

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

Lê Thanh Minh^{1,2}

¹ Trường ĐH công nghệ thông tin, ĐHQG Hồ Chí Minh, Việt Nam

² 21520063@gm.uit.edu.vn

Tóm tắt

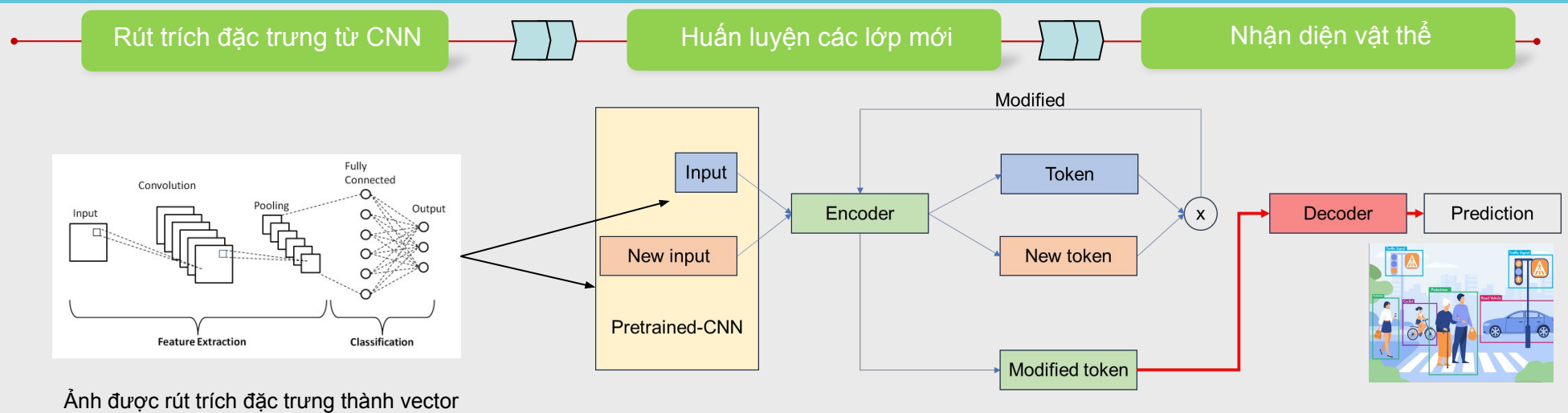
Chúng tôi đề xuất một mô hình có thể khai thác dữ liệu bên ngoài để mô hình có thể mở rộng, bao gồm:

- Xây dựng mô hình tiệm tiến có thể khai thác nhiều dữ liệu bên ngoài dựa trên kiến trúc Transformer
- Xây dựng bộ dữ liệu phù hợp dựa trên các bộ dữ liệu phổ biến để có thể đánh giá hiệu suất của mô hình
- Sử dụng mô hình đã huấn luyện sẵn thuộc họ CNN để trích xuất đặc trưng đầu vào cho mô hình

Động lực ?

- Hầu hết các hướng nghiên cứu gần đây chỉ tập trung vào cải thiện độ chính xác trên tập dữ liệu cụ thể, nhưng chưa tính đến việc cải thiện mô hình dựa vào lượng dữ liệu bên ngoài
- Số lớp đối tượng trong một tập dữ liệu cụ thể đã giới hạn ở số lượng đối tượng gắn nhãn, không bao quát hết được số lượng đối tượng có trong thực tế

Tổng quát



Mô Tả

1. Xây dựng bộ dữ liệu

- Thu thập dữ liệu từ các nguồn uy tín, được sử dụng phổ biến cho nhiều bài báo khoa học ở các hội nghị lớn
- Chia tập dữ liệu này thành hai phần : Base (chứa 50% lớp đối tượng của tập dữ liệu) và Extra (Gồm 10 phần nhỏ là Batch, mỗi Batch 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.

Ví dụ: Tập dữ liệu ImagetNet100 có tổng cộng 100 class, được chia làm 2 phần Base (50 class) và Extra (10 tập nhỏ, mỗi tập chứa 5 class)

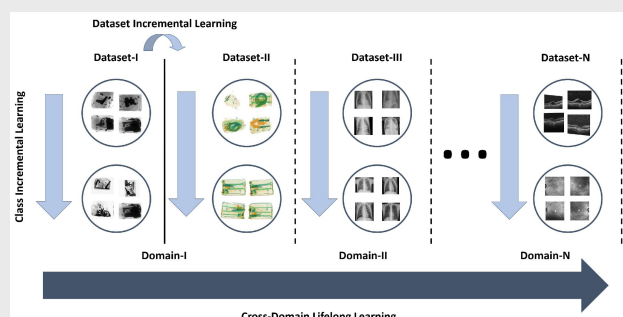


Figure 1. Cách từng batch được sử dụng trong quá trình học tiệm tiến

2. Xây dựng mô hình tiệm tiến

- Chúng ta trích xuất đặc trưng của một ảnh từ mạng CNN
- Huấn luyện phần dữ liệu Base để Encoder học đặc trưng
- Với từng tập Batch của tập dữ liệu Extra, một token của tập batch được tạo ra dựa trên encoder đã học cũ từ phần Base, từ đó so sánh được sự khác biệt giữa các lớp đối tượng mới để và điều chỉnh lại Encoder
- Cuối cùng, tinh chỉnh lại mô hình về phần rút trích đặc trưng và số lượng đặc trưng mà Encoder có thể tạo để mô hình có hiệu suất tốt nhất

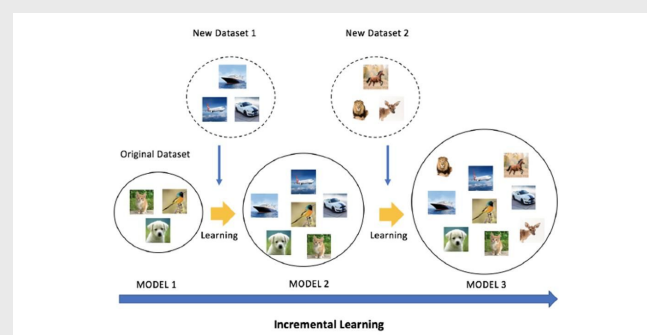


Figure 2. Pipeline huấn luyện của mô hình tiệm tiến