

Infraestrutura Computacional para Apoio à Avaliação Fecal com Inteligência Artificial

Lucas Chasseraux Tauil¹, José Remo Ferreira Brega¹

¹ Programa de Pós-Graduação em Ciência da Computação – Faculdade de Ciências
Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP)
Bauru – SP – Brasil

lucas.tauil@unesp.br, remo.brega@unesp.br

Resumo. Este artigo apresenta uma infraestrutura computacional para apoio à avaliação fecal com Inteligência Artificial (IA). O trabalho parte de uma revisão sistemática da literatura (RSL), conduzida em seis bases científicas, que identificou 62 registros e incluiu 15 estudos após triagem. Os achados indicaram predominância de redes neurais convolucionais em tarefas com imagens, uso recorrente da Escala de Bristol ou de variações pediátricas, baixa disponibilidade de bases públicas e limitações de replicabilidade. A partir dessas lacunas, foi especificado um protótipo modular em três camadas, contemplando gestão web, treinamento de modelos e serviço de IA. Experimentos preliminares com bases substitutas indicaram viabilidade técnica do pipeline, com acurácia de 0,8652 no ensaio com imagens de fezes humanas. Os resultados não representam validação clínica, mas evidenciam uma base computacional inicial para estudos futuros com rastreabilidade, versionamento e auditoria.

Abstract. This paper presents a computational infrastructure to support stool assessment with Artificial Intelligence (AI). The work starts from a systematic literature review (SLR), conducted in six scientific databases, which retrieved 62 records and included 15 studies after screening. The findings indicated the predominance of convolutional neural networks in image-based tasks, the recurrent use of the Bristol scale or pediatric variations, limited availability of public datasets, and replicability constraints. Based on these gaps, a three-layer modular prototype was specified, comprising web management, model training, and an AI service. Preliminary experiments with substitute datasets indicated the technical feasibility of the pipeline, reaching 0.8652 accuracy in the experiment with human stool images. The results do not represent clinical validation, but provide an initial computational basis for future studies with traceability, versioning, and auditing.

1. Introdução

A avaliação da consistência fecal é um método clínico não invasivo e de baixo custo para acompanhamento da saúde intestinal. A Escala de Fezes de Bristol é uma das referências mais empregadas para padronizar essa avaliação, mas sua aplicação ainda depende da observação humana e pode sofrer variações entre avaliadores [Lewis and Heaton 1997, Blake et al. 2016]. Em aplicações baseadas em imagens, essa limitação cria oportunidade para soluções computacionais capazes de apoiar a classificação fecal de forma mais padronizada e rastreável.

Estudos recentes têm explorado IA, especialmente redes neurais convolucionais, para análise de imagens fecais e de dados relacionados à saúde intestinal [Choy et al. 2021b, Yu et al. 2022, Zhou et al. 2023, Zhu et al. 2024]. Apesar dos resultados promissores desses classificadores, a RSL indicou que a consolidação de soluções aplicáveis depende também de fatores estruturais, como disponibilidade de bases públicas, replicabilidade, restrições éticas de compartilhamento de imagens sensíveis e explicitação do fluxo técnico que conecta coleta de dados, treinamento, publicação de modelos e inferência.

Este artigo combina uma revisão sistemática da literatura (RSL) com o desenvolvimento de um protótipo modular para apoio à avaliação fecal com IA. O objetivo não é apresentar validação clínica definitiva, mas mostrar como os achados da RSL foram traduzidos em requisitos computacionais e materializados em uma infraestrutura para gestão de dados, treinamento, publicação de modelos e inferência.

2. RSL como insumo para requisitos

A RSL seguiu um protocolo inspirado em diretrizes para revisões secundárias [Kitchenham and Charters 2007]. Foram consultadas IEEE Xplore, ACM Digital Library, Scopus, PubMed, Bireme e Lilacs, em março de 2025, nos metadados dos artigos. A string-base foi: ("Artificial Intelligence"OR "Machine Learning"OR "Deep Learning") AND ("feces"OR "stool"OR "bowel movement") AND ("Bristol Stool Scale"OR "Stool Consistency"OR "Stool Form Scale").

Foram incluídos estudos a partir de 2020, com texto completo disponível, relacionados ao uso de IA, aprendizado de máquina ou aprendizado profundo na análise fecal, com uso da Escala de Bristol, variações pediátricas ou critérios equivalentes, e descrição metodológica suficiente. Foram excluídos trabalhos duplicados, anteriores a 2020, sem acesso integral, sem relação adequada com IA aplicada à avaliação fecal, ou resumos curtos e pôsteres sem detalhamento metodológico. O processo resultou em 62 registros recuperados, 25 duplicados removidos, 37 avaliados por título e resumo, 26 lidos integralmente e 15 estudos incluídos.

