

Detecção de Blisters Farmacêuticos com Redes Neurais para Automação da Distribuição por Dose Unitária

Luana Bortko Rodrigues¹, Oscar E. Fugita², Leila Cristina C. Bergamasco^{1,3}

¹ Departamento de Ciência da Computação
Centro Universitário FEI – São Bernardo do Campo, SP – Brasil

²Hospital Universitário
Universidade de São Paulo– São Paulo, SP – Brasil

³Programa de Pós-Graduação em Engenharia Elétrica
Centro Universitário FEI – São Bernardo do Campo, SP – Brasil

{uniflurodrigues, leila.cristina}@fei.edu.br; oscareh@hu.usp.br

Abstract. *This study investigates the performance of the YOLOv8 and YOLOv11 architectures in detecting pharmaceutical tablets in blister packs, evaluating the impact of different datasets and analyzing how the origin and structural diversity of the images influence model performance. Experiments were conducted using datasets obtained from the internet, from the University Hospital of the University of São Paulo (HU-USP), and from a combined dataset. The results indicate similar performance between the evaluated architectures and suggest that dataset diversity may contribute to improving the models' generalization capability. These findings provide support for the development of automated solutions for the medication unit-dose repackaging process in hospital environments.*

Resumo. *Este estudo investiga o desempenho das arquiteturas YOLOv8 e YOLOv11 na detecção de comprimidos em blisters farmacêuticos, avaliando o impacto de diferentes conjuntos de dados e analisando como a origem e a diversidade estrutural das imagens influenciam o desempenho dos modelos. Para isso, foram realizados experimentos utilizando conjuntos de dados provenientes da internet, do Hospital Universitário da Universidade de São Paulo (HU-USP) e um conjunto combinado. Os resultados indicam desempenho semelhante entre as arquiteturas avaliadas e sugerem que a diversidade dos dados pode contribuir para ampliar a capacidade de generalização dos modelos. Esses resultados fornecem subsídios para o desenvolvimento de soluções de automação no processo de fracionamento de medicamentos em ambiente hospitalar.*

1. Introdução

O Hospital Universitário da Universidade de São Paulo (HU-USP) é um hospital-escola localizado na Cidade Universitária, responsável pela dispensação anual de milhares de medicamentos, evidenciando a relevância do setor de farmácia hospitalar na gestão assistencial [Universidade de São Paulo 2023].

O sistema de distribuição de medicamentos exerce papel estratégico ao impactar o controle de estoque, os custos institucionais e a segurança do paciente. Entre os modelos

existentes, destaca-se a distribuição por dose unitária, considerada uma evolução operacional por organizar doses individualizadas prontas para administração, contribuindo para maior controle da farmacoterapia [Andrade and Freitas 2019].

Entretanto, sua implementação demanda etapas adicionais, tendo como principal desafio o fracionamento físico dos comprimidos, processo que requer elevada precisão e tempo operacional. Apesar da relevância do tema, poucos estudos abordam a automação dessa etapa. Nesse contexto, [Tani et al. 2024] propuseram um método automatizado para cálculo de trajetórias de corte em *blisters* farmacêuticos, utilizando a rede YOLOv8 para detecção dos comprimidos e Diagramas de Voronoi para definição das regiões de corte, alcançando *mean Average Precision* (mAP) de 99%. Contudo, a validação foi conduzida em condições controladas, permanecendo em aberto a avaliação da capacidade de generalização da arquitetura YOLO em cenários reais.

Dessa forma, o presente artigo investiga o desempenho da arquitetura YOLOv11 na detecção de comprimidos em *blisters*, avaliando diferentes conjuntos de dados, incluindo imagens provenientes da internet e imagens coletadas no HU-USP. O objetivo é analisar como a diversidade estrutural e a origem das imagens influenciam o desempenho e a capacidade de generalização do modelo, complementando limitações apontadas na literatura recente.

2. Trabalhos Relacionados

As arquiteturas da família YOLO (*You Only Look Once*) têm sido amplamente empregadas em tarefas de detecção de objetos na área médica, com aplicações que incluem detecção de lesões cutâneas, identificação de nódulos pulmonares e diagnóstico em exames radiológicos [Ragab et al. 2024, Soni and Rai 2024]. Além das aplicações diagnósticas, estudos também demonstram o uso de YOLO para detecção de objetos clínicos em ambientes de procedimento, como a identificação de instrumentos cirúrgicos [Jiang et al. 2023]. Apesar dos resultados promissores, a literatura aponta desafios recorrentes, como a necessidade de conjuntos de dados amplos e balanceados, dificuldades na detecção de objetos pequenos, sensibilidade a variações de iluminação e limitações de generalização quando aplicados a domínios distintos [Ragab et al. 2024, Soni and Rai 2024].

