

Sistema Interativo de Contagem de Dedos usando YOLOv11 com Aplicação Web via Anvil Uplink

Juan Santos Trigo Nasser,¹ Ana Paula A. C. Shiguemori,²
IFSP, Jacareí, SP

Resumo

A contagem automática de dedos é uma aplicação de visão computacional com relevância em interfaces gestuais, ensino interativo e inclusão digital [1]. Métodos que permitem interação sem contato físico têm ganhado importância em contextos educacionais, industriais e de acessibilidade, favorecendo experiências naturais e seguras [2].

Este trabalho propõe um sistema interativo para detecção e contagem de dedos utilizando a rede YOLOv11, com implementação web via Anvil [5], permitindo acesso remoto sem necessidade de instalação local. O conjunto de dados foi inicialmente composto por 680 imagens rotuladas manualmente, expandido para 1.632 imagens por meio de técnicas de aumento de dados no framework Roboflow [3]. A divisão seguiu a proporção Train Set (1.428), Validation Set (136) e Test Set (68), garantindo equilíbrio entre treinamento, ajuste fino e avaliação. O modelo foi treinado por 10 épocas, com batch size de 16 e imagens redimensionadas para 640×640 pixels. O otimizador utilizado foi AdamW (learning rate = 0,002, momentum = 0,9, weight decay = 0,0005), aliado a estratégias de data augmentation como mosaic, flip horizontal (50%) e RandAugment, além de parâmetros de warm-up (3 épocas, warmup momentum = 0,8, warmup bias lr = 0,1). O modelo atingiu mAP@0.5 de 97,3%, apresentando precisão (P) de 97,5% e recall (R) de 97,3%, mantendo desempenho consistente mesmo diante de oclusões parciais, variações de iluminação e diversidade de poses. A média de processamento por imagem foi de 13,8 ms (72 FPS), confirmando sua aplicabilidade em tempo real. A interface interativa permite que o usuário utilize gestos para responder a desafios numéricos, demonstrando potencial para ensino de matemática, inclusão digital e controle gestual em ambientes industriais.

A análise crítica dos resultados indica que, embora o modelo apresente alto desempenho, sua robustez depende da diversidade do dataset e da qualidade da rotulagem. Futuras investigações podem explorar conjuntos de dados mais variados, integração com sensores adicionais e adaptação a cenários complexos, como controle remoto de drones e análise geoespacial, ampliando o espectro de aplicações práticas da proposta.

¹juanperuib6@gmail.com

²anapaula.acs@ifsp.edu.br

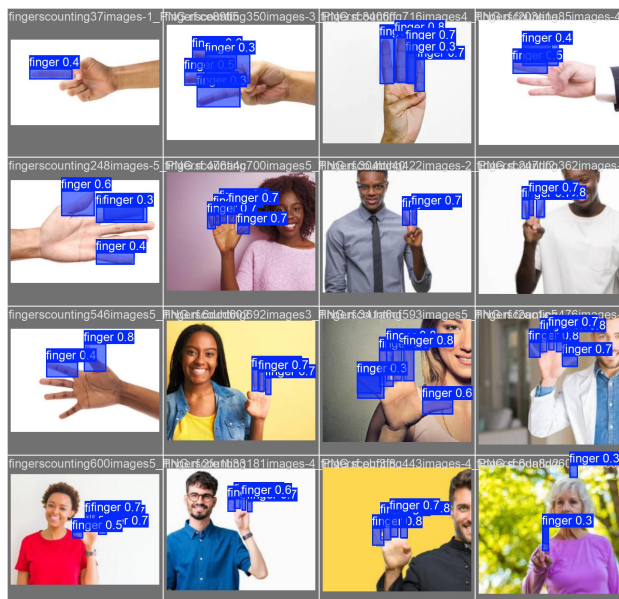


Figura 1. Exemplo de saída do modelo YOLOv11 para detecção de dedos. As caixas azuis indicam as regiões detectadas como “finger”, acompanhadas dos valores de confiança da predição.

Referências

- [1] Redmon, J., Divvala, S., Girshick, R. e Farhadi, A. *You Only Look Once: Unified, Real-Time Object Detection*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016.
- [2] Glorot, X. e Bengio, Y. *Understanding the difficulty of training deep feedforward neural networks*. Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, 2010.
- [3] Roboflow Inc. *Roboflow: Image Data Management for Computer Vision*, 2025. Disponível em: <https://roboflow.com/>.
- [4] Keskin, C., Agarwal, A. e Forsyth, D. A. *Real-time hand tracking and modeling from a single depth image*. 2011 International Conference on Computer Vision, 1120–1127, IEEE, 2011.
- [5] Luff, M. e Davies, I. *Anvil: A Platform for Rapid Development of Web Apps with Python*. Proceedings of the 19th Python in Science Conference, 2020. Disponível em: <https://anvil.works>.