

THÔNG TIN CHUNG CỦA NHÓM

- Link YouTube video của báo cáo (tối đa 5 phút):

https://youtu.be/y0Sz_Nzz1ik

- Link slides (dạng .pdf đặt trên Github của nhóm):

https://github.com/hoho303/CS519.O11/fsod_devit_slides.pdf

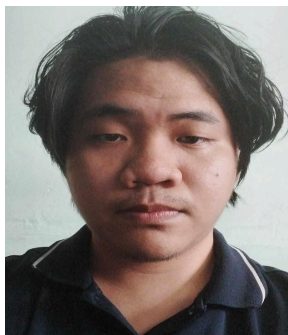
<ul style="list-style-type: none">● Họ và Tên: Nguyễn Hoàng Thuận● MSSV: 21521501 	<ul style="list-style-type: none">● Lớp: CS519.O11● Tự đánh giá (điểm tổng kết môn): 10/10● Số buổi vắng: 3● Số câu hỏi QT cá nhân: 1● Số câu hỏi QT của cả nhóm: ?● Link Github: https://github.com/hoho303/CS519.O11/● Mô tả công việc và đóng góp của cá nhân cho kết quả của nhóm:<ul style="list-style-type: none">○ Lên ý tưởng○ Làm poster○ Làm video YouTube
<ul style="list-style-type: none">● Họ và Tên: Mai Nhất Tùng● MSSV: 21520523	<ul style="list-style-type: none">● Lớp: CS519.O11● Tự đánh giá (điểm tổng kết môn): 10/10● Số buổi vắng: 3● Số câu hỏi QT cá nhân: 1● Số câu hỏi QT của cả nhóm: ?● Link Github:



<https://github.com/hoho303/CS519.O11/>

- Mô tả công việc và đóng góp của cá nhân cho kết quả của nhóm:
 - Lên ý tưởng
 - Viết report
 - Làm slide

- Họ và Tên: Lương Toàn Bách
- MSSV: 21521845



- Lớp: CS519.O11
- Tự đánh giá (điểm tổng kết môn): 10/10
- Số buổi vắng: 3
- Số câu hỏi QT cá nhân: ?
- Số câu hỏi QT của cả nhóm: ?
- Link Github:
<https://github.com/hoho303/CS519.O11/>
- Mô tả công việc và đóng góp của cá nhân cho kết quả của nhóm:
 - Lên ý tưởng
 - Viết report
 - Làm slide

ĐỀ CƯƠNG NGHIÊN CỨU

TÊN ĐỀ TÀI (IN HOA)

PHÁT HIỆN ĐỐI TƯỢNG TRONG ẢNH VỚI ÍT DỮ LIỆU

TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)

FEW-SHOT OBJECT DETECTION

TÓM TẮT *(Tối đa 400 từ)*

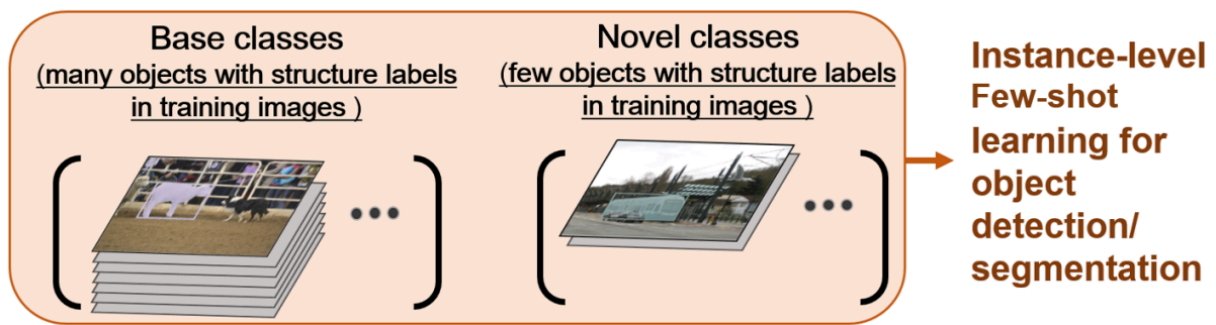
Trong lĩnh vực thị giác máy tính, bên cạnh bài toán phân loại (Classification) và phân đoạn (Segmentation) ảnh thì phát hiện đối tượng (Object Detection) cũng được xem là một trong những vấn đề quan trọng cần được giải quyết, với mục tiêu là phát hiện và xác định các đối tượng có trong hình ảnh hoặc video như: con người, xe cộ, chó, mèo,... Thực tế bài toán này đã được nghiên cứu rất nhiều trong những năm qua, và cũng đã đạt được vô số những tiến bộ đáng kể. Tuy nhiên, đa phần các phương pháp gần đây đều dựa trên cải tiến thông qua mô hình học sâu (Deep Learning), thường đòi hỏi nguồn dữ liệu huấn luyện lớn và đa dạng. Chính vì vậy sẽ gây trở ngại rất lớn khi áp dụng vào trong đời sống, khi mà luôn tồn tại những đối tượng mới lạ (Novel Objects) thường ít hoặc không tồn tại trong những bộ dữ liệu phổ biến, chẳng hạn như các loài động vật quý hiếm, các bệnh hiếm gặp, hoặc các hiện tượng tự nhiên kỳ lạ. Điều này đã dẫn đến sự cần thiết để giải quyết một bài toán khó hơn, gần với thực tế hơn đó là phát hiện đối tượng ít mẫu. Việc phát hiện và nghiên cứu các đối tượng ít mẫu này có ý nghĩa quan trọng trong việc bảo tồn, phát triển và hiểu biết về thế giới xung quanh. Nhưng việc phát hiện và phân tích các đối tượng ít mẫu là một nhiệm vụ khó khăn, đòi hỏi các phương pháp và kỹ thuật nghiên cứu tiên tiến. Do đó, nghiên cứu về lĩnh vực này có thể mang lại những đóng góp quan trọng cho khoa học.