No cenário de fracionamento de *blisters* farmacêuticos, tais limitações tornam-se particularmente relevantes, uma vez que a elevada diversidade de formatos, disposições internas e características estruturais das embalagens pode impactar diretamente o desempenho dos modelos de detecção. Nesse contexto, destaca-se o trabalho de Tani et al. [Tani et al. 2024], que aplicou a YOLOv8 associada ao Diagrama de Voronoi para estimar regiões de corte entre cápsulas, alcançando um mAP de 99%. Contudo, os próprios autores apontam a necessidade de novos testes para avaliar a capacidade de generalização do modelo frente à diversidade de tipos de *blisters*, uma vez que falhas foram observadas em parte dos tipos de *blisters* analisados, como em casos de cápsulas posicionadas diagonalmente nas embalagens. Nesse sentido, o presente estudo investiga o impacto da origem e da variabilidade estrutural dos conjuntos de dados no desempenho de modelos YOLO aplicados à detecção de comprimidos em *blisters*.

3. Materiais e Métodos

3.1. Conjuntos de Dados

Foram utilizados três conjuntos de dados para o treinamento e avaliação dos modelos, todos configurados para a detecção de uma única classe correspondente a comprimidos (*pills*). O primeiro conjunto foi formado a partir da combinação de duas base de dados públicas disponíveis na plataforma Roboflow [Roboflow 2024, Yoloultra 2025]. O conjunto final contém 4364 imagens e 74974 instâncias anotadas, apresentando maior volume de dados e ampla variação de poses.

O segundo conjunto corresponde à base de dados HU-USP, composto por imagens capturadas no HU-USP. Esse conjunto contém 342 imagens e 4342 instâncias anotadas de diferentes medicamentos em *blisters*, registrados sob múltiplas orientações (horizontal, vertical e diagonal) e com variações no formato e organização das embalagens. A Figura 1 apresenta exemplos representativos dos conjuntos de dados utilizados.

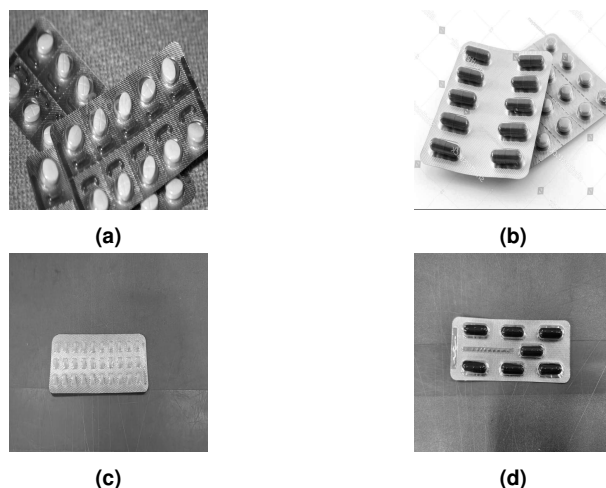


Figura 1. Exemplos representativos dos conjuntos de dados utilizados. Imagens a e b correspondem à base de dados pública, enquanto as figuras c e d pertencem ao conjunto de dados do HU-USP.

O terceiro conjunto foi gerado a partir da combinação dos dois anteriores, permitindo avaliar o impacto da diversidade e da origem dos dados no desempenho dos modelos. Em todos os conjuntos, os dados foram divididos em 80% para treinamento e 20% para teste.

3.2. Desenvolvimento

Os experimentos foram conduzidos em Python utilizando a biblioteca Ultralytics para implementação das arquiteturas YOLOv8n e YOLOv11n, inicializadas com pesos pré-treinados no conjunto de dados COCO [Lin et al. 2014] e ajustadas por meio de *fine-tuning*. O treinamento foi realizado com resolução de entrada de 640×640 pixels, *batch size* de 16 e 4 *workers* para carregamento de dados. A divisão entre conjuntos de treino e teste foi definida por arquivos de configuração no formato YAML.

Os experimentos foram executados em um computador com processador AMD Ryzen 7 5800X e 16 GB de memória RAM, sem utilização de GPU dedicada. Devido às

limitações computacionais, o treinamento foi realizado por 10 épocas para todos os experimentos. Essa configuração permitiu avaliar o comportamento inicial de convergência dos modelos e comparar o impacto da arquitetura e da origem dos conjuntos de dados sob condições experimentais equivalentes.

