. **Journal of Dentistry of Tehran University of Medical Sciences**, v. 12, n. 1, p. 60–69, 2015.

OBERMEYER, Z.; EMANUEL, E. J. Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. **New England Journal of Medicine**, v. 375, n. 13, p. 1216–1219, 29 set. 2016.

OLMEZ, H.; GORGULU, S.; AKIN, E.; BENGI, A. O.; TEKDEMIR, İ.; ORS, F. Measurement accuracy of a computer-assisted three-dimensional analysis and a conventional two-dimensional method. **The Angle Orthodontist**, v. 81, n. 3, p. 375–382, maio 2011.

ONGKOSUWITO, E. M.; KATSAROS, C.; VAN'T HOF, M. A.; BODEGOM, J. C.; KUIJPERS-JAGTMAN, A. M. The reproducibility of cephalometric measurements: a comparison of analogue and digital methods. **European Journal of Orthodontics**, v. 24, p. 655–665, 2002.

PARK, J.-H.; HWANG, H.-W.; MOON, J.-H.; YU, Y.; KIM, H.; HER, S.-B.; SRINIVASAN, G.; ALJANABI, M. N. A.; DONATELLI, R. E.; LEE, S.-J. Automated identification of cephalometric landmarks: Part 1—Comparisons between the latest deep-learning methods YOLOV3 and SSD. **The Angle Orthodontist**, v. 89, n. 6, p. 903–909, 1 nov. 2019.

PARMAR, C.; GROSSMANN, P.; BUSSINK, J.; LAMBIN, P.; AERTS, H. J. W. L. Machine Learning methods for Quantitative Radiomic Biomarkers. **Scientific Reports**, v. 5, n. 1, p. 1-11, out. 2015.

RAVIKUMAR, D.; N., S.; RAMAKRISHNA, M.; SHARNA, N.; ROBINDRO, W. Evaluation of McNamara's analysis in South Indian (Tamil Nadu) children between 8–12 years of age using lateral cephalograms. **Journal of Oral Biology and Craniofacial Research**, v. 9, n. 2, p. 193–197, abr. 2019.

RICKETTS, R. M. The Influence Of Orthodontic Treatment On Facial Growth And Development. **Angle Orthod**, v. 30, n. 1, p. 103-33, 1960.

RIEDEL, R. A. The Relation of Maxillary structures to cranium in malocclusion and in normal occlusion. **Angle Orthod**, v. 22, n. 3, p. 142-5, 1952.

ROSSATO, P. H.; FERNANDES, T. M. F.; URNAU, F. D. A.; CASTRO, A. C. DE; CONTI, F.; ALMEIDA, R. R. DE; OLTRAMARI-NAVARRO, P. V. P. Dentoalveolar effects produced by different appliances on early treatment of anterior open bite: A randomized clinical trial. **The Angle Orthodontist**, v. 88, n. 6, p. 684–691, 1 nov. 2018.

SANTORO, M.; JARJOURA, K.; CANGIALOSI, T. J. Accuracy of digital and analogue cephalometric measurements assessed with the sandwich technique. **American Journal of Orthodontics and Dentofacial Orthopedics**, v. 129, n. 3, p. 345–351, mar. 2006.

SANTOS, R. M. G.; DE MARTINO, J. M.; HAITER NETO, F.; PASSERI, L. A. Cone beam computed tomography-based cephalometric norms for Brazilian adults. **International Journal of Oral and Maxillofacial Surgery**, v. 47, n. 1, p. 1–8, 2017.

SCHWENDICKE, F.; GOLLA, T.; DREHER, M.; KROIS, J. Convolutional neural networks for dental image diagnostics: A scoping review. **Journal of Dentistry**, v. 91, p. 103-226, dez. 2019.

SHAHIDI, S.; OSHAGH, M.; GOZIN, F.; SALEHI, P.; DANAEI, S. M. Accuracy of computerized automatic identification of cephalometric landmarks by a designed software. **Dentomaxillofacial Radiology**, v. 42, n. 1, p. 20110187–20110187, jan. 2013.

