

MÔ HÌNH HỌC TIỆM TIẾN ĐỂ KHAI THÁC TẬP DỮ LIỆU BÊN NGOÀI TRONG BÀI TOÁN NHẬN DIỆN VẬT THỂ

Lê Thanh Minh - 21520063

Tóm tắt

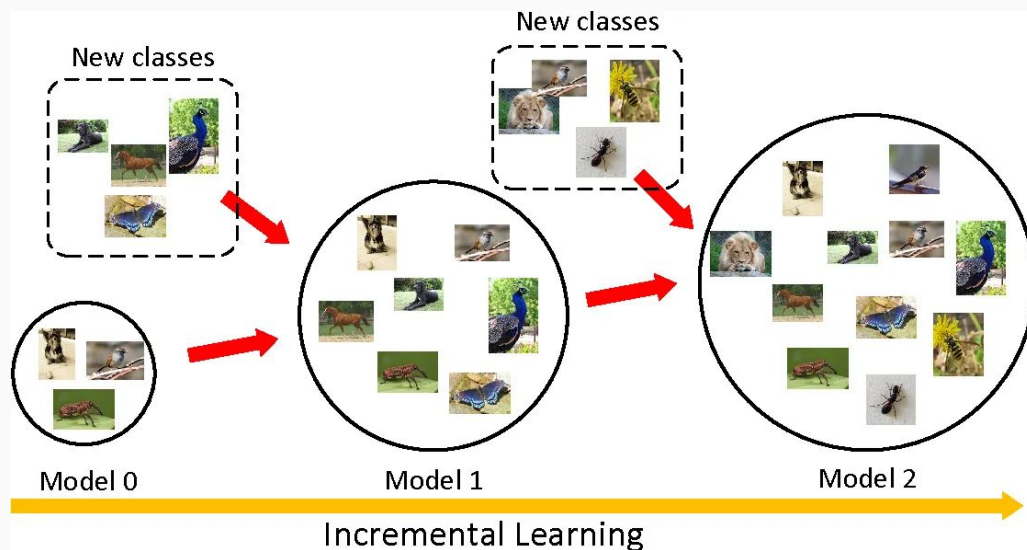
- Lớp: CS519.O11
- Link Github của nhóm:
<https://github.com/Kryst4lize/CS519.O11>
- Link YouTube video:
<https://www.youtube.com/channel/UCvpoCxCfSWmzeNyJ6HEzu-A>



Lê Thanh Minh - 21520063

Giới thiệu

Học tiệm tiến (Incremental Learning) là phương pháp phát triển các hệ thống trí tuệ nhân tạo có khả năng học liên tục để giải quyết các nhiệm vụ mới từ **dữ liệu mới**, đồng thời bảo toàn kiến thức đã học từ các dữ liệu đã biết trước đó.



- Số lớp đối tượng chưa bao quát hết hết
- Các mô hình hiện tại chưa sử dụng dữ liệu bên ngoài

Mục tiêu

- Thu thập dữ liệu và xây dựng một bộ dữ liệu UIT_dataset phù hợp với việc đánh giá mô hình học tiệm tiến
- Xây dựng mô hình tiệm tiến dựa trên kiến trúc Transformer
- Tinh chỉnh mô hình sao cho đạt hiệu suất cao nhất

Nội dung và Phương pháp

Nội dung 1: Thu thập dữ liệu và xây dựng bộ dữ liệu phù hợp cho mô hình học tiệm tiến.