Nội dung sẽ được nghiên cứu: Tìm hiểu tổng quan về bài toán và các hướng tiếp cận chung của bài toán phát hiện đối tượng ít mẫu; chạy thử nghiệm mô hình DE-ViT, ghi nhận và đánh giá kết quả trên các tập dữ liệu ít mẫu phổ biến (Pascal VOC, COCO, LVIS); đề xuất những cải tiến mới để nâng cao độ chính xác cho bài toán phát hiện đối tượng ít mẫu dựa trên mô hình DE-ViT; thử nghiệm triển khai mô hình DE-ViT với các tập dữ liệu khác nhau.

GIỚI THIỆU *(Tối đa 1 trang A4)*

- Trong lĩnh vực thị giác máy tính, bên cạnh bài toán phân loại (Classification) và phân đoạn (Segmentation) ảnh thì phát hiện đối tượng (Object Detection) cũng được xem là một trong những vấn đề quan trọng cần được giải quyết, với mục tiêu là phát hiện và xác định các đối tượng có trong hình ảnh hoặc video như: con người, xe cộ, chó, mèo,... Thực tế bài toán này đã được nghiên cứu rất nhiều trong những năm qua, và cũng đã đạt được vô số những tiến bộ đáng kể. Tuy nhiên, đa phần các phương pháp gần đây đều dựa trên cải tiến thông qua mô hình học sâu (Deep Learning), thường đòi hỏi nguồn dữ liệu huấn luyện lớn và đa dạng. Chính vì vậy sẽ gây trở ngại rất lớn khi áp dụng vào thực tế, khi mà luôn tồn tại những đối tượng mới lạ (Novel Objects), thường ít hoặc không tồn tại trong những bộ dữ liệu phổ biến. Điều này đã dẫn đến sự cần thiết để giải quyết một bài toán khó hơn, gần với thực tế hơn đó là phát hiện đối tượng ít mẫu.

- Trong bài toán phát hiện đối tượng ít mẫu (Few-shot Object Detection), mục tiêu của nó là phát hiện được những đối tượng mới lạ (Novel Objects) khi chỉ tồn tại một số ít các mẫu (instances) cho lớp đối tượng đó dựa trên nguồn dữ liệu đa dạng sẵn có. Tập dữ liệu này có thể được hiểu là tập các đối tượng cơ sở (Base Objects) và không bị hạn chế về số lượng mẫu tồn tại trong các bộ dữ liệu phổ biến.



MỤC TIÊU

- Tìm hiểu tổng quan về bài toán và các hướng tiếp cận hiện nay của bài toán phát hiện đối tượng ít mẫu.
- Đề xuất những cải tiến mới để nâng cao độ chính xác cho bài toán phát hiện đối tượng ít mẫu dựa trên mô hình DE-ViT.
- Áp dụng mô hình DE-ViT vào các bài toán cụ thể trong thực tế.

