



Doctor Bone: treinamento de redes neurais para auxílio na identificação da idade óssea

Doctor Bone: training neural networks to assist in determining bone age

Doctor Bone: entrenamiento de redes neuronales para ayudar en la determinación de la edad óssea

Rodrigo Lages Barbosa¹, Heitor de Castro Teixeira e Martins², Felipe Cassiano Barbosa¹, Beatriz Torres da Costa¹, Rolf Freitas Matela¹, José Fernando Rodrigues Ferreira Neto³, Yuri Nekan Soares Fontes⁴, João Alexandre Lobo Marques⁵, João Batista Furlan Duarte⁶, Joel Sotero da Cunha Neto⁷

¹ Estagiário, Vice-Reitoria de Pesquisa, Universidade de Fortaleza, Fortaleza (CE), Brasil.

² Graduando em ciência da computação, Centro de Ciências Tecnológicas, Universidade de Fortaleza, Fortaleza (CE), Brasil.

³ Coordenador de projetos, Vice-Reitoria de Pesquisa, Universidade de Fortaleza, Fortaleza (CE), Brasil.

⁴ Analista de projetos, Vice-Reitoria de Pesquisa, Universidade de Fortaleza, Fortaleza (CE), Brasil.

⁵ Doutor, Laboratory of Applied Neurosciences, University of Saint Joseph, Macao, China.

⁶ Doutor, Centro de Ciências Tecnológicas, Universidade de Fortaleza, Fortaleza (CE), Brasil.

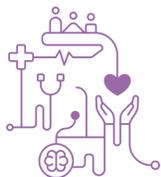
⁷ Mestre, Centro de Ciências Tecnológicas, Universidade de Fortaleza, Fortaleza (CE), Brasil.

Autor correspondente: Prof. Joel Sotero da Cunha Neto

E-mail: joelsotero@unifor.br

Resumo

Objetivo: Explorar a aplicação de inteligência artificial (IA) na predição da idade óssea a partir de imagens de raios-X. Método: Utilizou-se a Metodologia Interdisciplinar para o Desenvolvimento de Tecnologias em Saúde (MIDTS) para desenvolver uma ferramenta de predição. O treinamento foi realizado com redes neurais convolucionais (CNNs) usando um conjunto de dados de 14.036 imagens de raios-X. Resultados: A ferramenta alcançou um coeficiente de determinação (R^2) de 0,94807 e um Erro Médio Absoluto (MAE) de 6,97, destacando sua precisão e potencial de aplicação clínica. Conclusão: O projeto demonstrou grande potencial para aprimorar a predição da idade



óssea, com possibilidades de evolução conforme a base de dados aumenta e a IA se torna mais sofisticada.

Descritores: Inteligência artificial; Auxílio ao Diagnóstico; Idade Óssea

Abstract

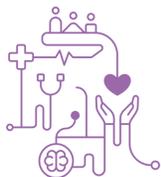
Objective: To explore the application of artificial intelligence (AI) in predicting bone age from X-ray images. **Method:** The Interdisciplinary Methodology for the Development of Health Technologies (MIDTS) was used to develop a prediction tool. Training was conducted with convolutional neural networks (CNNs) using a dataset of 14,036 X-ray images. **Results:** The tool achieved a coefficient of determination (R^2) of 0.94807 and a Mean Absolute Error (MAE) of 6.97, highlighting its accuracy and clinical potential. **Conclusion:** The project demonstrated great potential to enhance bone age prediction, with possibilities for evolution as the database grows and AI becomes more sophisticated.

Keywords: Artificial intelligence; Diagnostic Aid; Bone Age

Resumen

Objetivo: Explorar la aplicación de la inteligencia artificial (IA) en la predicción de la edad ósea a partir de imágenes de rayos X. **Método:** Se utilizó la Metodología Interdisciplinaria para el Desarrollo de Tecnologías en Salud (MIDTS) para desarrollar una herramienta de predicción. El entrenamiento se realizó con redes neuronales convolucionales (CNN) utilizando un conjunto de datos de 14,036 imágenes de rayos X. **Resultados:** La herramienta alcanzó un coeficiente de determinación (R^2) de 0.94807 y un Error Medio Absoluto (MAE) de 6.97, destacando su precisión y potencial clínico. **Conclusión:** El proyecto demostró un gran potencial para mejorar la predicción de la edad ósea, con posibilidades de evolución a medida que la base de datos crezca y la IA se vuelva más sofisticada.