3.3. Métricas de Avaliação

O desempenho dos modelos foi avaliado por meio das métricas *Precision*, *Recall* e *mean Average Precision* (mAP), amplamente empregadas em tarefas de detecção de objetos. A métrica *Precision* quantifica a proporção de detecções corretas em relação ao total de predições realizadas pelo modelo, enquanto *Recall* indica a proporção de objetos corretamente identificados em relação ao número total de objetos presentes no conjunto de validação. Dentre as métricas adotadas, destaca-se o mAP@0.5, que corresponde à média da *Average Precision* (AP) calculada para predições com Intersection over Union (IoU) superior a 0,5. O IoU mede o grau de sobreposição entre a *bounding box* prevista e a *bounding box* de referência (*ground truth*), sendo utilizado como critério para determinar se uma detecção é considerada correta.

4. Resultados

A Tabela 1 apresenta os resultados obtidos pelos modelos YOLOv8n e YOLOv11n nos diferentes conjuntos de dados, considerando as métricas *Precision*, *Recall* e mAP@0.5.

Tabela 1. Desempenho dos modelos nos diferentes conjuntos de dados

Modelo	Conjuntos de dados	<i>Precision</i>	<i>Recall</i>	mAP@0.5
YOLOv8n	Público	0.958	0.971	0.988
YOLOv8n	HU	0.989	0.987	0.994
YOLOv8n	Combinado	0.959	0.974	0.988
YOLOv11n	Público	0.957	0.975	0.987
YOLOv11n	HU	0.986	0.985	0.994
YOLOv11n	Combinado	0.959	0.970	0.986

De modo geral, os resultados indicam desempenho semelhante entre as duas arquiteturas avaliadas, com diferenças pouco expressivas nas métricas consideradas. Em média, o modelo YOLOv8n apresentou valores ligeiramente superiores quando comparado ao YOLOv11n, embora a diferença observada tenha sido pequena. Embora pequenas variações nos valores de *Precision*, *Recall* e mAP@0.5 possam ser observadas entre YOLOv8n e YOLOv11n, uma análise estatística pareada entre os resultados obtidos nos diferentes conjuntos de dados não evidenciou ganho estatisticamente significativo (teste t pareado, $p > 0.05$), indicando desempenho comparável entre as duas arquiteturas.

Observou-se também uma pequena variação no desempenho dos modelos de acordo com o conjunto de dados utilizado para treinamento. Os modelos treinados com o conjunto de imagens coletadas no HU-USP apresentaram valores ligeiramente superiores de *Precision*, *Recall* e mAP@0.5 em comparação com aqueles treinados exclusivamente com o conjunto de dados público. Entretanto, deve-se considerar que o conjunto HU-USP possui menor número de imagens, o que limita conclusões mais amplas sobre o impacto do tamanho e da origem dos dados no desempenho dos modelos.

O conjunto combinado, formado pela união da base de dados pública e hospitalar, apresentou desempenho comparável aos demais cenários. Esse resultado sugere que a inclusão de diferentes fontes de dados pode ampliar a diversidade estrutural do treinamento sem causar degradação significativa no desempenho do modelo.

Além da análise quantitativa apresentada na Tabela 1, foi realizada também uma análise qualitativa das predições geradas pelos modelos. Para isso, imagens do conjunto HU-USP foram submetidas ao modelo treinado exclusivamente com o conjunto de dados público, permitindo avaliar seu comportamento em dados provenientes de um ambiente hospitalar real.

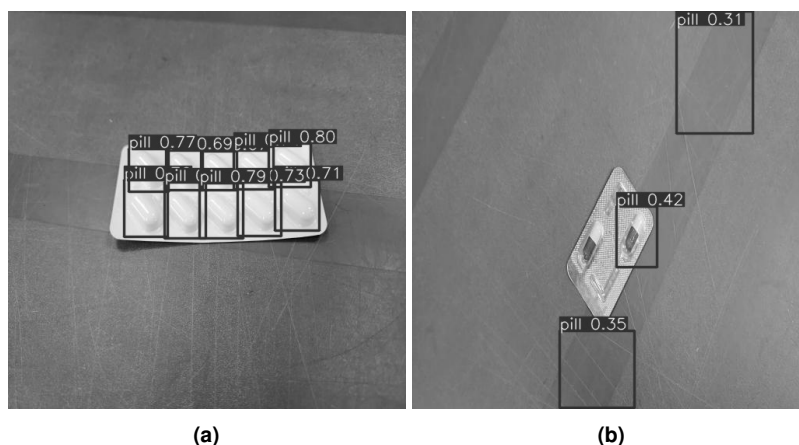


Figura 2. Exemplos de predições do modelo treinado com o conjunto de dados público aplicadas a imagens do HU-USP.

