

14º Congresso de Inovação, Ciência e Tecnologia do IFSP - 2023

VISÃO COMPUTACIONAL APLICADA À DETECÇÃO DE OBJETOS PARA A GESTÃO DO FLUXO DE PESSOAS

JÚPITER MENDES MICHELATO¹, TIAGO SOARES VITOR²

¹ Graduando em Engenharia de Controle e Automação, Bolsista PIBITI, IFSP, Câmpus São João da Boa Vista, jupiter.michelato@aluno.ifsp.edu.br.

² Professor de Engenharia de Controle e Automação, IFSP, Câmpus São João da Boa Vista, tiagosvitor@ifsp.edu.br.

Área de conhecimento (Tabela CNPq): 1.03.00.00-7 Ciência da computação

RESUMO: O projeto propõe o desenvolvimento de um sistema inteligente, baseado em visão computacional, para a detecção de pessoas em *frames* de vídeo. Uma detecção precisa é fundamental na solução de problemas de monitoramento e contagem automatizada do fluxo de pessoas. Tradicionalmente, o monitoramento humano apresenta limitações em termos de confiabilidade e eficiência, devido a fadiga e possíveis erros. Logo, uma abordagem computacional que empregue técnicas de processamento de imagens e algoritmos de aprendizado de máquina pode ser usada para o desenvolvimento de uma solução que seja eficiente e robusta para o problema. Para isso, a biblioteca de visão computacional OpenCV foi utilizada, viabilizando o uso do COCO *dataset* (banco de imagens) no treinamento de um modelo de classificação baseado na arquitetura de rede neural chamada SSD. O sistema implementado foi submetido a testes em diversos cenários e complexidades, apresentando eficiência e robustez na detecção de pessoas em *frames* de vídeo com altos níveis de confiança. Assim, essa solução se comprovou promissora, podendo ser incorporada a um sistema de rastreamento que compute o fluxo de pessoas e contribua na tomada de decisões do gestor.

PALAVRAS-CHAVE: detecção de objetos; processamento de imagem, visão computacional; OpenCV; automação.

COMPUTER VISION APPLIED TO OBJECT DETECTION FOR PEOPLE FLOW MANAGEMENT

ABSTRACT: The project proposes the development of an intelligent system, based on computer vision, for detection of people in video frames. Accurate detection is critical in problems of automated people flow monitoring and counting. Traditionally, human monitoring has limitations in terms of reliability and efficiency, due to fatigue and possible errors. Therefore, a computational approach that employs image processing techniques and machine learning algorithms can be used to develop a solution that is efficient and robust for the problem. For this, the OpenCV computer vision library was used, enabling the use of the COCO dataset (image bank) in training a classification model based on the neural network architecture called SSD. The implemented system was tested in different scenarios and complexities, showing efficiency and robustness in detecting people in video frames with high levels of confidence. Thus, this solution has proven to be promising, and can be incorporated into a tracking system that computes the flow of people and contributes to the manager's decision-making.

KEYWORDS: object detection; image processing, computer vision; OpenCV; automation.

INTRODUÇÃO

A crescente disponibilidade de câmeras de vigilância e o avanço na área de visão computacional têm despertado interesse na aplicação dessas tecnologias para o monitoramento automatizado do fluxo de pessoas em diversos contextos. No entanto, a eficácia desses sistemas em ambientes complexos e dinâmicos é uma questão desafiadora. Este trabalho objetiva explorar técnicas de processamento de imagens e aprendizado de máquina para criar um sistema que possa detectar pessoas de forma precisa. Sistemas como este podem alimentar técnicas de rastreamento para a contagem do número de pessoas em movimento.

Várias aplicações do monitoramento de fluxo de pessoas podem ser listadas na literatura. Estudos quanto à circulação de pessoas mostraram que é possível condicionar a configuração espacial dos espaços públicos de forma que interfiram na vitalidade desses espaços (Arrais; Medeiros, 2016). Um projeto foi desenvolvido para automatizar a entrega de equipamentos para laboratórios de ensino em ambiente universitário a fim de desafogar o fluxo de pessoas (Araújo et al., 2017). Com isso, o controle de abastecimento foi melhorado, gerando mais segurança e agilidade na retirada dos kits. No contexto pandêmico (Oliveira; Oliveira; Júnior, 2021), o fluxo de pessoas é controlado para o melhor manejo da doença.

Tradicionalmente, operadores manuais precisam acompanhar constantemente o *feed* de vídeo para fazer o levantamento de informações do movimento das pessoas. Tais processos são propensos a erros humanos devido à fadiga, exaustão, desatenção ou complexidade da supervisão. Portanto, tal abordagem é ineficaz em condições que requerem maior confiabilidade.