Descriptores: Inteligencia artificial; Ayuda Diagnóstica; Edad Óssea

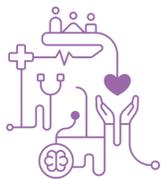


Introdução

A análise de raios X de mão é essencial para diversas aplicações clínicas, incluindo a avaliação do crescimento infantil e a detecção de anomalias ósseas (1). Uma das avaliações é a idade óssea, que é crucial na prática médica, principalmente em pediatria, ortopedia e endocrinologia. Em pediatria, a avaliação da idade óssea é essencial para monitorar o crescimento e desenvolvimento, diagnosticar distúrbios do crescimento e avaliar a eficácia da terapia com hormônio do crescimento (2). Além disso, em ortopedia, a avaliação precisa da idade óssea é vital para avaliar a maturidade esquelética, o que é crucial para o planejamento cirúrgico e para prever o resultado de intervenções ortopédicas (3). Além disso, em endocrinologia, a avaliação da idade óssea apoia no diagnóstico de distúrbios endócrinos e no monitoramento dos efeitos desses distúrbios no desenvolvimento esquelético (2).

A avaliação manual da idade óssea está associada a vários desafios e limitações, incluindo o potencial para erros e variabilidade interpessoal nas interpretações de especialistas. Estes desafios suscitaram uma necessidade urgente de métodos automatizados para aliviar o trabalho manual e resolver o problema da variabilidade intra e interobservador (4), que pode levar a inconsistências nos resultados da avaliação (5). Um dos principais desafios na validação de métodos automatizados de avaliação da idade óssea é a ausência de um padrão ouro para a classificação da idade óssea, tornando difícil comparar diretamente as classificações automatizadas com as classificações manuais e demonstrar a substituição da classificação manual por métodos automatizados (6).

As abordagens manuais e semiautomáticas de análises teciduais frequentemente apresentam subjetividade intra e interobservador, além de variabilidade nos métodos de análise de imagens, evidenciando as limitações da avaliação manual na histomorfometria óssea (7). A interpretação da idade óssea torna-se ainda mais desafiadora na presença de anormalidades morfológicas significativas, como nas displasias esqueléticas, ressaltando as limitações da avaliação manual da idade óssea nesses casos (8).

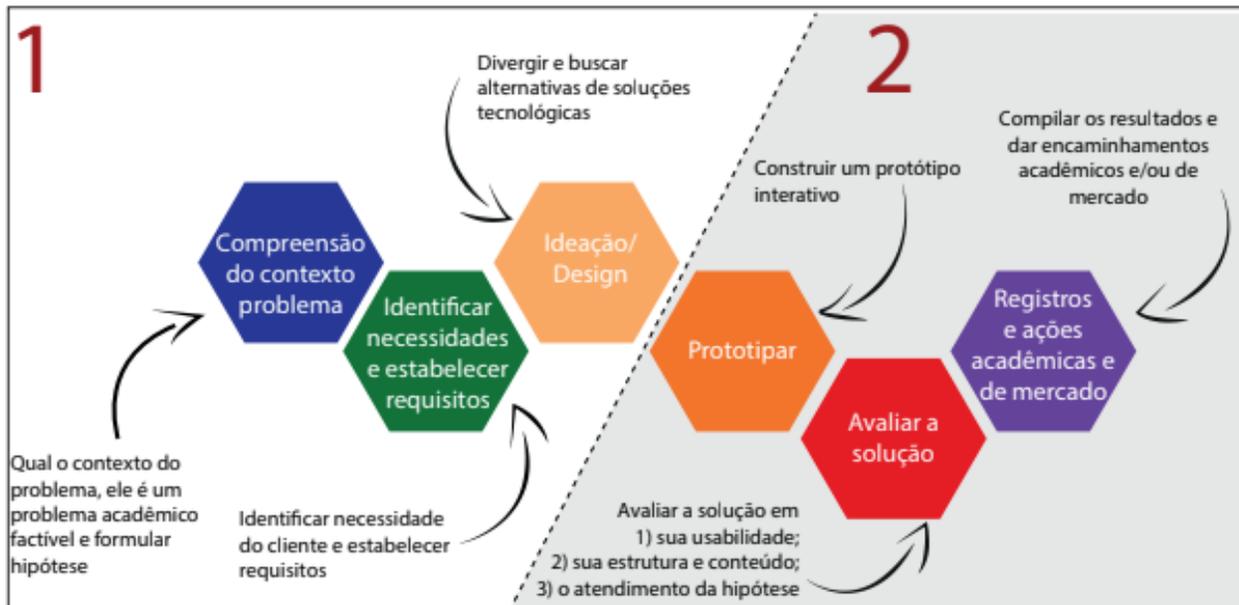


A IA tem revolucionado diversas áreas da medicina, oferecendo novas ferramentas e metodologias para diagnóstico e tratamento de doenças. Um dos campos que mais se beneficia dessa tecnologia é a análise de imagens médicas, onde algoritmos de aprendizado profundo (*Deep Learning*) têm demonstrado capacidades superiores na interpretação e avaliação de exames de imagem (9). Nesse contexto, o projeto *Doctor Bone* surge como uma tecnologia disruptiva, voltada para a predição da idade óssea, um componente crítico na prática médica pediátrica e ortopédica.

