

BÁO CÁO ĐỒ ÁN CUỐI KỲ

Môn học

**CS519 - PHƯƠNG PHÁP LUẬN
NGHIÊN CỨU KHOA HỌC**

Lớp học

CS519.M11 - CS519.M11.KHCL

Giảng viên

PGS.TS. LÊ ĐÌNH DUY


Thời gian

09/2021 - 12/2021

----- *Trang này cố tình để trống* -----

THÔNG TIN CHUNG CỦA NHÓM

- Link YouTube video của báo cáo (tối đa 5 phút):
https://www.youtube.com/watch?v=o8mCTXNJ6r8&ab_channel=Th%E1%BA%AFngTr%C6%B0%C6%A1ngTh%C3%A0nh
- Link slides (dạng .pdf đặt trên Github của nhóm):
<https://github.com/erwin24092002/CS519.N11/blob/main/Slide.pdf>
- *Mỗi thành viên của nhóm điền thông tin vào một dòng theo mẫu bên dưới*
- *Sau đó điền vào Đề cương nghiên cứu (tối đa 5 trang), rồi chọn Turn in.*

<ul style="list-style-type: none">● Họ và Tên: Trương Thành Thắng● MSSV: 20521907 	<ul style="list-style-type: none">● Lớp: CS519.N11● Tự đánh giá (điểm tổng kết môn): 9/10● Số buổi vắng: 2 (có xin phép qua mail)● Số câu hỏi QT cá nhân: 10● Số câu hỏi QT của cả nhóm: 3● Link Github: https://github.com/erwin24092002/CS519.N11● Mô tả công việc và đóng góp của cá nhân cho kết quả của nhóm:<ul style="list-style-type: none">○ Đề xuất bài toán.○ Lên ý tưởng cải tiến.○ Khảo sát, nghiên cứu các mô hình phân lớp.○ Viết chính Đề cương nghiên cứu.○ Làm slide Powerpoint.
<ul style="list-style-type: none">● Họ và Tên: Ngô Văn Tấn Lưu● MSSV: 20521591	<ul style="list-style-type: none">● Lớp CS519.N11● Tự đánh giá (điểm tổng kết môn): 9/10● Số buổi vắng: 2



- Số câu hỏi QT cá nhân: 10
- Số câu hỏi QT của cả nhóm: 3
- Link Github:
<https://github.com/tanluuuuuuu/CS519.N11>
- Mô tả công việc và đóng góp của cá nhân cho kết quả của nhóm:
 - Đóng góp ý tưởng.
 - Khảo sát, nghiên cứu các mô hình phát hiện đối tượng có hướng.
 - Bổ sung phần Nội dung và Phương pháp nghiên cứu.
 - Làm video Youtube
 - Làm poster.

ĐỀ CƯƠNG NGHIÊN CỨU

TÊN ĐỀ TÀI (IN HOA)

GIẢM THIỂU HỘP GIỚI HẠN NHẦM LẦN TRONG BÀI TOÁN PHÁT HIỆN ĐỐI TƯỢNG CÓ HƯỚNG TRONG KHÔNG ẢNH

TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)

REDUCING CONFUSION IN AERIAL IMAGE ORIENTED OBJECT DETECTION VIA BOUNDING BOX CONSOLIDATION

TÓM TẮT (Tối đa 400 từ)

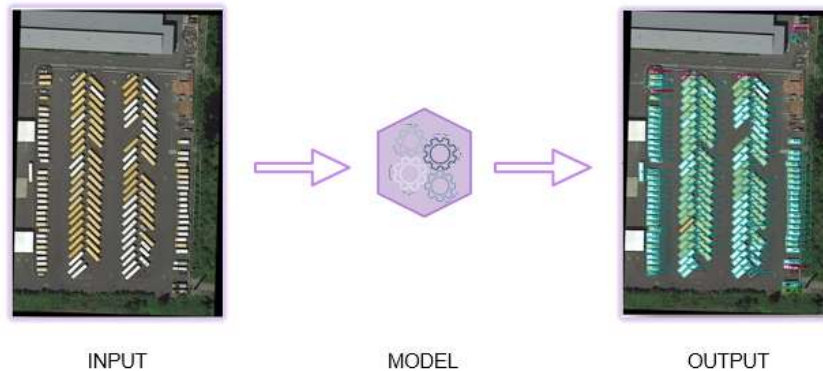
"Giảm thiểu hộp giới hạn nhầm lẫn trong bài toán phát hiện đối tượng có hướng trong không ảnh" là nghiên cứu giải quyết vấn đề nhiều hộp giới hạn được gán cho cùng một đối tượng trong quá trình phát hiện đối tượng có hướng trong không ảnh, điều có thể dẫn đến sự không chính xác và nhầm lẫn. Nghiên cứu đề xuất một cách tiếp cận hai bước kết hợp các mô hình phát hiện đối tượng có hướng và mô hình phân loại.