Os estudos incluídos contemplaram abordagens com imagens, dados estruturados, escalas BSFS/BITSS e diferentes níveis de replicabilidade [Chang et al. 2021, Choy et al. 2021b, Yang et al. 2022, Yu et al. 2022, Huysentruyt et al. 2023, Dawkins et al. 2022, Nakazawa et al. 2023, Choy et al. 2021a, Zhu et al. 2024, Van Den Berg et al. 2024, Jung et al. 2024, Zhu et al. 2023, Zhou et al. 2024, Zhou et al. 2023, Fujimoto et al. 2023]. Os achados indicaram três tendências: predominância de CNNs e *transfer learning* em estudos com imagens; uso recorrente da Escala de Bristol ou da BITSS; e replicabilidade limitada pela escassez de bases públicas, indisponibilidade de código e restrições éticas de compartilhamento de imagens fecais.

A partir desses achados, foram definidos quatro requisitos centrais para a solução: organizar projetos, classes e imagens de forma rastreável; separar treinamento e inferência; publicar modelos com versionamento e metadados; e manter histórico das previsões para auditoria técnica.

3. Arquitetura e *pipeline* da solução

Para atender aos requisitos derivados da RSL, a solução foi estruturada em três camadas, conforme a Figura 1. A camada de gestão web concentra o cadastro de projetos, classes e imagens, controle de acesso, exportação de conjuntos de dados e associação entre projeto e modelo ativo. A camada de treinamento recebe o conjunto exportado, executa treino, validação e teste, e gera os artefatos de publicação. A camada de serviço de IA carrega os modelos publicados e disponibiliza *endpoints* de inferência.



Figura 1. Arquitetura modular da solução proposta, com separação entre gestão web, treinamento e serviço de IA.

A escolha por CNNs e *transfer learning* decorre dos achados da RSL. CNNs são adequadas para extração hierárquica de padrões visuais, enquanto o *transfer learning* reduz a dependência de grandes bases rotuladas, aspecto relevante em um domínio com baixa disponibilidade de dados públicos e restrições éticas para compartilhamento de imagens fecais. No protótipo, o treinamento foi implementado em Python com TensorFlow/Keras, utilizando ResNet50V2, aumento de dados, *fine-tuning* e geração automática de métricas e artefatos.

Na gestão web, as imagens são persistidas no servidor e associadas a registros no banco de dados. São armazenados dados sobre projetos, classes, imagens, usuários, versões de modelos e histórico de inferências. Para cada imagem, mantêm-se metadados como projeto, classe, caminho lógico, extensão, tipo MIME, tamanho, dimensões, status e usuário responsável. As predições registram projeto, imagem submetida, rótulo previsto, confiança, versão do modelo e resposta retornada pelo serviço, permitindo auditoria técnica posterior.

O fluxo técnico é composto por cinco etapas: organização das imagens por projeto e classe, exportação do conjunto de treinamento, execução do treino/validação/teste, publicação versionada dos artefatos do modelo e inferência via *endpoint*. Essa sequência traduz os requisitos identificados na RSL em responsabilidades de software.

4. Resultados preliminares e análise

Os experimentos preliminares tiveram como objetivo verificar a viabilidade computacional do *pipeline*, e não validar clinicamente a solução final. Como a base oficial ainda se encontra em consolidação, foram utilizados dois conjuntos substitutos. O ensaio mais aderente ao domínio-alvo utilizou 870 imagens de fezes humanas, organizadas em três classes, e atingiu acurácia de teste de 0,8652. Adicionalmente, um conjunto de maturação de bananas, com 5347 imagens e seis classes, foi empregado apenas como teste técnico auxiliar em cenário multiclasse com variação de cor, textura e aparência, atingindo acurácia de 0,9753.

Os resultados devem ser interpretados como evidências iniciais de viabilidade computacional. O protótipo executou o ciclo completo previsto: organização e exportação dos dados, identificação das classes, treinamento, avaliação em teste, geração de artefatos, publicação do modelo e disponibilização para inferência. Assim, a contribuição principal não está na obtenção isolada de valores de acurácia, mas na demonstração de que o fluxo técnico proposto é executável e coerente com os requisitos de rastreabilidade, desacoplamento e versionamento.

