

Detecção de Alérgenos em Imagens de Alimentos com YOLOv8

Rodrigo Eduardo Seger¹, Luciano Costa Blomberg¹, Ana Trindade Winck¹

¹Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA)
Porto Alegre – RS – Brazil

rodrigoes@ufcspa.edu.br, lucianocb@ufcspa.edu.br, anatw@ufcspa.edu.br

Abstract. *Food allergies are adverse reactions to the ingestion of certain foods. The correct identification of allergens presents challenges, especially when proper labeling is not available. This study evaluates four variants of the YOLOv8 deep learning architecture for allergen detection in images of prepared foods. A total of 9,984 images from public datasets were combined, with the best results obtained using the YOLOv8m architecture, achieving precision = 0.76 and recall = 0.65. Although modest, the results indicate a promising use of the architecture for allergen identification.*

Resumo. *Alergias alimentares são reações adversas à ingestão de alimentos. A correta identificação de alérgenos traz desafios, principalmente quando não é possível contar com a rotulagem adequada. Este trabalho avalia quatro variantes da arquitetura de deep learning YOLOv8 para detecção de alérgenos em imagens de alimentos preparados. Foram combinadas 9.984 imagens extraídas de datasets públicos, obtendo-se os melhores resultados com a arquitetura YOLOv8m, com precisão=0,76 e revocação=0,65. Os resultados, embora modestos, indicam um uso promissor da arquitetura para identificação de alérgenos.*

1. Introdução

Alergias alimentares são reações imunológicas adversas desencadeadas por proteínas específicas presentes em alimentos. São reconhecidas como uma questão de saúde pública pois apresentam um desafio para a alimentação das pessoas alérgicas, afetando cerca de 2% da população adulta com hipersensibilidade alimentar, enquanto quase 1% enfrenta alergias alimentares diagnosticadas [Solé et al. 2018]. Atualmente, o tratamento consiste na exclusão do alimento envolvido. No Brasil, as Resoluções da Diretoria Colegiada (RDC) da ANVISA nº 26/2015 [Brasil 2015] e nº 727/2022 [Brasil 2022] destacam um conjunto de 18 grupos de alimentos com maior prevalência para alergias. Dentre esses, em [Pereira et al. 2024] são destacados leite, ovo, trigo e soja como os 4 principais.

Existem trabalhos que se dedicam a apontar a presença de alérgenos em ingredientes de alimentos embalados, através da inspeção dos rótulos, como descrito em [Menezes et al. 2024, Menin et al. 2024, Gadelha et al. 2024], buscando auxiliar a população alérgica no consumo de alimentos industrializados. No entanto, quando se trata de alimentos já preparados, a população alérgica pode desconhecer os ingredientes mais utilizados em sua preparação, especialmente quando são consumidos fora de casa. Nesse contexto, o desenvolvimento de um modelo de Deep Learning que seja capaz de

identificar potenciais alérgenos em pratos de alimentos através de imagens é uma abordagem promissora, pois pode auxiliar na escolha alimentar dos pacientes alérgicos, ao fornecer informações sobre possíveis alérgenos presentes em alimentos preparados.

Em [Mishra et al. 2022], é apresentado o Allergen30, uma solução de Deep Learning que busca detectar alérgenos em imagens de alimentos. Os autores fornecem um dataset com cerca de 5.000 imagens distribuídas em aproximadamente 30 classes de alimentos. Essas imagens são em sua maioria bem tratadas, mas não retratam fotos reais de um prato a ser consumido. Ainda, em [Bossard et al. 2014] é proposto o Food-101, um dataset para classificação de imagens de alimentos preparados em relação ao tipo de alimento, mas não focados em alergia. Os autores disponibilizam um dataset com 100.000 imagens de alimentos distribuídas em 101 categorias. Embora não relacionadas a alérgenos, as imagens estão mais próximas do que seria um prato real ser consumido.