SHETTIGAR, P.; SHETTY, S.; D. NAIK, R.; M. BASAVARADDI, S.; K. PATIL, A. A Comparative Evaluation of Reliability of an Android-based App and Computerized Cephalometric Tracing Program for Orthodontic Cephalometric Analysis. **Biomedical and Pharmacology Journal**, v. 12, n. 1, p. 341–346, 28 mar. 2019.

SOFFER, S.; BEN-COHEN, A.; SHIMON, O.; AMITAI, M. M.; GREENSPAN, H.; KLANG, E. Convolutional Neural Networks for Radiologic Images: A Radiologist's Guide. **Radiology**, v. 290, n. 3, p. 590–606, mar. 2019.

SILVA, T. P.; HUGHES, M. M.; MENEZES, L. DOS S.; MELO, M. DE F. B. DE; TAKESHITA, W. M.; FREITAS, P. H. L. DE. Artificial Intelligence-Based Cephalometric Landmark Annotation and Measurements According to Arnett's Analysis: Can we trust a bot to do that? **Dentomaxillofacial Radiology**, v. 50, n. 20200548, p. 1-6, 21 abr. 2021.

STEINER, C. C. Cephalometrics for you and me. **American Journal of Orthodontics**, v. 39, n. 10, p. 729–755, out. 1953.

STORNIOLO-SOUZA, J. M.; SEMINARIO, M. P.; PINZAN-VERCELINO, C. R. M.; PINZAN, A.; JANSON, G. McNamara analysis cephalometric parameters in White-Brasilians, Japanese and Japanese-Brasilians with normal occlusion. **Dental Press Journal of Orthodontics**, v. 26, n. 1, p. 1–22, 2021.

SUBRAMANIAN, A. K.; CHEN, Y.; ALMALKI, A.; SIVAMURTHY, G.; KAFLE, D. Cephalometric Analysis in Orthodontics Using Artificial Intelligence—A Comprehensive Review. **BioMed Research International**, v. 2022, p. 1–9, 16 jun. 2022.

TAMADA, I.; NAKAJIMA, H.; OGATA, H.; NAKAJIMA, T.; SAKAMOTO, T.; ISHII, T. How to revise and utilise the cephalogram for craniofacial dysostosis—Modification of the porion and the McNamara line. **Journal of Cranio-Maxillofacial Surgery**, v. 38, n. 6, p. 441–451, set. 2010.

TURING, A. M. I. Computing Machinery And Intelligence. **Mind**, v. LIX, n. 236, p. 433–460, 1 out. 1950.

WANG, C.W.; HUANG, C.T.; LEE, J.H.; et al. Evaluation and Comparison of Anatomical Landmark Detection Methods for Cephalometric X-Ray Images: A Grand Challenge. **IEEE Transactions on Medical Imaging**, v. 34, n. 9, p. 1890–1900, set. 2015.

WANG, C.W.; HUANG, C.T.; LEE, J.H.; et al. A benchmark for comparison of dental radiography analysis algorithms. **Medical Image Analysis**, v. 31, p. 63–76, jul. 2016.

YAJI, A.; PRASAD, S.; PAI, A. Artificial Intelligence in Dento-Maxillofacial Radiology. **Acta Scient Dent Scienc**, v. 3, n. 1, p. 116-121, 2019.

YU, H. J.; CHO, S. R.; KIM, M. J.; KIM, W. H.; KIM, J. W.; CHOI, J. Automated Skeletal Classification with Lateral Cephalometry Based on Artificial Intelligence. **Journal of Dental Research**, v. 99, n. 3, p. 249–256, 1 mar. 2020.

ZAMRIK, O. M.; ISERI, H. The reliability and reproducibility of an Android cephalometric smartphone application in comparison with the conventional method. **Angle Orthodontist**, v. 91, n. 2, p. 236–242, 1 mar. 2021.