NỘI DUNG VÀ PHƯƠNG PHÁP

Nội dung 1: Tìm hiểu tổng quan về bài toán và các hướng tiếp cận hiện nay của bài toán phát hiện đối tượng ít mẫu.

Phương pháp nghiên cứu:

- Tìm hiểu tổng quan về bài toán.
- Tìm hiểu tổng quan các hướng tiếp cận tiên tiến chính của bài toán hiện có.
- Tìm hiểu tổng quan các bộ dữ liệu hiện có.

Kết quả dự kiến:

- Báo cáo kết quả tìm hiểu tổng quan
- Báo cáo những vấn đề của các mô hình này ở hiện tại, ưu nhược điểm của các phương pháp đã có

Thời gian dự kiến: 4 tuần

Nội dung 2: Chạy thử nghiệm mô hình DE-ViT, ghi nhận và đánh giá kết quả trên

các tập dữ liệu ít mẫu phổ biến (Pascal VOC, COCO, LVIS).

Phương pháp thực hiện:

- Nghiên cứu kiến trúc và cách hoạt động của mô hình DE-ViT.
- Cài đặt và chạy thử nghiệm mô hình DE-ViT.
 - Ghi nhận và đánh giá kết quả của mô hình
 - Đánh giá ưu điểm, nhược điểm và những điều cần cải tiến của mô hình

Kết quả dự kiến:

- Báo cáo các kết quả nghiên cứu về mô hình DE-ViT.
- Báo cáo phân tích những kết quả sau khi chạy mô hình
- Báo cáo ưu nhược điểm và những điều cần cải tiến của mô hình

Thời gian dự kiến: 6 tuần

Nội dung 3: Đề xuất những cải tiến mới để nâng cao độ chính xác cho bài toán phát hiện đối tượng ít mẫu dựa trên mô hình DE-ViT.

Phương pháp nghiên cứu:

- Nghiên cứu tài liệu, tìm hiểu và lựa chọn các phương pháp nhằm nâng cao độ chính xác cho bài toán phát hiện đối tượng ít mẫu dựa trên mô hình DE-ViT.
- Triển khai những cải tiến mới và chạy thực nghiệm trên các tập dữ liệu phát hiện đối tượng ít mẫu phổ biến (Pascal VOC, COCO, LVIS).

Kết quả dự kiến:

- Báo cáo chi tiết các phương pháp áp dụng để cải tiến mô hình DE-ViT.

Thời gian dự kiến: 10 tuần

Nội dung 4: Thử nghiệm triển khai mô hình DE-ViT với các tập dữ liệu khác nhau.

Phương pháp nghiên cứu:

- Tìm hiểu, lựa chọn hoặc xây dựng những tập dữ liệu khác nhau ngoài những tập dữ liệu phổ biến về bài toán phát hiện đối tượng ít mẫu.
- Thực nghiệm đánh giá mô hình đề xuất trên các bộ dữ liệu đó.

Kết quả dự kiến:

- Bảng so sánh các kết quả đánh giá.

Thời gian dự kiến: 4 tuần

KẾT QUẢ MONG ĐỢI

Tài liệu báo cáo tổng hợp về mô hình phát hiện đối tượng ít mẫu DE-ViT, và độ hiệu quả của mô hình trên các tập dữ liệu phổ biến.

Phương pháp cải tiến mới dựa trên mô hình DE-ViT.

TÀI LIỆU THAM KHẢO (*Định dạng DBLP*)

- [1]. Jiaming Han, Yuqiang Ren, Jian Ding, Ke Yan, Gui-Song Xia: Few-Shot Object Detection via Variational Feature Aggregation. AAAI 2023: 7-2
- [2]. Xiaopeng Yan, Ziliang Chen, Anni Xu, Xiaoxi Wang, Xiaodan Liang, Liang Lin: Meta R-CNN : Towards General Solver for Instance-level Few-shot Learning. IEEE/ICCV 2019: 27-10
- [3]. Xin Wang, Thomas E. Huang, Trevor Darrell, Joseph E. Gonzalez, Fisher Yu: Frustratingly Simple Few-Shot Object Detection. ICML 2020: 13-7
- [4]. Limeng Qiao, Yuxuan Zhao, Zhiyuan Li, Xi Qiu Jianan Wu, Chi Zhang: DeFRCN: Decoupled Faster R-CNN for Few-Shot Object Detection. IEEE/ICCV 2021: 10-10
- [5]. Xinyu Zhang, Yuting Wang & Abdeslam Boularias: Detect Every Thing with Few Examples. 2023