Considerado as características das imagens de cada dataset, neste trabalho é proposta a construção de um novo dataset composto por 10.000 imagens, selecionadas em proporções iguais a partir de [Mishra et al. 2022] e [Bossard et al. 2014]. As imagens foram anotadas e organizadas no Roboflow, contemplando uma diversidade de pratos contendo os dois principais alérgenos: leite e ovo. A partir desse dataset, foram treinados modelos de Deep Learning baseados na arquitetura YOLOv8, aplicada à detecção automática de alérgenos em imagens de alimentos preparados.

2. Metodologia

A combinação do Allergen30 [Mishra et al. 2022] com o Food-101 [Bossard et al. 2014] foi motivada pelas limitações do primeiro que, embora apresente anotações consistentes, suas imagens são pouco realistas, com baixa presença de ruídos ou distorções típicas de cenários reais. Por outro lado, o Food-101 reúne imagens em contextos mais variados, incluindo pratos prontos e ingredientes fotografados sob diferentes condições visuais. Para a construção do novo dataset, foram selecionadas 10.000 imagens, considerando rótulos que continham leite, ovo ou outros, com a distribuição apresentada na Tabela 1.

Tabela 1. Quantidade de imagens por classe no dataset

Classe	Quantidade
Leite	3.600
Ovos	3.600
Leite e Ovos	800
Outros	2.000
Total	10.000

As imagens do novo dataset foram anotadas com auxílio da ferramenta Roboflow, com a criação de caixas delimitadoras ao redor dos objetos de interesse, facilitando a identificação dos alérgenos. As novas classes foram definidas com base na presença dos alérgenos leite e ovo na receita original do prato ou da refeição retratada em cada imagem.

Após a coleta e anotação, as imagens foram divididas no formato 70-20-10 para treinamento, validação e teste. Foram avaliadas as arquiteturas Nano, Small, Medium e Large do YOLOv8, onde cada uma está relacionada à quantidade de parâmetros, que dizem respeito os pesos e vieses aplicados durante a convolução. Os treinamentos iniciais

foram realizados por 100 épocas, com imagens redimensionadas para 640x640 pixels e batch ajustado automaticamente, usando o otimizador SGD e taxa de aprendizado inicial de 0,01 com decaimento progressivo. Cada arquitetura do modelo Yolo foi avaliada com base nas medidas de Precisão e Revocação, além das medidas de mAP50 e mAP50-95, apresentadas pela arquitetura. Essas duas últimas medidas computam a precisão média da detecção de objetos (mean average precision), em que a mAP50 permite um limiar de 0,5 em relação à caixa de detecção do objeto e a mAP50-95 um limiar de 0,5 a 0,95.

3. Resultados

Das 10.000 imagens selecionadas, 16 foram identificadas como ruidosas e retiradas do dataset. A Tabela 2 mostra a distribuição das imagens, considerando combinações de classes. Há predominância de imagens contendo apenas um alérgeno, enquanto combinações de Leite e Ovos ou envolvendo múltiplos rótulos são menos frequentes.

Tabela 2. Distribuição das imagens por classes no dataset anotado

Classe / Combinação de Classes	Quantidade de imagens
Leite	3.659
Ovos	2.706
Outros	2.776
Leite e Ovos	480
Leite e Outros	124
Outros e Ovos	202
Leite, Outros e Ovos	37
Total	9.984

O Desempenho dos Modelos YOLOv8 Treinados podem ser observados na Tabela 3, que apresenta, para cada modelo, os parâmetros (em milhões), as medidas mAP50, mAP50-95, precisão e revocação. Embora a variação de qualidade seja discreta, as arquiteturas Medium e Large apresentam um melhor desempenho em relação à Nano e Small.