O resultado com imagens de fezes humanas é o mais relevante para a pesquisa, mas deve ser analisado com cautela, pois o conjunto utilizado é substituto e simplifica a classificação em três grupos amplos. Dessa forma, os resultados não permitem conclusão clínica final, mas indicam uma base técnica viável para a continuidade dos experimentos com a base oficial.

5. Conclusão

Este artigo apresentou uma infraestrutura computacional para apoio à avaliação fecal com IA, concebida a partir de lacunas identificadas em uma RSL e materializada em um protótipo modular. Os achados da literatura motivaram uma arquitetura capaz de organizar dados, separar treinamento e inferência, publicar modelos versionados e registrar previsões para auditoria técnica. Os resultados preliminares indicam que o *pipeline* implementado é tecnicamente viável e executa o ciclo completo de organização de dados, treinamento, publicação de modelo e inferência. A principal limitação atual é a ausência de validação final com a base oficial de imagens e com as classes de interesse definidas para o domínio-alvo. Como próximos passos, pretende-se consolidar essa base, executar novos experimentos e aprofundar a avaliação dos modelos no contexto de aplicação.

Referências

- Blake, M. R., Raker, J. M., and Whelan, K. (2016). Use of the Bristol Stool Form Scale in clinical and research settings: A systematic review. *The American Journal of Gastroenterology*, 111(1):75–83.
- Chang, J. et al. (2021). Towards an ambient estimation of stool types to support nutrition counseling for people affected by the geriatric frailty syndrome. *Journal of Ambient Intelligence and Humanized Computing*, 12:11315–11330.
- Choy, C. A. et al. (2021a). Machine learning supports automated digital image scoring of stool consistency in diapers. *Journal of Pediatric Gastroenterology and Nutrition*, 73(2):147–153.

- Choy, Y. P., Hu, G., and Chen, J. (2021b). Detection and classification of human stool using deep convolutional neural networks. *IEEE Access*, 9:160485–160496.
- Dawkins, J. B. et al. (2022). Gut metabolites predict *Clostridioides difficile* recurrence. *Microbiome*, 10(1):1–14.
- Fujimoto, T. et al. (2023). Effect of daily consumption of a fermented milk containing *Lactobacillus paracasei* strain Shirota (LcS) on stool consistency in united states adults with hard or lumpy stools: A randomized controlled trial. *The Journal of Nutrition*, 153(6):1836–1845.
- Huysentruyt, K. et al. (2023). Technician-scored stool consistency spans the full range of the Bristol Scale in a healthy US population and differs by diet and chronic stress load. *The Journal of Nutrition*, 153(4):1199–1207.
- Jung, J. et al. (2024). Artificial intelligence- and physician-interpreted stool image characteristics correlate with C-Reactive Protein among inpatients with acute severe ulcerative colitis: A pilot study. *Inflammatory Bowel Diseases*, 30(3):397–406.
- Kitchenham, B. and Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. Technical Report EBSE-2007-01, EBSE Technical Report.
- Lewis, S. J. and Heaton, K. W. (1997). Stool form scale as a useful guide to intestinal transit time. *Scandinavian Journal of Gastroenterology*, 32(9):920–924.
- Nakazawa, H. et al. (2023). A mountable toilet system for personalized health monitoring via the analysis of excreta. *Nature Biomedical Engineering*, 7:541–553.
- Van Den Berg, M. M. et al. (2024). The importance of objective stool classification in fecal ¹H-NMR metabolomics: Exponential increase in stool crosslinking is mirrored in systemic inflammation and associated to fecal acetate and methionine. *Metabolomics*, 20(1).
- Yang, X. et al. (2022). Stool image analysis for digital health monitoring by smart toilets. *IEEE Transactions on Biomedical Engineering*, 69(5):1655–1666.
- Yu, X. et al. (2022). Human stools classification for gastrointestinal health based on an improved ResNet18 model with dual attention mechanism. *Computers in Biology and Medicine*, 145:105440.
- Zhou, H. et al. (2023). Generation and application of a convolutional neural network algorithm in evaluating stool consistency in diapers. *Acta Paediatrica*, 112(3):682–689.
- Zhou, Y. et al. (2024). Long-term, automated stool monitoring using a novel smart toilet: A feasibility study. *Neurogastroenterology & Motility*, 36(2):e14764.
- Zhu, Q. et al. (2023). Integrative analysis of $\gamma\delta$ T cells and dietary factors reveals predictive values for autism spectrum disorder in children. *Brain, Behavior, and Immunity*, 111:148–159.
- Zhu, Y. et al. (2024). The effectiveness of artificial intelligence in assisting mothers with assessing infant stool consistency in a breastfeeding cohort study in china. *Nutrients*, 16(2):200.