# THÔNG TIN CHUNG CỦA NHÓM

- Link YouTube video của báo cáo (tối đa 5 phút): *https://youtu.be/c7RB9x_B3l4*
- Link slides (dạng .pdf đặt trên Github của nhóm): *https://github.com/huynt654/CS519.O21.KHTN---Scientific-research-methodology/blob/main/slides.pdf*

| | |
|---|---|
| • Họ và Tên: Phan Nguyễn Hữu Phong<br>• MSSV: 22521090<br> <br> | • Lớp: CS519.O21.KHTN<br>• Tự đánh giá (điểm tổng kết môn): 8.5/10<br>• Số buổi vắng: 1<br>• Số câu hỏi QT cá nhân: 9<br>• Số câu hỏi QT cả nhóm: 15<br>• Link Github: https://github.com/fonzi22<br>• Mô tả công việc:<br>- Lên ý tưởng đồ án.<br>- Làm slide.<br>- Chỉnh sửa poster. |
| • Họ và Tên: Nguyễn Tiến Huy<br>• MSSV: 22520567<br> | • Lớp: CS519.O21.KHTN<br>• Tự đánh giá (điểm tổng kết môn): 9/10<br>• Số buổi vắng: 0<br>• Số câu hỏi QT cá nhân: 6<br>• Số câu hỏi QT cả nhóm: 15<br>• Link Github: https://github.com/huynt654<br>• Mô tả công việc:<br>- Lên ý tưởng đồ án.<br>- Viết proposal.<br>- Làm video thuyết trình. |

- **ĐỀ CƯƠNG NGHIÊN CỨU**

| |
|---|
| **TÊN ĐỀ TÀI (IN HOA)**<br>HỌC CÓ NHẬN THỨC VỀ NGỮ CẢNH ĐỂ TĂNG CƯỜNG PHÁT HIỆN VI PHẠM GIAO THÔNG VỀ MŨ BẢO HIỂM |
| **TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)**<br>CONTEXT-AWARE LEARNING FOR ENHANCING TRAFFIC HELMET VIOLATION DETECTION |
| **TÓM TẮT** *(Tối đa 400 từ)*<br>Trong bối cảnh đô thị hiện đại ngày càng tăng, tai nạn xe cộ trở thành một trong những nguyên nhân quan trọng gây thiệt hại về người và tài sản, với đặc biệt đáng chú ý là tai nạn xe máy. Một yếu tố quan trọng dẫn đến làm cho các vụ tai nạn này dẫn đến các hậu quả nghiêm trọng là việc cả người lái xe và người ngồi trên xe không đội mũ bảo hiểm. Với sự phát triển mạnh mẽ của học máy và học sâu, các phương pháp phát hiện đối tượng hiện tại đã cho thấy hiệu quả đáng kể. Tuy nhiên, khi đối mặt với các bối cảnh môi trường phức tạp, các phương pháp này vẫn bộc lộ một số hạn chế. Để giải quyết vấn đề này, nghiên cứu này trình bày một hệ thống phát hiện mũ bảo hiểm giao thông mới tận dụng việc học có nhận thức ngữ cảnh về mặt ngữ nghĩa. Nó giúp khai thác các pixel ngữ nghĩa về vị trí của chúng và đối tượng là gì. Hơn nữa, chúng tôi đề xuất một cơ chế chú ý (Attention Mechanism) mới, Spatial-Channel Attention Module (SCAM), một cơ chế chú ý cải tiến tập trung hiệu quả vào các đối tượng ở nhiều kích thước trong không gian. Chúng tôi hy vọng cách tiếp cận mới liên quan đến việc học tập theo ngữ cảnh có thể cho kết quả cao trong việc phát hiện vi phạm giao thông theo thời gian thực. Bên cạnh đó, chúng tôi thực hiện so sánh phương pháp này với các mô hình deep learning khác để đánh giá hiệu suất. |
| **GIỚI THIỆU** *(Tối đa 1 trang A4)* |

Hiện nay, số người bị thương hoặc tử vong do tai nạn giao thông ngày càng gia tăng, đặc biệt là ở các nước chưa phát triển. Nguyên nhân có thể do sự thụ động hoặc lỗi của người tham gia giao thông. Việc sử dụng mũ bảo hiểm có thể làm giảm tai nạn giao thông và giảm chấn thương đầu. Vì vậy, ở hầu hết các nước, việc sử dụng mũ bảo hiểm cho người tham gia giao thông đã trở thành hợp pháp. Tuy nhiên, việc giám sát mũ bảo hiểm giao thông phụ thuộc rất nhiều vào lực lượng cảnh sát, việc này gây tốn kém về năng lượng và tiền bạc. Vì lý do này, mỗi quốc gia cần có một hệ thống phát hiện mũ bảo hiểm tự động. Tuy nhiên, một vấn đề khác là mỗi quốc gia sẽ có điều kiện giao thông khác nhau, dẫn đến hiệu suất của mô hình bị ảnh hưởng khi môi trường thay đổi giữa các quốc gia. Để giải quyết những vấn đề này, mô hình nhận thức được bối cảnh ngữ nghĩa phức tạp được đề xuất.