Métodos

Para desenvolver a solução, foi utilizada a Metodologia Interdisciplinar para o Desenvolvimento de Tecnologias em Saúde - MIDTS (10), organizada em duas fases, conforme mostra a Figura 1. Primeiramente, a metodologia visa uma compreensão detalhada do problema através da análise do contexto, identificação das necessidades e geração de ideias. A etapa seguinte envolve a criação de protótipos, avaliação e documentação dos resultados, onde a solução proposta é aprimorada e, se necessário, pode retornar às fases anteriores até ser considerada adequada e seus resultados serem registrados.

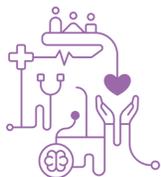
Figura 1 - Representação do MIDTS (10).



Reuniões com profissionais da área durante a fase inicial mostraram que a análise da idade óssea em raios X carece de um padrão ouro e está sujeita a viés médico. Esse problema se reflete na variabilidade e inconsistência dos diagnósticos, que dependem fortemente da experiência individual dos radiologistas. A falta de uniformidade nos laudos pode levar a erros na avaliação do crescimento infantil e na detecção de anomalias ósseas, impactando negativamente o tratamento dos pacientes.

A segunda etapa do MIDTS (10) focou na identificação das necessidades dos clientes, abrangendo uma meticulosa coleta de requisitos essenciais para o desenvolvimento do projeto, optando-se por um sistema em forma de aplicativo, visando a facilidade de acesso por parte dos usuários. Esses requisitos incluíam a capacidade de fazer *upload* de exames radiográficos através da câmera ou galeria de imagens e analisá-los usando um modelo de IA pré-treinado para determinar a idade óssea. Além disso, o aplicativo deveria permitir que os usuários reportassem possíveis erros de previsão, fornecendo a idade óssea corrigida.

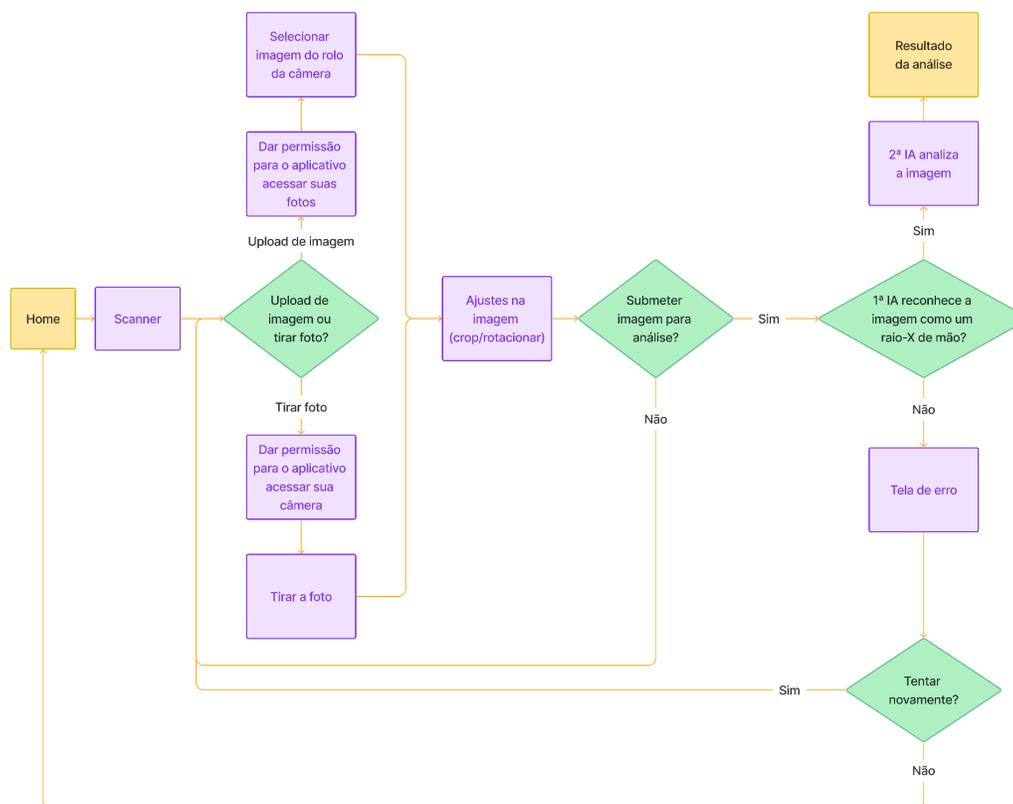
O aplicativo *Doctor Bone* foi idealizado durante a terceira fase do método. Seus fluxos foram projetados para fornecer clareza e objetividade, visando reduzir o número de telas necessárias e aprimorar a usabilidade. O design foi desenvolvido de acordo com as



heurísticas de usabilidade para design de interface do usuário propostas por Nielsen (11) e as diretrizes do Material Design (12), com o propósito de oferecer uma experiência de usuário consistente. Ao todo, foram desenvolvidas 19 telas.