ANEXO A – APROVAÇÃO COMITÊ DE ÉTICA



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: AVALIAÇÃO DA CONFIABILIDADE E REPRODUTIBILIDADE DAS MEDIDAS LINEARES E ANGULARES EM RADIOGRAFIAS CEFALOMÉTRIAS LATERAIS POR MEIO DE SOFTWARE COM INTELIGÊNCIA ARTIFICIAL

Pesquisador: Wilton Mitsunari Takeshita

Área Temática:

Versão: 1

CAAE: 47835221.5.0000.5546

Instituição Proponente: FUNDACAO UNIVERSIDADE FEDERAL DE SERGIPE

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 4.919.920

Apresentação do Projeto:

As informações elencadas nos campos "Apresentação do Projeto", "Objetivo da Pesquisa" e "Avaliação dos Riscos e Benefícios" foram retiradas do arquivo "Informações Básicas da Pesquisa" (PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1770101.pdf) e do "Projeto Detalhado / Brochura Investigador" (Projeto_CEFBOT_CEP.pdf), postados em 09/06/2021.

INTRODUÇÃO

A Cefalometria é uma grande aliada na busca de informações para o diagnóstico ortodôntico e cirúrgico visto que a partir dela são realizadas análises do desenvolvimento craniofacial, estudo das maloclusões dentárias e das discrepâncias esqueléticas (COLLINS et al., 2007). Utilizada desde a introdução do cefalostato, que permitiu a padronização do posicionamento da cabeça do paciente durante a execução da técnica radiográfica, esse exame radiográfico de análise bidimensional tomou-se uma ferramenta clínica e de pesquisa essencial na Odontologia (BROADBENT, 1931; CHIEN et al., 2009; OLMEZ et al., 2011). A avaliação cefalométrica deve ser realizada de forma cautelosa pois erros podem levar a grandes discrepâncias no diagnóstico e planejamento (KUMAR et al., 2008; LUDLOW et al., 2009). O tempo demandado pelo método manual dos traçados cefalométricos e a subjetividade das marcações realizadas pelo profissional representam fatores limitantes na prática clínica (CHEN et al., 2004). Nesse cenário, softwares que

Endereço: Rua Cláudio Batista s/nº
Bairro: Sanatório **CEP:** 49.060-110
UF: SE **Município:** ARACAJU
Telefone: (79)3194-7208 **E-mail:** cep@academico.ufs.br

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

| Tipo Documento | Arquivo | Postagem | Autor | Situação |
|---|---|------------------------|-------------------------------|----------|
| Informações Básicas do Projeto | PB_INFORMAÇÕES_BASICAS_DO_PROJETO_1770101.pdf | 09/06/2021 00:39:48 | | Aceito |
| Projeto Detalhado / Brochura Investigador | Projeto_CEFBOT_CEP.pdf | 09/06/2021 00:32:23 | Wilton Mitsunari Takeshita | Aceito |
| Outros | TERMO_DE_AUTORIZAÇÃO_acesso_arquivos.pdf | 09/06/2021 00:31:12 | Wilton Mitsunari Takeshita | Aceito |
| Declaração de Pesquisadores | TERMO_DE_COMPROMISSO_pibic_2021.pdf | 09/06/2021 00:28:13 | Wilton Mitsunari Takeshita | Aceito |
| TCLE / Termos de Assentimento / Justificativa de Ausência | DispensaTCLE_radiografia_arquivo.pdf | 09/06/2021 00:27:05 | Wilton Mitsunari Takeshita | Aceito |
| Declaração de Instituição e | Autorizacao_UFS_PIBIC2021.pdf | 09/06/2021 00:26:46 | Wilton Mitsunari Takeshita | Aceito |

Endereço: Rua Claudio Batista s/n°
 Bairro: Sanatório CEP: 49.060-110
 UF: SE Município: ARACAJU
 Telefone: (79)3194-7208 E-mail: oep@academico.ufs.br

Página 06 de 07



Continuação do Parecer: 4.919.920

| | | | | |
|----------------|-------------------------------------|------------------------|-------------------------------|--------|
| Infraestrutura | Autorizacao_UFS_PIBIC2021.pdf | 09/06/2021 00:26:46 | Wilton Mitsunari Takeshita | Aceito |
| Folha de Rosto | Folhaderosto_assinada_carimbada.pdf | 09/06/2021 00:25:23 | Wilton Mitsunari Takeshita | Aceito |

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

ARACAJU, 20 de Agosto de 2021

Assinado por:
FRANCISCO DE ASSIS PEREIRA
 (Coordenador(a))