Soluções usando Visão Computacional podem superar as limitações dos métodos tradicionais de contagem manual e oferecer informações estratégicas para a tomada de decisões em diversos segmentos da sociedade, como planejamento urbano, segurança pública e gestão de espaços públicos (Maregoni; Stringhini, 2009). Portanto, o objetivo deste trabalho é desenvolver uma solução que permita a detecção de pessoas em frames para um posterior rastreamento.

Para este fim, este trabalho utiliza a biblioteca de visão computacional OpenCV (Opencv Team, 2022) e explora um conjunto de imagens, COCO *dataset* (Lin, 2014; Coco Dataset, 2021), para o treinamento do algoritmo de detecção, alinhado com o trabalho de Liu et al. (2016) que introduziu uma arquitetura de rede neural chamada de *Single Shot MultiBox Detector* (SSD). O SSD é capaz de detectar objetos presentes na imagem usando múltiplas caixas. Após o desenvolvimento do sistema, vários testes foram realizados para a validação em cenários diversos, mostrando-se promissor para a detecção de pessoas.

MATERIAL E MÉTODOS

O software, escrito em linguagem Python, se encontra na Figura 1. Inicialmente, deve-se importar a biblioteca OpenCV. Ela fornece diversas funções e ferramentas essenciais para processamento de imagens, detecção de objetos e análise de vídeo em tempo real. Além da biblioteca, alguns recursos são essenciais para a implementação da capacidade de se reconhecer objetos/pessoas. São eles:

COCO dataset: *Common Objects in Context* é um dos maiores bancos de imagens, tendo sido criado para obter o “estado da arte” dos algoritmos de reconhecimento de objetos. As imagens em seu banco de dados são imagens complexas do dia a dia, sendo ótimas para treinar inteligências artificiais com imagens em seu contexto natural (Lin, 2014). Suas imagens são separadas e categorizadas em 80 categorias diferentes, fazendo com que uma inteligência artificial treinada usando este *dataset* seja capaz de reconhecer 80 tipos de objetos diferentes.

ssd_mobilenet_v3_large_coco_2020_01_14: Este arquivo, disponível no OpenCV, é uma rede de SSD, sendo utilizada na detecção de objetos, já sendo treinada utilizando o COCO *dataset*. Os SSD's foram criados em 2016 e utilizam uma única rede neural profunda para detectar imagens, sendo mais fácil treinar devido aos métodos que utiliza para a geração das caixas de detecção, compilando todo o trabalho computacional gerado na criação destas caixas em uma única rede.

frozen_inference_graph.pb: *Freezing* é o processo de identificar, salvar e disponibilizar em um único arquivo todos os dados necessários, como gráficos, pesos, parâmetros de treinamento, entre outros, não podendo mais ser treinado. Esses dados são utilizados para inferência no *TensorFlow* (Google Brain, 2015) e podem ser usados em outros frameworks.

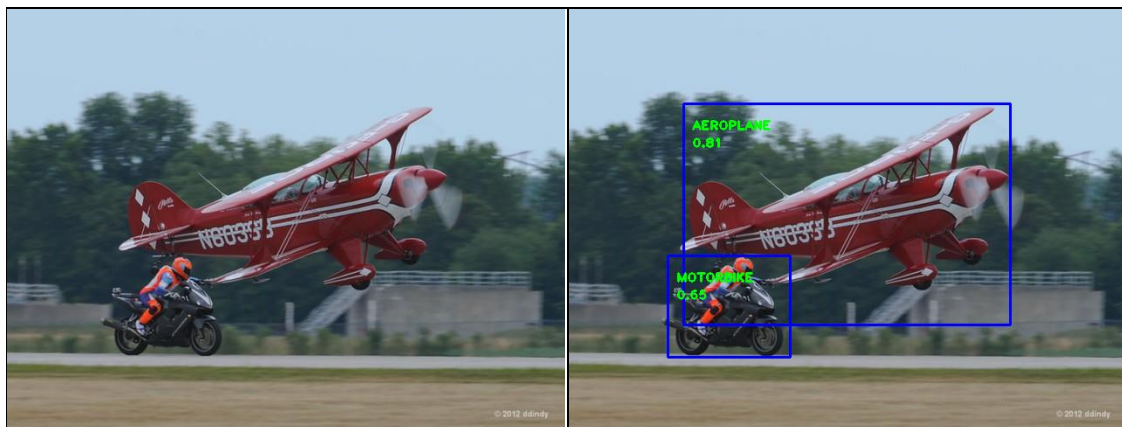


FIGURA 3. Detecção de avião, motocicleta e pessoa.



FIGURA 4. Detecção de carro e caminhão.



FIGURA 5. Detecção de pessoas e bola.