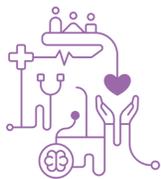
O fluxo principal da aplicação consiste em tirar uma foto com o sistema a qual seria então submetida a duas redes pré-treinadas: uma de classificação para identificar se a imagem enviada é ou não um raio X de mão, e outra de regressão para realizar a predição da idade óssea. A representação desse fluxo pode ser vista na Figura 2.

Figura 2 - Fluxo principal da aplicação.



Com os requisitos definidos e a aplicação idealizada, o projeto avançou para a fase de prototipação do MIDTS (10), na qual foram desenvolvidas as IAs do sistema. Para o treinamento destas, foram utilizados exclusivamente modelos de Redes Neurais Convolucionais (CNNs). Os modelos escolhidos para a tarefa de regressão, que visa a predição da idade óssea, incluíram *EfficientNetV2M*, *EfficientNetV2S*, *InceptionResNetV2*, *EfficientNetV2B3*, *ResNet152V2* e *InceptionV3*. Para a classificação, foi utilizado o *VGG16*, que em testes iniciais apresentou resultados promissores e foi escolhido de início. A escolha dos modelos foi limitada pelo poder de processamento do *hardware* utilizado para treino, mas também levou em consideração o desempenho superior dessas redes em tarefas de visão computacional, comprovado em diversas competições e *benchmarks* (13).

Os dados utilizados para o treinamento e validação das IAs foram obtidos através da Sociedade de Radiologia da América do Norte (RSNA). A RSNA reuniu um extenso conjunto de dados de imagens de raio X e promoveu um desafio aberto para que



cientistas desenvolvessem soluções de IA para o reconhecimento de raio X de mão e predição de idade óssea (14 pp498-503).

O conjunto de dados foi dividido em 10.719 imagens de raio X para treinamento, 1.892 imagens para validação e 1.425 imagens para teste. Às imagens, estava associada uma idade óssea, expressa em meses, bem como o gênero do indivíduo, totalizando 14.036 imagens no formato PNG.

A análise da distribuição de dados de gênero revela uma distribuição relativamente equilibrada, conforme ilustrado pelo gráfico apresentado na Figura 3-A. Esse equilíbrio é fundamental para garantir a representatividade e a validade dos resultados nas análises. No entanto, quando examinamos a distribuição etária das imagens, representada em meses, os dados não apresentam a mesma uniformidade. O gráfico da distribuição por faixa etária, ilustrado na Figura 3-B, evidencia uma concentração desigual em determinadas faixas etárias, principalmente entre 90 e 180 meses, o que pode introduzir vieses e limitar a generalização dos resultados.

Não foram empregadas técnicas específicas de manipulação de dados para regularizar a distribuição das amostras entre diferentes faixas etárias. No entanto, foram utilizadas técnicas de aumento de dados (*data augmentation*) em todas as faixas etárias durante alguns treinamentos, com o objetivo de aumentar a variabilidade dos dados e melhorar a robustez do modelo. As técnicas de aumento de dados aplicadas incluíram: translação aleatória (com fator de altura variando entre -0.2 e 0.2, e fator de largura variando entre -0.15 e 0.15), espelhamento horizontal aleatório, rotação aleatória (com fator de rotação de até 15%) e zoom aleatório (com fator de altura variando entre -0.2 e 0.1, e fator de largura variando entre -0.3 e 0.1). As técnicas foram escolhidas para simular condições de aquisição das imagens, visando maior generalização do modelo.

Figura 3 - Gráficos de distribuição dos dados por sexo (A) e faixa etária (B).