- **Phương pháp thực hiện:**

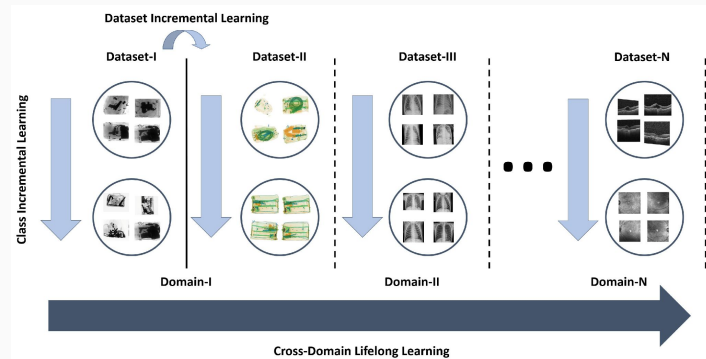
- + Thu thập dữ liệu từ các nguồn uy tín, được sử dụng làm thước đo cho nhiều bài báo khoa học ở các hội nghị uy tín
- + Chia tập dữ liệu này thành hai phần : Base (chứa 50% lớp đối tượng của tập dữ liệu); Extra (Gồm 10 phần nhỏ (Batch), mỗi phần là tập dữ liệu chứa 5% số lớp đối tượng của tập dữ liệu đó). Các class trong phần Base và các Batch trong phần Extra không trùng nhau.

Nội dung và Phương pháp

Nội dung 2: Xây dựng mô hình tiệm tiến dựa trên kiến trúc Transformer và tinh chỉnh mô hình

- Phương pháp thực hiện

- + Chúng ta trích xuất đặc trưng của ảnh cần huấn luyện thông qua mạng pre-trained CNN
- + Đối với phần dữ liệu Base, chúng ta mong muốn huấn luyện trên mô hình Transformer để tạo ra một encoder có thể biểu diễn được đặc trưng của các lớp phần Base
- + Đối với phần Extra, mỗi Batch sẽ được đưa vào encoder để tạo ra một token mới, so sánh chúng đối với token cũ từ phần Base, từ đó so sánh được sự khác biệt giữa các đặc trưng của các lớp đối tượng mới để điều chỉnh lại encoder sao cho có thể học được các lớp mới và giữ nguyên các lớp cũ
- + Tinh chỉnh tham số sao cho hiệu suất của mô hình tốt nhất



Kết quả dự kiến

- Xây dựng một mô hình học tiệm tiến có có kết quả cạnh tranh về mặt hiệu suất hoặc thời gian hơn so với các phương pháp SotA trước đó (State of the Art) với độ chính xác hơn từ 5-10%
- Kết quả thực nghiệm, đánh giá, so sánh so với các phương pháp khác
- Kết quả sau khi đánh giá đủ tốt và mức độ đóng góp lớn để có thể được chấp thuận tại các hội nghị lớn như CVPR2024 sắp tới

Tài liệu tham khảo

- [1]. Alex Krizhevsky: Learning Multiple Layers of Features from Tiny Images, 2009
- [2]. Olga Russakovsky, *Jia Deng*, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg and Li Fei-Fei. ImageNet Large Scale Visual Recognition Challenge. IJCV, 2015
- [3]. Hou, Saihui and Pan, Xinyu and Loy, Chen Change and Wang, Zilei and Lin, Dahua, “Learning a unified classifier incrementally via rebalancing,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019
- [4]. Yan, Shipeng, Jiangwei Xie, and Xuming He. "DER: Dynamically Expandable Representation for Class Incremental Learning.", *CPVR*, 2021. /abs/2103.16788.
- [5]. A. Vaswani, N. Shazeer, N. Parmar, et al., “Attention is all you need,” CoRR, vol. abs/1706.03762, 2017. arXiv: 1706.03762. [Online]. Available: <http://arxiv.org/abs/1706.03762>.
- [6]. K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings, Y. Bengio and Y. LeCun, Eds., 2015. [Online]. Available: <http://arxiv.org/abs/1409.1556> .
- [7]. K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” CoRR, vol. abs/1512.03385, 2015. arXiv: 1512.03385. [Online]. Available: <http://arxiv.org/abs/1512.03385> .
- [8]. C. Szegedy, W. Liu, Y. Jia, et al., “Going deeper with convolutions,” CoRR, vol. abs/1409.4842, 2014. arXiv: 1409 . 4842. [Online]. Available: <http://arxiv.org/abs/1409.4842>
- [9]. K. O’Shea and R. Nash, “An introduction to convolutional neural networks,” CoRR, vol. abs/1511.08458, 2015. arXiv: 1511.08458. [Online].