



The ecological footprint of medical AI

Daniel Truhn¹ · Gustav Müller-Franzes¹ · Jakob Nikolas Kather^{2,3,4}

Received: 30 June 2023 / Revised: 23 July 2023 / Accepted: 1 August 2023 / Published online: 15 August 2023
© The Author(s) 2023

Artificial intelligence (AI) is progressively being woven into the fabric of clinical medicine. Most notably, radiology, as an image-based medical specialty, offers numerous applications of AI, including image reconstruction, image analysis, and clinical decision-making. A rich body of research shows the potential of AI in radiology workflows, and indeed, many AI-based medical devices are already available and have received regulatory approval in major markets for clinical routine use. Existing AI models can support radiologists in a wide range of tasks, reaching from image acquisition [1] to diagnosis [2] and outcome prediction [3]. In the future, the development of foundation models, i.e., large multi-modal transformer neural networks that can manipulate text, images, and other data types, will conceivably increase the use of AI in radiology and medicine even more [4]. The advancements in large language models (LLMs), e.g., ChatGPT, are a testament to these technical developments, as these LLMs already have the capability to reason about clinically relevant topics across a wide range of domains.

Behind this effectiveness of contemporaneous AI models are two primary resources: access to large high-quality datasets and the ability to train huge neural networks with up to hundreds of billions of parameters. As these neural networks are trained and are being deployed at scale, a high computational demand is exerted, which consumes electricity and hence leaves a carbon footprint.

With the alarming verdict from the Intergovernmental Panel on Climate Change (IPCC-2021) suggesting that

global warming of 1.5 °C and 2 °C will be surpassed this century unless drastic CO₂ and other greenhouse gas emissions reductions are undertaken, we cannot overlook the energy and carbon costs involved in developing and using AI models in medicine. According to the United Nations Sustainable Development Goals Report 2022, taking urgent action to combat climate change is a top priority for our societies and the key instrument is to reduce and in the near future completely remove any carbon emissions.

How big a problem are medical AI models in this regard? To help comprehend the scale of this ecological impact, let us examine some relatable comparisons. On an annual basis, an average inhabitant of EU-27 contributes about 6.8 tons of CO₂ emissions, as per greenhouse gas emission statistics [5]. A round-trip flight between Munich and New York in economy class emits around 2.1 tons of CO₂ per person, according to Stiftung Myclimate [6]. And if we look at medical equipment, an MRI scanner running for 1 year results in around 58.3 tons of CO₂ emissions which amounts to 14.6 kg CO₂ per scan for a utilization of 4000 scans per year [7].

Consider, now, the carbon footprint of training a large language model like GPT-3, the predecessor of the first ChatGPT model. Training such an LLM leaves the same carbon footprint of roughly 262 persons taking a round trip flight from Munich to New York or running one MRI machine in a typical hospital environment for 9.5 years [8]. Comparatively, the carbon emission of conversing with a mid-sized language model in 20 interactions (i.e., giving 20 prompts) typically is equivalent to about a millionth of a one-person round trip transatlantic flight [8]. Of note, this number depends on multiple factors, such as the size of the language model, or the infrastructure used for deployment, but we estimate, that this number is correct within few orders of magnitude even for larger contemporary language models. Let us examine a more specific example from medical research, where smaller AI models are frequently used. Consider a state-of-the-art network undergoing training on a large database comprising hundreds of thousands of images for 100 hours. This equates roughly to the typical

✉ Daniel Truhn
dtruhn@ukaachen.de

¹ Department of Diagnostic and Interventional Radiology, University Hospital RWTH Aachen, Aachen, Germany

² Else Kroener Fresenius Center for Digital Health, Technical University Dresden, Dresden, Germany

³ Department of Medicine I, University Hospital Dresden, Dresden, Germany

⁴ Medical Oncology, National Center for Tumor Diseases (NCT), University Hospital Heidelberg, Heidelberg, Germany

daily carbon footprint of a European citizen [9]. This is modest. In the future, when AI models will increasingly be applied the carbon footprint of deploying the model will be more relevant. Let us assume that an AI model is run on the images of a patient that has just undergone an MR examination. In our experience, applying a model on a new examination takes a fraction of a second, but to account for bigger models in the future, we estimate that the model will occupy a state-of-the-art GPU in the cloud for at most 10 s. This amounts to a carbon footprint of 0.5 g [9]. Compare this to the carbon emission for the MRI scan itself which amounts to 14,600 g [7].

The magnitude of this carbon footprint seems modest, although, with AI models poised to play a central role in tomorrow's healthcare at a massive scale, these small environmental impacts can add up to relevant amounts and thus represent an opportunity for improvement.