Các mô hình phát hiện đối tượng có hướng được sử dụng trong nghiên cứu được thiết kế để xử lý các thách thức đặc biệt trong việc phát hiện các đối tượng trong không ảnh, trong khi mô hình phân loại được sử dụng để gán nhãn lại các đầu ra của hộp giới hạn có hướng và giảm sự nhầm lẫn trong kết quả. Kết quả đầu ra từ cả hai mô hình được tổng hợp và non-maximum suppression được áp dụng để loại bỏ các hộp giới hạn dư thừa hoặc không chính xác.

Nghiên cứu sử dụng các mô hình phát hiện đối tượng có hướng như S2ANet [2], R3Det [1], ReDet, Oriented RCNN [5] và RoI Transformer [3] và kết hợp nó với mô hình phân loại EfficientNet [7]. Bộ dữ liệu DOTA [6] được sử dụng để chứng minh rằng phương pháp hai bước có hiệu quả trong việc giảm nhầm lẫn và cải thiện độ chính xác của việc phát hiện đối tượng trong không ảnh. Cách tiếp cận của dự án có ý nghĩa thực tế quan trọng đối với các ứng dụng khác nhau, bao gồm ứng phó với thiên tai, quy hoạch đô thị và nông nghiệp.

GIỚI THIỆU (Tối đa 1 trang A4)

Không ảnh là một công cụ quan trọng cho nhiều ứng dụng, bao gồm ứng phó thảm họa, quy hoạch đô thị và nông nghiệp. Tuy nhiên, việc phát hiện các vật thể trong không ảnh là một thách thức, vì các vật thể có thể xuất hiện ở các góc độ và hướng khác nhau, đồng thời nhiều vật thể có thể xuất hiện ở gần nhau. Một vấn đề phổ biến trong phát hiện đối tượng trong không ảnh là việc gán nhiều hộp giới hạn cho cùng một đối tượng, chỉ với một trong số các hộp giới hạn đó được gán nhãn chính xác, dẫn đến nhầm lẫn và không chính xác trong kết quả. Điều này có thể có ý nghĩa quan trọng đối với các ứng dụng trong thế giới thực, đặc biệt là bài toán đếm số lượng.



Mô hình bài toán.

Input: 1 không ảnh. Output: các bounding box tối thiểu có hướng chứa các đối tượng được phát hiện và phân lớp, trong đó mỗi đối tượng chỉ được bao bởi một bounding box với nhãn chính xác.

Để giải quyết vấn đề này, nghiên cứu sử dụng phương pháp hai bước kết hợp các mô hình phát hiện đối tượng có hướng như S2ANet [2], R3Det [1], ReDet [4], RCNN và RoI Transformer [3] với mô hình phân loại, cụ thể là EfficientNet [7]. Các mô hình phát hiện đối tượng có hướng được thiết kế để xử lý các thách thức phát hiện đối tượng trong không ảnh, trong khi mô hình phân loại được sử dụng để gán nhãn lại các đầu ra là hộp giới hạn có hướng và giảm sự nhầm lẫn trong kết quả. Kết quả đầu ra từ cả hai mô hình được tổng hợp và non-maximum suppression được áp dụng để loại bỏ các hộp giới hạn dư thừa hoặc không chính xác.

Nghiên cứu sử dụng bộ dữ liệu DOTA [6] để chứng minh rằng cách tiếp cận hai bước

này có hiệu quả trong việc giảm nhầm lẫn và cải thiện độ chính xác của việc phát hiện đối tượng trong các không ảnh. Bằng cách tận dụng các thế mạnh của các mô hình phân loại và phát hiện đối tượng có hướng, nghiên cứu có thể là một cách tiếp cận mạnh mẽ để cải thiện khả năng phát hiện đối tượng trong các lĩnh vực đầy thách thức như không ảnh.

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

1. Xây dựng được mô hình 2 bước mà có thể linh động thay đổi module phát hiện đối tượng có hướng và module phân lớp.
2. Khảo sát lần lượt các mô hình S2ANet [2], R3Det [1], ReDet [4], Oriented RCNN [5], RoI Transformer [3] làm module phát hiện và EfficientNet [7] làm module phân loại trên bộ dữ liệu DOTA [6].
3. Phân tích được ưu điểm cũng như hạn chế của mô hình 2 bước đã đề ra.

NỘI DUNG VÀ PHƯƠNG PHÁP

Nội dung 1: Nghiên cứu điểm yếu của các phương pháp phát hiện đối tượng có hướng trước đây.

