

HỌC CÓ NHẬN THỨC VỀ NGŨ CẢNH ĐỂ TĂNG CƯỜNG PHÁT HIỆN VI PHẠM GIAO THÔNG VỀ MŨ BẢO HIỂM

Nguyễn Tiến Huy - 22520567
Phan Nguyễn Hữu Phong - 22521090

Tóm tắt

- Link Github của nhóm: <https://github.com/fonzi22/CS519.021.KHTN/>
- Link YouTube video: https://www.youtube.com/watch?v=c7RB9x_B3l4
- Các thành viên:



Phan Nguyễn Hữu Phong
22521090



Nguyễn Tiến Huy
22520567

Giới thiệu

- Xuất phát từ mối lo ngại, liệu mô hình có thực sự nhận biết được đặc điểm môi trường xung quanh đối tượng cần phát hiện hay không.
- Các phương pháp phát hiện đối tượng dựa trên deep learning chỉ tập trung vào việc học cách “phát hiện” các đối tượng trong các lớp được xác định trước đó.
- Vì vậy, chúng tôi đặt ra câu hỏi: Liệu có thể tìm hiểu ngữ cảnh trong bài toán phát hiện người tham gia giao thông có đội mũ bảo hiểm hay không?

Mục tiêu

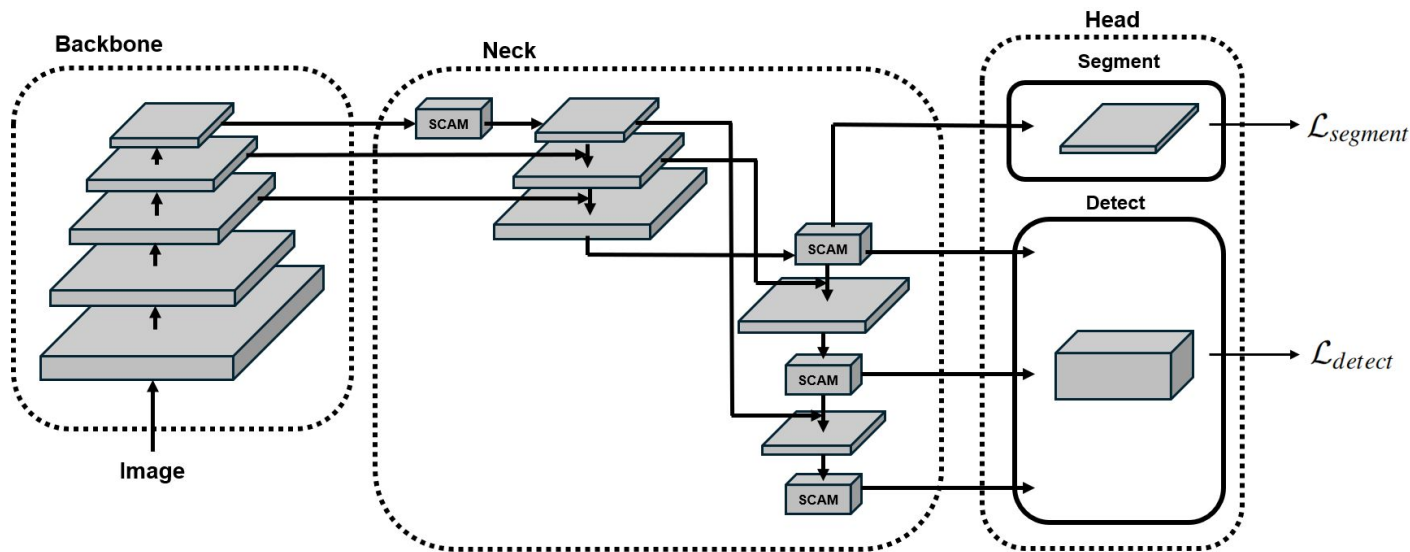
- Xây dựng hệ thống phát hiện vi phạm đội mũ bảo hiểm xe máy với độ chính xác cao và có thời gian suy luận nhanh nhanh trong thời gian thực.
- Cải thiện hiệu suất của mô hình phát hiện đối tượng bằng cách kết hợp thêm thông tin về ngữ cảnh và một cơ chế chú ý mới Spatial-Channel Attention Module (SCAM).
- Xây dựng một mô hình nhẹ với ít tham số giúp giảm thời gian huấn luyện cũng như thời gian suy luận nhưng vẫn đảm bảo hiệu suất so với các mô hình deep learning có số lượng tham số lớn.