We, as medical professionals, including radiologists, can have an influence by our purchase decisions and by proxy on the model providers. This can push the industry towards reducing the computational load for AI models. Models can be condensed to improve efficiency. The utilization of large, centralized data centers can offer up to twice the efficiency due to factors like optimized computational time usage and efficient cooling system [10]. Moreover, specialized hardware optimized for neural network inference could further reduce energy consumption. A transition towards clean energy sources for data centers is inevitable to remove carbon emissions. This is also economically reasonable, as wind and solar power are much cheaper than fossil fuel-based electricity.

Still, these actions represent changes at a small level, and often at an individual level. The magnitude of greenhouse gas emissions, the urgency of mitigating climate change, and the existential threat to our planet demand collective action. The healthcare sector, where we as medical doctors exert a considerable influence, is a relevant part of our society that could help to shift society on a sustainable path. An important area where such influence can be employed is investment decisions—be it on a personal level, or through indirect control and lobbying on a collective level. Large sums of money of medical professionals are for example tied up in pension funds. If we consciously decide to steer these investments away from fossil fuels and towards sustainable industries the positive impact on our planet could be substantial. This action, known as divestment, could result in reduced carbon emissions. In fact, divestment symbolizes the type of collective actions that can bring about tangible change, but there are others such as using the considerable trust society places in healthcare professionals and engaging collectively in initiatives such as the Physicians for Social Responsibility (<http://www.psr.org>) to spread awareness of the problem and potential solutions.

In conclusion, while AI's role in radiology and its associated carbon footprint are indeed relevant topics, the potential for more substantial impact goes beyond individual actions. AI in healthcare is poised to grow, but alongside its development, we must take conscious, collective steps towards sustainability at a larger scale. By incorporating these considerations into our professional and financial decisions, we can contribute to a more sustainable healthcare system and a liveable planet for future generations.

Acknowledgements In accordance with the COPE (Committee on Publication Ethics) position statement of 13 February 2023 (<https://publicationethics.org/cope-position-statements/ai-author>), the authors hereby disclose the use of the following artificial intelligence models during the writing of this article. GPT-4 (OpenAI) for checking spelling and grammar.

Funding Open Access funding enabled and organized by Projekt DEAL. The authors state that this work has not received any funding.

Declarations

Guarantor The scientific guarantor of this publication is Daniel Truhn.

Conflict of interest JNK declares consulting services for Owkin, France; DoMore Diagnostics, Norway and Panakeia, UK and has received honoraria for lectures by Eisai, Roche, MSD, BMS, Bayer, and Fresenius. No other potential conflicts of interest are reported by the authors.

Statistics and biometry No complex statistical methods were necessary for this paper.

Informed consent No Informed Consent was necessary.

Ethical approval Institutional Review Board approval was not required because no experiments on humans or animals were performed.

Study subjects or cohorts overlap Not applicable.

Methodology

- Commentary

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Kaniewska M, Deininger-Czermak E, Getzmann JM et al (2023) Application of deep learning-based image reconstruction in MR imaging of the shoulder joint to improve image quality and reduce scan time. *Eur Radiol* 33:1513–1525
2. Khader F, Han T, Müller-Franzes G et al (2023) Artificial intelligence for clinical interpretation of bedside chest radiographs. *Radiology* 307:e220510
3. Jiang X, Zhao H, Saldanha OL et al (2023) An MRI deep learning model predicts outcome in rectal cancer. *Radiology* 307:e222223
4. Moor M, Banerjee O, Abad ZSH et al (2023) Foundation models for generalist medical artificial intelligence. *Nature* 616:259–265
5. Greenhouse gas emission statistics - carbon footprints. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Greenhouse_gas_emission_statistics_-_carbon_footprints. Accessed 27 Jun 2023
6. Stiftung myclimate Calculate your CO2 emissions now! https://co2.myclimate.org/en/portfolios?calculation_id=5923442. Accessed 27 Jun 2023
7. Woolen SA, Becker AE, Martin AJ et al (2023) Ecodesign and operational strategies to reduce the carbon footprint of MRI for energy cost savings. *Radiology* 307:e230441
8. Luccioni AS, Viguier S, Ligozat A-L (2022) Estimating the carbon footprint of BLOOM, a 176B parameter language model. <https://arxiv.org/abs/2211.02001>
9. Lacoste A, Luccioni A, Schmidt V, Dandres T (2019) Quantifying the carbon emissions of machine learning. <https://arxiv.org/abs/1910.09700>
10. Patterson D, Gonzalez J, Le Q et al (2021) Carbon emissions and large neural network training. <https://arxiv.org/abs/2104.10350>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.