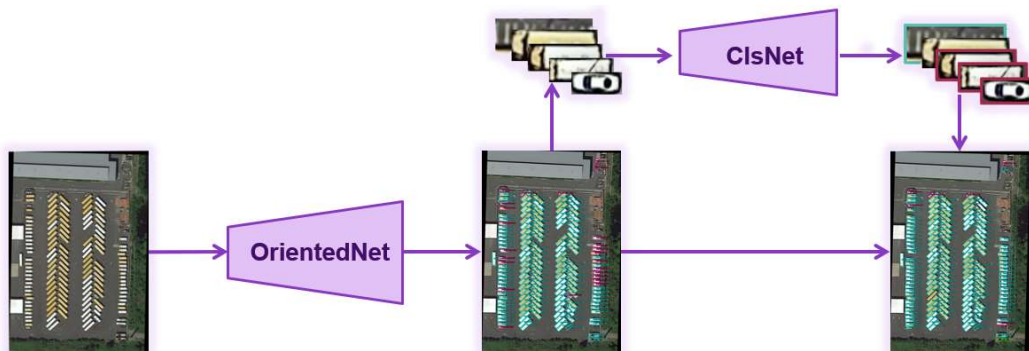
- Nghiên cứu các phương pháp phát hiện đối tượng có hướng S2ANet [2], R3Det [1], Oriented RCNN [5], RoI Transformer [3], ReDet [4]. Do đến nay có rất nhiều phương pháp phát hiện đối tượng có hướng, chúng tôi tiến hành nghiên cứu các phương pháp công bố trong 4 năm trở lại đây và làm cơ sở cho các thử nghiệm cải tiến.
- Thực nghiệm các phương pháp và trực quan hoá kết quả trên bộ dữ liệu DOTA [6] để phân tích được vấn đề gán nhiều hộp giới hạn cho cùng một đối tượng.

Nội dung 2: Nghiên cứu các phương pháp giảm thiểu hộp giới hạn nhầm lẫn trong bài toán phát hiện đối tượng có hướng.

- Nghiên cứu các phương pháp giảm thiểu hộp giới hạn như Loại hộp giới hạn nhầm lẫn bằng Điểm tin cậy, Multi-class Nonmaximum suppression.
- Thử nghiệm các phương pháp giảm thiểu hộp giới hạn nhầm lẫn và rút ra được hạn chế của mỗi phương pháp đó.

Nội dung 3: Đề xuất mô hình 2 bước để giảm thiểu hộp giới hạn nhằm lẫn trong bài toán phát hiện đối tượng có hướng.

- Nghiên cứu và xây dựng mô hình kết hợp mô hình phát hiện đối tượng có hướng và mô hình phân lớp, tăng cường chất lượng phân lớp đầu ra.
- Nghiên cứu phương pháp phân lớp EfficientNet [7].
- Thử nghiệm lần lượt các mô hình S2ANet [2], R3Det [1], ReDet [4], Oriented RCNN [5], RoI Transformer [3] làm module phát hiện và EfficientNet [7] làm module phân loại trên bộ dữ liệu DOTA [6], chứng minh sự hiệu quả của cách tiếp cận 2 bước được đề xuất.



Tổng quan phương pháp 2 bước.

KẾT QUẢ MONG ĐỢI

1. Báo cáo tổng hợp các tác động tích cực của cách tiếp cận 2 bước này trong việc giảm thiểu hộp giới hạn nhằm lẫn trong bài toán phát hiện đối tượng có hướng trong không ảnh.
2. Chương trình demo kết quả của phương pháp.

TÀI LIỆU THAM KHẢO (Định dạng DBLP)

- [1]. Yang, Xue, Junchi Yan, Ziming Feng, and Tao He. "R3det: Refined single-stage detector with feature refinement for rotating object." In Proceedings of the AAAI conference on artificial intelligence, vol. 35, no. 4, pp. 3163-3171. 2021.
- [2]. Han, Jiaming, Jian Ding, Jie Li, and Gui-Song Xia. "Align deep features for oriented object detection." IEEE Transactions on Geoscience and Remote Sensing 60 (2021): 1-11.

- [3]. Ding, Jian, Nan Xue, Yang Long, Gui-Song Xia, and Qikai Lu. "Learning roi transformer for oriented object detection in aerial images." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2849-2858. 2019.
- [4]. Han, Jiaming, Jian Ding, Nan Xue, and Gui-Song Xia. "Redet: A rotation-equivariant detector for aerial object detection." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2786-2795. 2021.
- [5]. Xie, Xingxing, Gong Cheng, Jiabao Wang, Xiwen Yao, and Junwei Han. "Oriented R-CNN for object detection." In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 3520-3529. 2021.
- [6]. Ding, Jian, Nan Xue, Gui-Song Xia, Xiang Bai, Wen Yang, Michael Ying Yang, Serge Belongie et al. "Object detection in aerial images: A large-scale benchmark and challenges." IEEE transactions on pattern analysis and machine intelligence 44, no. 11 (2021): 7778-7796.
- [7]. Tan, Mingxing, and Quoc Le. "Efficientnet: Rethinking model scaling for convolutional neural networks." In International conference on machine learning, pp. 6105-6114. PMLR, 2019.