Gráfico A - Distribuição por sexo

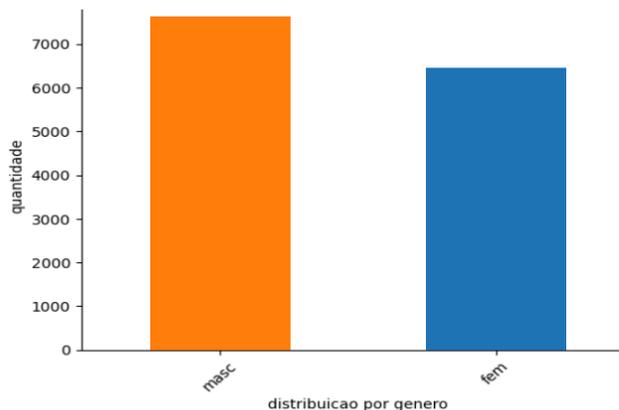
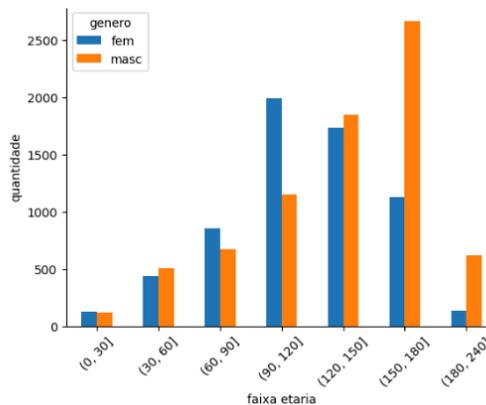


Gráfico B - Distribuição de faixa etária por gênero



Os treinamentos e testes das redes neurais foram realizados utilizando um computador com a seguinte configuração de *hardware*: processador AMD Ryzen 9 7950X com 16 núcleos e 32 threads, 64 GB de memória RAM DDR5 e uma GPU GeForce RTX 4080 com 16 GB de memória. A memória RAM foi utilizada como cache para acelerar o treinamento, o que ajudou a evitar sobrecarga no processador, especialmente ao aplicar *data augmentation* nas imagens. O conjunto de dados de treinamento foi constantemente alterado, seguindo caráter aleatório, para garantir a diversidade das amostras durante o processo de treinamento, o que contribuiu para a eficácia do modelo.

Para otimizar o tempo de treinamento do modelo regressivo, foi feito, a princípio, o treinamento das redes para o gênero masculino. A hipótese era que um modelo eficaz para o gênero masculino também seria eficaz para o feminino e para ambos os gêneros combinados. Esse método se mostrou eficiente, uma vez que o melhor modelo masculino, medido pelo erro absoluto médio (MAE), foi quase idêntico ao feminino, diferenciando-se apenas em dois hiperparâmetros: o bloco de atenção e a última camada densa. Para o conjunto geral, o melhor modelo masculino foi exatamente igual em todos os hiperparâmetros.

Várias configurações foram testadas, com e sem *data augmentation*. Apesar de ambos os métodos terem mostrado bons resultados, o melhor modelo foi obtido sem *data augmentation*. Isso indicou que o modelo era capaz de aprender características precisas a



partir dos dados brutos, sem a necessidade de aumentar artificialmente a quantidade de dados.

Em termos de arquitetura da rede neural, fez-se o uso de modelos pré-treinados disponíveis no Keras (13) com pesos da ImageNet1k (15), configurando o treinamento como refinamento do modelo para adaptá-lo à sua tarefa específica (*fine-tuning*). A última camada do modelo importado, padrão nos modelos de classificação, foi substituída por uma camada de *pooling*, sendo testadas as variações *GlobalMaxPooling2D* e *GlobalAveragePooling2D*, com a última apresentando melhor desempenho. Após essa camada, havia três camadas densas, que poderiam ter ou não *dropout* (0, 0,3 ou 0,5). Os blocos de atenção "*Squeeze and Excite*" e "CBAM" também foram implementados para avaliar seu impacto no desempenho. O número de neurônios seguiu um padrão de redução gradual: a primeira e a segunda camadas variaram entre 512 e 1024 neurônios, enquanto a última camada variou entre 128 e 256 neurônios. A escolha de potências de 2 serviu para fins de otimização computacional e padronização.

O tamanho do *batch* variou de 8 (*EfficientNetV2M*) a 100 (*InceptionV3*), limitados pela capacidade da GPU. Para melhorar o processo de treinamento, foi feito o uso de diversas funções do Keras (13), como *EarlyStopping* para interromper o treinamento quando a melhoria nas métricas de validação cessava, *ModelCheckpoint* para salvar o modelo sempre que sua performance melhorava, *CSVLogger* para registrar métricas de treinamento em um arquivo CSV para análises posteriores, e *ReduceLRonPlateau* para reduzir a taxa de aprendizado quando uma métrica específica parava de melhorar.

Para a fase de avaliação, várias sessões de treinamento foram realizadas e os melhores modelos de cada arquitetura foram salvos. As principais métricas utilizadas para a comparação e análise de desempenho das redes foram: a precisão e o *recall* para os modelos de classificação, e o coeficiente de determinação (R^2) e Erro Médio Absoluto (MAE) para os de regressão.