Nội dung và Phương pháp

- **Nắm bắt các phương pháp hiện tại:** Thông qua việc đọc các bài báo nghiên cứu khoa học, để cập nhật kiến thức.
- **Sử dụng và thực nghiệm các mô hình về bài toán:** tìm hiểu, xây dựng ra pipeline cơ bản để giải quyết bài toán trước khi đi tìm câu trả lời cho câu hỏi nghiên cứu.
- **Tìm hiểu sâu vào ‘Học có nhận thức ngữ cảnh trong Deep learning’:** Tập trung vào các bài báo tạp chí, hội nghị liên quan đến vấn đề đang quan tâm, và rút trích ra những keyword, nội dung quan trọng.

Nội dung và Phương pháp

- **Xây dựng pipeline cơ bản:** Tìm hiểu các hệ thống phát hiện đối tượng dựa trên các mạng học sâu gần đây và lựa chọn 1 hệ thống phù hợp với điều kiện về tài nguyên, cụ thể là Yolo.



Nội dung và Phương pháp

- **Thu thập và tiền xử lý dữ liệu:** Là bước chuẩn bị dữ liệu cho việc huấn luyện mô hình. Ngoài ra, để dữ liệu được chuẩn hoá, nên dữ liệu được qua bước tiền xử lý.
- **Huấn luyện mô hình:** Từ dữ liệu đã thu thập và pipeline cơ bản (dựa trên sự tìm hiểu, nhóm quyết định sử dụng YOLO), nhóm tiến hành huấn luyện mô hình phát hiện đối tượng.

Nội dung và Phương pháp

- **Cải tiến (tinh chỉnh mô hình với các quan sát):** Tìm hiểu điều gì thực sự ảnh hưởng tới việc đưa ra dự đoán, và nhận thấy, khi sử dụng các segmentation map trong quá trình học. Từ đó có thể đưa ra những cải tiến tốt hơn cho mô hình.
- **Thực nghiệm kết quả:** Tiến hành cài đặt 'observation' và thực nghiệm sự ảnh hưởng của chúng trong việc đưa ra các dự đoán phát hiện về đối tượng vi phạm giao thông về mũ bảo hiểm.

Kết quả dự kiến

- Việc kết hợp thông tin về ngữ cảnh cũng như áp dụng cơ chế chú ý SCAM giúp cải thiện được hiệu suất của mô hình phát hiện vi phạm mã bảo hiểm so với việc không sử dụng các phương pháp này.
- Mô hình có số lượng tham số nhỏ với ít hơn 10 triệu tham số, giúp cải thiện đáng kể thời gian huấn luyện cũng như thời gian suy luận.
- Chứng minh mô hình cho thấy độ hiệu quả cao khi so sánh với các mô hình học sâu khác trên cùng một tập dữ liệu.

Tài liệu tham khảo

- [1]. Al-Sultan, S., Al-Bayatti, A.H., Zedan, H.: Context-aware driver behavior detection system in intelligent transportation systems. *IEEE Transactions on Vehicular Technology* 62(9), 4264–4275 (2013).
- [2] Siris, A., Jiao, J., Tam, G. K., Xie, X., & Lau, R. W. (2021). Scene context-aware salient object detection. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 4156-4166).
- [3] Fifty, C., Duan, D., Junkins, R. G., Amid, E., Leskovec, J., Ré, C., & Thrun, S. (2023). Context-aware meta-learning. *arXiv preprint arXiv:2310.10971*.
- [4] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- [5] Park, J., Woo, S., Lee, J. Y., & Kweon, I. S. (2018). Bam: Bottleneck attention module. *arXiv preprint arXiv:1807.06514*.
- [6] Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). Cbam: Convolutional block attention module. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 3-19).