Sự phát triển của các phương pháp thị giác máy tính trong lĩnh vực hệ thống giao thông thông minh dựa trên deep learning, đặc biệt là Mạng CNN, cho kết quả ấn tượng trên nhiều tác vụ khác nhau như truy xuất hình ảnh, tạo hình ảnh, v.v.

Tận dụng khả năng học tập của các mô hình học sâu, các phương pháp đạt được hiệu suất cao, tuy nhiên chúng rất phức tạp và đòi hỏi nhiều tính toán. Việc áp dụng chúng trong thực tế để thay thế sự giám sát của cảnh sát bằng hệ thống phát hiện vi phạm đội mũ bảo hiểm giao thông trở nên không thực tế. Do đó, để có thể áp dụng vào thực tế, chúng tôi đề xuất một mô hình nhỏ hơn mang lại kết quả tương đương với các mô hình lớn hơn nhiều.

Trong bài báo này, chúng tôi tập trung phát triển các thuật toán tự động phát hiện vi phạm giao thông liên quan đến việc đội mũ bảo hiểm. Tóm lại, những đóng góp chính của chúng tôi được trình bày như sau:

- Phát triển 1 hệ thống phát hiện được thiết kế đặc biệt để phân biệt cả người tham gia giao thông và việc sử dụng mũ bảo hiểm, áp dụng phương pháp học tập theo ngữ cảnh.
- Giới thiệu một cơ chế chú ý mới Spatial-Channel Attention Module, được nghiên cứu trên nhiều quy mô, được thiết kế để tăng cường độ chính xác của việc phát hiện.
- Đáng chú ý, mô hình đề xuất thể hiện hiệu suất tương đương với các mô hình hiện đại trong khi vẫn duy trì kiến trúc gọn nhẹ thuận lợi cho các ứng dụng thực tế trong thế giới thực.

**Hình 1.** Tổng quan kiến trúc của mô hình phát hiện vi phạm mũ bảo hiểm.

**MỤC TIÊU** *(Viết trong vòng 3 mục tiêu)*

1. 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.
2. Cải thiện hiệu suất của mô hình phát hiện đối tiện bằng cách kết hợp thêm thông tin về ngữ cảnh và một cơ chế chú ý mới (SCAM).
3. 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**

***Nội dung***

Trong quá trình tìm hiểu, nhóm chúng em đặt ra câu hỏi: "Liệu việc biết thông tin ngữ cảnh có giúp cải thiện khả năng phát hiện đối tượng?"

Để trả lời và tìm giải pháp cho câu hỏi trên: nhóm chúng em quyết định bài toán cần giải quyết là phát hiện vi phạm giao thông về mũ bảo hiểm, để thực nghiệm cho câu hỏi trên, cũng như kiểm chứng cho giải pháp. Để cụ thể hoá quá trình tìm hiểu, nhóm em lên kế hoạch thực hiện bao gồm như sau:

- **Hiểu bài toán:** là việc nhóm đi tìm vấn đề thực sự đang nằm ở đâu, điều gì ảnh hưởng tới bài toán, những ràng buộc liên quan đến bài toán.
- **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:** là hoạt động 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.

***Phương pháp***

- **Thu thập và tiền xử lý dữ liệu:** Là bước chuẩn bị dữ liệu cho việc huyến luyện mô hình. Ngoài ra, để dữ liệu được chuẩn hoá, nên dữ liệu được qua bước tiền xử lý.

- **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.

- **Huyến 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 huyến luyện mô hình phát hiện đối tượng.

Hình 2: Ví dụ về segmentation map (ảnh phân đoạn)

- **Cải tiến (tinh chỉnh mô hình với các quan sát):** Sau đó, nhóm em 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, nó sẽ giúp khai thác các pixel ngữ nghĩa, mối tương quan giữa các pixel với nhau, đặc biệt, nó trả lời cho câu hỏi: "where are pixels belong "và "what are the objects?".

- **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.

Hình 3: Pipeline dự kiến thực hiện

**KẾT QUẢ MONG ĐỢI**

1. 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 mũ bảo hiểm so với việc không sử dụng các phương pháp này.
2. 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.
3. 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** *(Định dạng DBLP)*

[1]. Context-Aware Meta-Learning - *Christopher Fifty, Dennis Duan, Ronald Guenther Junkins, Ehsan Amid, Jure Leskovec, Christopher Re, Sebastian Thrun* - ICLR 2024

[2] Scene Context-Aware Salient Object Detection - ***Avishek Siris, Jianbo Jiao, Gary K.L. Tam, Xianghua Xie, Rynson W.H. Lau***; Proceedings of the IEEE/CVF ICCV, 2021, pp. 4156-4166