A Figura 2 apresenta exemplos dessas predições. Observa-se que, embora o modelo seja capaz de identificar corretamente a maioria dos comprimidos presentes nos *blisters*, algumas dificuldades ocorrem em determinadas configurações estruturais e ângulos de captura, especialmente em imagens registradas em orientação diagonal. Na Figura 2(b), observa-se um exemplo em que o *blister* está posicionado diagonalmente, resultando em erros de detecção. Esses casos ilustram desafios de generalização quando modelos treinados em um domínio são aplicados a dados provenientes de contextos distintos. Esses resultados contribuem para compreender o impacto da origem e da diversidade estrutural dos conjuntos de dados no desempenho de modelos YOLO aplicados à detecção de comprimidos em *blisters*.

5. Conclusão

Este trabalho investigou o desempenho das arquiteturas YOLOv8 e YOLOv11 na detecção de comprimidos em *blisters* farmacêuticos, analisando o impacto da origem e da diversidade estrutural dos conjuntos de dados utilizados no treinamento. Os resultados indicaram desempenho semelhante entre as arquiteturas avaliadas, com diferenças pouco expressivas nas métricas de avaliação, além de sugerir que a diversidade dos dados pode contribuir para ampliar a capacidade de generalização dos modelos. É importante destacar que parte das contribuições desse trabalho foi a coleta de dados diretamente no ambiente hospitalar do HU-USP com o objetivo de garantir a personalização do treinamento do modelo para dados reais.

Como continuidade desta pesquisa, pretende-se integrar os modelos de detecção a métodos de cálculo de trajetórias de corte para o fracionamento automatizado de *blisters* farmacêuticos. Entre as abordagens a serem exploradas estão o uso de Diagramas de Voronoi para delimitação das regiões de corte e a aplicação de técnicas de segmentação semântica para melhorar a identificação das áreas dos comprimidos. Espera-se que essas etapas contribuam para o desenvolvimento de uma solução mais robusta voltada à automação do processo de distribuição por dose unitária no ambiente hospitalar.

Agradecimentos

O presente trabalho foi realizado com apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Brasil. Processo nº 2025/05640-5 e da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Referências

- Andrade, L. S. d. and Freitas, C. d. B. L. (2019). Farmacoterapia: análise da implantação de distribuição individualizada de medicamentos em um hospital municipal do rio de janeiro. *Academus Revista Científica da Saúde*, 4(2):18–25.
- Jiang, K., Pan, S., Yang, L., Yu, J., Lin, Y., and Wang, H. (2023). Surgical instrument recognition based on improved yolov5. *Applied Sciences*, 13(21).
- Lin, T.-Y. et al. (2014). Microsoft coco: Common objects in context. In *ECCV*.
- Ragab, M. G., Abdulkadir, S. J., Muneer, A., Alqushaibi, A., Sumiea, E. H., Qureshi, R., Al-Selwi, S. M., and Alhussian, H. (2024). A comprehensive systematic review of yolo for medical object detection (2018 to 2023). *IEEE Access*, 12:57815–57836.
- Roboflow (2024). blister dataset. <https://universe.roboflow.com/teste-wjjov/blister-uurot>. Acesso em 2026-02-22.
- Soni, A. and Rai, A. (2024). Yolo for medical object detection (2018–2024). In *2024 IEEE 3rd International Conference on Electrical Power and Energy Systems (ICE-PES)*, pages 1–7.
- Tani, A. A., Norberto, J. M., da Silva, M. R., and Durand-Petiteville, A. (2024). Automated computation of cutting paths for unit-dose repackaging of pharmaceutical blister packs. In *2024 IEEE 20th International Conference on Automation Science and Engineering (CASE)*, page 1678–1683. IEEE.
- Universidade de São Paulo (2023). Anuário estatístico da usp. <https://uspdigital.usp.br/anuario/AnuarioControle>. Acesso em 06 mar. 2025.
- Yoloultra (2025). Tablet blister dataset. <https://universe.roboflow.com/yoloultra/tablet-blister-rhrec>. Acesso em 2026-02-22.