O artigo visa documentar os métodos e resultados obtidos a partir do processo de treinamento das redes neurais para o projeto, englobando a parte final do método.

Resultados e Discussão

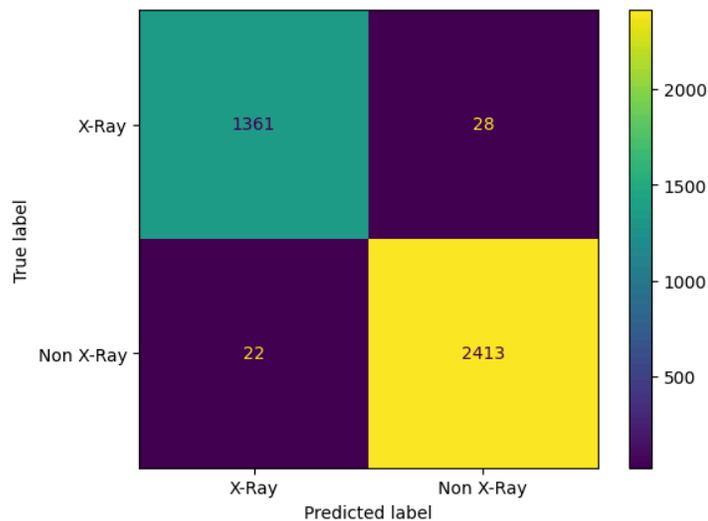


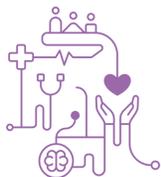
No que diz respeito à IA de classificação, o modelo *VGG16* apresentou resultados promissores. O modelo exibiu uma taxa de precisão de 97,9% em imagens de raios X da mão e 99% para as demais imagens, destacando sua capacidade de distinção entre as diferentes categorias.

A precisão, que mede a proporção de observações positivas corretamente previstas em relação ao total de positivos previstos, foi de 98,8%, indicando um alto nível de confiança na capacidade do modelo de identificar com precisão as imagens de raios X da mão. Enquanto isso, o *recall*, que avalia a proporção de positivos reais que foram corretamente identificados pelo modelo, refletiu essa excelência com 99%. Com um escore F1 de 98,9%, o modelo demonstra um equilíbrio entre precisão e *recall*, indicando confiabilidade tanto na identificação de verdadeiros positivos quanto na minimização de falsos positivos. A Figura 4 mostra a matriz de confusão para o modelo de classificação.

Para os modelos de regressão, a principal métrica de avaliação foi o coeficiente de determinação (R^2), seguido pelo MAE. Cada modelo apresentou resultados variados em termos de precisão e explicabilidade dos dados de resposta, de forma que nem todos chegaram na fase de avaliação.

Figura 4 - Matriz de confusão para *VGG16*.





O modelo *EfficientNetV2M* foi inicialmente testado nos dados masculinos e femininos, mas seus resultados não foram tão promissores quanto os de outros modelos avaliados. Além disso, o tempo de treinamento necessário foi significativamente maior em comparação com outras redes. Por esses motivos, a arquitetura foi abandonada para testes futuros.

O *EfficientNetV2B3* destacou-se como um dos melhores modelos, alcançando um MAE de 6,9 e um R^2 de quase 95%. Esses resultados indicam que o modelo não só é preciso na predição da idade óssea, mas também explica bem a variabilidade dos dados de resposta, tornando-se uma forte opção para aplicação prática.

O *InceptionResNetV2* também apresentou resultados satisfatórios, com valores de MAE próximos aos obtidos pelo *InceptionV3* e um R^2 elevado. Este modelo mostrou-se eficaz, equilibrando a precisão e a capacidade de explicação dos dados.

O *InceptionV3* obteve o melhor coeficiente de determinação (R^2), sugerindo que quase 95% da variação dos dados de resposta é explicada pelo modelo. Apesar de seu desempenho ligeiramente inferior em termos de MAE, a alta métrica torna-o uma opção viável para melhorias futuras.

Por outro lado, o *EfficientNetV2S* mostrou o melhor MAE, aproximadamente 6,5 em média, mas teve um R^2 ligeiramente inferior aos modelos *InceptionV3* e *InceptionResNetV2*. Este resultado sugere que, embora o modelo classifique bem, ele não explica tão bem as variações nos dados de resposta, necessitando de ajustes adicionais para melhorar sua interpretabilidade.

A Figura 5 a seguir apresenta as melhores métricas dos modelos de regressão para o sexo masculino após os testes

Figura 5 - Métricas dos modelos de regressão.