Tabela 3. Desempenho dos Modelos YOLOv8 Treinados

Modelo	Parâmetros	mAP50	mAP50-95	Precisão	Revocação
YOLOv8n (Nano)	3,0	0,71	0,56	0,73	0,60
YOLOv8s (Small)	11,1	0,70	0,58	0,72	0,64
YOLOv8m (Medium)	25,8	0,74	0,62	0,76	0,65
YOLOv8l (Large)	43,6	0,73	0,62	0,76	0,64

4. Conclusão

Buscando uma solução voltada a minimização do risco de ingestão acidental de alérgenos de leite e ovos em alimentos, este trabalho propôs uma solução que combinou de imagens de datasets públicos, extraídas e pré-processadas para indução de modelos baseados em deep learning e que visaram a detecção dos referidos alérgenos. Os experimentos realizados, por meio da arquitetura YOLOv8, apontam viabilidade da aplicação de deep learning na detecção de alérgenos em imagens de alimentos. Ainda que com resultados modestos,

as métricas indicam que os modelos conseguem identificar, em diferentes pratos, regiões associadas à presença de leite e/ou ovo em suas receitas.

Em relação às limitações do estudo, destaca-se o número reduzido de alérgenos e arquiteturas de deep learning testadas. Essas escolhas foram influenciadas, principalmente, por restrições computacionais, de modo que, neste momento, priorizou-se a análise de um grupo restrito, composto pelos alérgenos mais prevalentes, conforme verificado no trabalho de [Pereira et al. 2024]. No entanto, para trabalhos futuros, considera-se a utilização de um conjunto mais amplo de alérgenos, bem como a experimentação com outras arquiteturas de deep learning, o que poderá proporcionar uma avaliação mais abrangente e robusta do desempenho preditivo. Por fim, a despeito das limitações mencionadas, entendemos que este estudo é um importante ponto de partida para a disponibilização dos modelos em aplicações voltadas a usuários finais, auxiliando em escolhas alimentares mais seguras e conscientes.

Referências

- Bossard, L., Guillaumin, M., and Van Gool, L. (2014). Food-101 – mining discriminative components with random forests. In Fleet, D., Pajdla, T., Schiele, B., and Tuytelaars, T., editors, *Computer Vision – ECCV 2014*, volume 8694 of *Lecture Notes in Computer Science*, pages 446–461, Cham, Switzerland. Springer.
- Brasil (2015). Resolução da diretoria colegiada nº 26, de 02 de julho de 2015. Diário Oficial da União.
- Brasil (2022). Resolução da diretoria colegiada nº 727, de 15 de julho de 2022. Diário Oficial da União.
- Gadelha, G. A., Pereira, R. A., Guedes, F. M., and Winck, A. T. (2024). Processamento de linguagem natural na identificação de alérgenos em rótulos alimentares: Uma aplicação no contexto brasileiro. *Journal of Health Informatics*, 16(Especial).
- Menezes, G. P., Pereira, R. A., Guedes, F. M., and Winck, A. T. (2024). Desenvolvimento de um aplicativo móvel para leitura e identificação de alérgenos em rótulos alimentares. In *Anais do XX Congresso Brasileiro de Informática em Saúde*, pages 91–95, São Paulo. Sociedade Brasileira de Informática em Saúde.
- Menin, G., Pereira, R. A., Guedes, F. M., and Winck, A. T. (2024). Api rest para reconhecimento óptico de caracteres em rótulos alimentares. *Journal of Health Informatics*, 16(Especial).
- Mishra, M. et al. (2022). Allergen30: Detecting food items with possible allergens using deep learning-based computer vision. *Food Analytical Methods*, 15:3045–3078.
- Pereira, R. A., Guedes, F. M., Gadelha, G. A., Menin, G., Menezes, G. P., Amantéa, S. L., and Winck, A. T. (2024). Alérgenos alimentares em rótulos de produtos alimentícios: uma unificação da nomenclatura. *Arquivos de Asma, Alergia e Imunologia*, 8(2):125–128.
- Solé, D., Silva, et al. (2018). Consenso brasileiro sobre alergia alimentar: 2018 - parte 2 - diagnóstico, tratamento e prevenção. documento conjunto elaborado pela sociedade brasileira de pediatria e associação brasileira de alergia e imunologia. *Arquivos de Asma, Alergia e Imunologia*, 2(1):39–82.