No geral, o sistema possui habilidade de identificar objetos em imagens. As limitações verificadas nos testes apontam dificuldades do algoritmo em trabalhar com cenários onde objetos estão sobrepostos. Também, objetos pequenos ou que estão parcialmente na figura podem ser mais difíceis de serem classificados. Entretanto, como o objetivo final é detectar pessoas, é possível concluir que o algoritmo é capaz de identificá-las com altos valores de confiança, quando a imagem não possui as características limitantes citadas.

Uma vez validada a solução para detecção de pessoas em imagens, é possível submetê-la a um frame de vídeo de pessoas se movimentando. A Figura 6 mostra o resultado para um determinado instante de tempo do vídeo *Slow Motion High Angle View Crowd Walking People* de mack2happy, do Canva (Cava Company, 2023). Podemos observar que o algoritmo foi eficiente em detectar todas as

peças do frame que estão completamente enquadradas. Também altos níveis de confiança foram calculados. Portanto, a solução proposta tem viabilidade para alimentar um sistema automação da contagem de pessoas através da visão computacional.



FIGURA 6. Detecção de pessoas e a utilização de centroide.

CONCLUSÕES

Os resultados alcançados demonstraram que o sistema desenvolvido apresenta uma promissora capacidade de identificação e classificação de objetos em imagens. A utilização do COCO dataset e a adoção da arquitetura SSD, por meio da biblioteca OpenCV, permitiram que o modelo obtivesse sucesso em diversas situações e em diferentes níveis de complexidade. Também, desafios e limitações foram evidenciados. O sistema apresentou certa limitação na detecção de objetos sobrepostos, parcialmente exibidos ou distantes. Entretanto, como o objetivo principal é a detecção de pessoas, a solução se mostrou eficiente com altos valores de confiabilidade.

As ferramentas tecnológicas utilizadas para a solução do problema se mostraram adequadas e robustas. Elas permitiram a criação de sistema promissor para a aplicação em estudos de gestão do fluxo de pessoas, destacando a visão computacional como uma importante abordagem para soluções inovadoras e eficientes em problemas reais. Trabalhos futuros poderão buscar a redução das limitações encontradas em certos cenários através do refinamento no modelo e uso de estratégias de sistemas inteligentes.

CONTRIBUIÇÕES DOS AUTORES

J.M.M. contribuiu com a pesquisa, desenvolvimento, implementação e teste de software. T.S.V. atuou na supervisão, metodologia e validação de dados. Ambos contribuíram na redação do manuscrito original. Todos os autores contribuíram com a revisão do trabalho e aprovaram a versão submetida.

AGRADECIMENTOS

Gostaríamos agradecer ao Instituto Federal de São Paulo pelo apoio à pesquisa de iniciação científica. Também, aos professores e colegas pelo suporte e troca de conhecimentos ao longo desse processo. Ao CNPq pela bolsa e incentivo.

REFERÊNCIAS

ARRAIS, J. S. S.; MEDEIROS, V. A. S. **A Configuração Espacial e a Vitalidade dos Espaços Públicos no Centro de Goiânia**. 7º Congresso Luso-Brasileiro para o Planejamento Urbano, Regional, Integrado e Sustentável, p. 12, out. 2016.

ARAÚJO, N. A. R. et al. **Automação do Controle do Fluxo de Equipamentos nos Laboratórios Universitários**. 16º Congresso Nacional de Iniciação Científica – CONIC-SEMESP, jan. 2017.

CANVA COMPANY. **Canva**. Disponível em: <https://www.canva.com/>, Acesso em 20 de agosto de 2023.

COCO dataset. COCO: Common Objects in Context. [S. l.], 2021. Disponível em: <https://cocodataset.org/#home>, Acesso em: 21 fev. 2023.

Google Brain. **TensorFlow**, 2015. Disponível em: <https://www.tensorflow.org/?hl=pt-br>, Acesso em: 21 fev. 2023.

LIN, T. et al. **Microsoft coco: Common objects in context**. Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, Proceedings, Part V 13. Springer International Publishing, p. 740-755, Sept. 2014.

LIU, W. et al. (2016). **SSD: Single Shot MultiBox Detector**. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds) Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science, vol 9905. Springer, 2016.

MAREGONI, M.; STRINGHINI, D. **Tutorial: Introdução à visão computacional usando openCV**. RITA, volume XVI, número 1. 2009. Acesso em: 21 fev. 2023.

OLIVEIRA, S. R. DE; OLIVEIRA, M. A. DE C.; JUNIOR, D. F. C. **A Vigilância por dados no contexto da Covid-19: implicações e desafios para as relações de trabalho**. Teoria Jurídica Contemporânea, v. 6, n. 0, 2021.

OPENCV TEAM. **OpenCV**. Disponível em: <https://opencv.org/>, Acesso em: 20 jun. 2022.