Modelo	Erro Médio Absoluto (em meses)	R ²
EfficientNetV2M	7,84	0,90613
EfficientNetV2B3	6,97	0,94807
InceptionResNetV2	6,92	0,94827
InceptionV3	7,02	0,94992
EfficientNetV2S	6,56	0,93942
ResNet152V2	7,43	0,93987

Com os resultados do treinamento para o sexo masculino, treinamentos similares foram feitos para o sexo masculino e de forma geral, sem diferenciação. As métricas de cada modelo seguiram padrões semelhantes e todos os dados podem ser observados nas Figuras 6, 7 e 8.

Figura 6 - Gráfico de resultados gerais da rede de regressão.

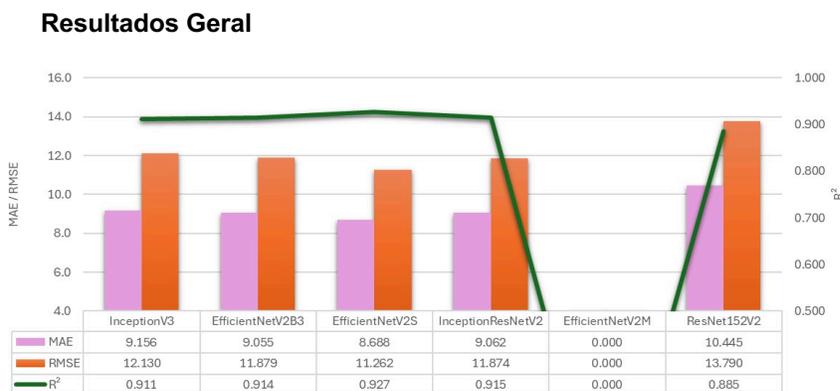


Figura 7 - Gráfico de resultados do sexo masculino da rede de regressão.



Resultados Masculino

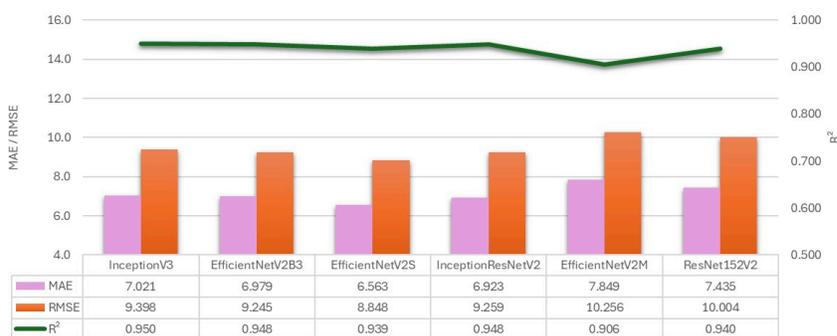
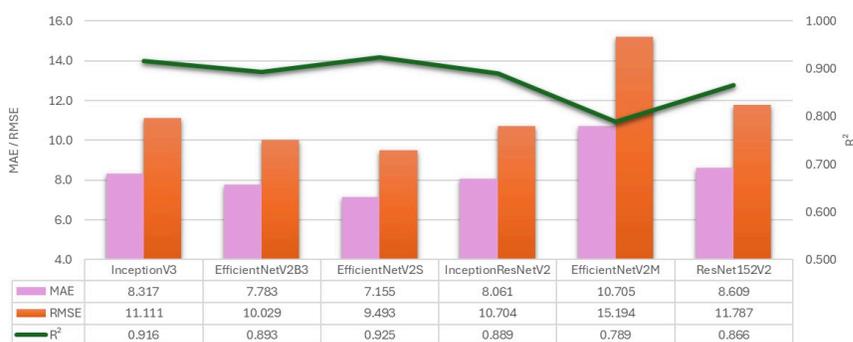


Figura 8 - Gráfico de resultados do sexo feminino da rede de regressão.

Resultados Feminino

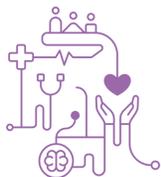


Conclusão

O projeto demonstrou a eficácia de utilizar modelos de IA para o reconhecimento de raios X de mão e a predição da idade óssea, com possibilidade de contribuição significativa para a prática médica pediátrica e ortopédica. Ao explorar as diferentes arquiteturas de redes neurais, conseguimos identificar os modelos mais adequados para essas tarefas específicas.

O modelo *VGG16* se destacou como promissor para a classificação de imagens de raio X. Graças à sua alta acurácia e precisão, ele foi considerado satisfatório para a tarefa e selecionado para o projeto *Doctor Bone*.

Entre os modelos de regressão testados, o *EfficientNetV2B3* destacou-se como a melhor opção, alcançando um MAE de 6,9 e um R² de quase 95%. Esses resultados



indicam que o modelo não só é preciso na predição da idade óssea, mas também explica bem a variabilidade dos dados de resposta. O *EfficientNetV2S* também mostrou um desempenho notável, com o melhor MAE de aproximadamente 6,5, embora seu R^2 fosse ligeiramente inferior, sugerindo que, embora classifique bem, necessita de melhorias para explicar melhor as variações nos dados.

O *InceptionV3* apresentou o melhor coeficiente de determinação (R^2), indicando uma alta capacidade de explicar a variabilidade nos dados de resposta, tornando-o uma opção viável apesar de seu desempenho ligeiramente inferior em termos de MAE. Esses resultados ressaltam a importância de equilibrar precisão e explicabilidade ao selecionar modelos de IA para aplicações clínicas.

Os resultados apontam a adoção dos modelos *EfficientNetV2B3* e *EfficientNetV2S* como as melhores opções para a predição da idade óssea. No entanto, outros modelos, como o *InceptionV3* e o *InceptionResNetV2*, também se mostraram promissores e podem ser explorados em futuros estudos com mais treinamentos e ajustes específicos.

Referências

1. Burns DAR, Silva LR, Júnior DC, Blank D, Vaz EDS, Borges WG. Tratado de pediatria. 4th ed. São Paulo: Manole Saúde; 2017.
2. Somkantha K, Theera-Umporn N, Auephanwiriyakul S, Williamson TH. Bone age assessment in young children using automatic carpal bone feature extraction and support vector regression. J Digit Imaging. 2011 Dec;24(6):1044-58. doi: 10.1007/s10278-011-9372-3.
3. Larson DB, Chen MC, Lungren MP, Halabi SS, Stence NV, Langlotz CP. Performance of a deep-learning neural network model in assessing skeletal maturity on pediatric hand radiographs. Radiology. 2018 Apr;287(1):313-22. doi: 10.1148/radiol.2017170236.
4. Rijn RR, Lequin MH, Thodberg HH. Bone age assessment: automated techniques coming of age? Acta Radiol. 2013 Nov;54(9):1024-9. doi: 10.1258/ar.2012.120443.
5. Savi FM, de Oliveira PT, Cestari TM, Granjeiro JM, Taga R. Histomorphometric evaluation of critical-sized bone defects using osteomeasure and aperio image analysis systems. Tissue Eng Part C Methods. 2019 Dec;25(12):732-741. doi: 10.1089/ten.tec.2019.0179.
6. Lee BD, Lee SA, Kim H, Cho J, Kim MS, Ko HK, et al. Automated bone age assessment using artificial intelligence: the future of bone age assessment. Korean J Radiol. 2021 May;22(5):792-800. doi: 10.3348/kjr.2020.0941.
7. Albanese A, Stanhope R, Fitzgerald F, Preece M. The use of a computerized method of bone age assessment in clinical practice. Horm Res. 1995;44(3):2-7. doi: 10.1159/000184665.



8. Maratová K, Chaloupková P, Šnajderová M, Krejčí H, Černá J, Novotná D, et al. A comprehensive validation study of the latest version of Bonexpert on a large cohort of Caucasian children and adolescents. *Front Endocrinol (Lausanne)*. 2023;14:1130580. doi: 10.3389/fendo.2023.1130580.
9. Ravi D, Wong C, Deligianni F, Berthelot M, Andreu-Perez J, Lo B, Yang GZ. Deep learning for health informatics. *IEEE J Biomed Health Inform*. 2017 Jan;21(1):4-21.
10. Filho JEV, Brasil CCP, Carneiro MC, Junior GS. MIDTS: método interdisciplinar para o desenvolvimento de tecnologias em saúde. In: Jorge MSB, Vergara CAC, Sampaio HADC, Moreira TMM, editors. *Tecnologias e-Health em Gestão em Saúde*. Curitiba: Editora CRV; 2021. p. 49-66.
11. Nielsen J. Ten usability heuristics [Internet]. [place unknown: publisher unknown]; 1994 Apr 24 [updated 2024 Jan 30; cited 2024 May 29]. Available from: <https://www.nngroup.com/articles/ten-usability-heuristics/>.
12. Material Design [Internet]. [place unknown: publisher unknown]; [date unknown] [cited 2024 May 29]. Available from: <https://m3.material.io/>.
13. Keras [Internet]. Keras 3 API Documentation. [place unknown: publisher unknown]; 2024 [cited 2024 May 24]. Available from: <https://keras.io/api/>.
14. Halabi SS, Prevedello LM, Kalpathy-Cramer J, Mamonov AB, Bilbily A, Cicero M, et al. The RSNA Pediatric Bone Age Machine Learning Challenge. *Radiology*. 2019 Feb;290(2):498-503.
15. ImageNet [Internet]. [place unknown: publisher unknown]; [date unknown] [cited 2024 May 28]. Available from: <https://www.image-net.org/>.