ÁP DỤNG TRANSFER LEARNING VỚI MẠNG VGG16 CHO CUỘC THI DIGIT RECOGNITION

Đinh Xuân Vũ – 18521662

1. **PHƯƠNG PHÁP SỬ DỤNG**

Như ta đã biết với bài toán nhận dạng chữ viết viết tay là một bài toán classification với 10 class khác nhau cho từng chữ số từ 0 đến 9. Bộ dataset được sử dụng là MNIST hand written.

Thay vì phải xây dựng một mô hình học sâu và huấn luyện lại từ đầu để có thể nhận biết được những đặc trưng và thực hiện phân loại thì ta có thể sử dụng những mô hình có sẵn và tinh chỉnh nó cho phù hợp với bài toán mà ta đang giải quyết và ta gọi đó là Transfer Learning.

Tuy nhiên ta cần phải xem xét đến những gì mà ta sử dụng để có thể giải quyết bài toán này. Đầu tiên là pretrained model VGG16 trên dataset ImageNet có kích thước là 224x224x3 (như hình dưới) trong khi dataset của ta chỉ là 28x28x1. Tiếp theo softmax layer có một vector 1000x1 chiều để giải quyết cho bài toán phân loại 1000 lớp trong khi ta chỉ cần vector 10x1 để giải quyết. Vì thế ta có thể sử dụng VGG16 để extract feature từ dữ liệu và đưa nó vào một mô hình đơn giản hơn để giải quyết bài toán và ta gọi đó là feature extraction.

Diagram

Description automatically generated

1. **CHI TIẾT CÀI ĐẶT**

Vấn đề đầu tiên ta gặp phải là VGG16 cần feed image vào với kích thước tối thiểu 32x32x3 (bức ảnh RGB với kích thước 32x32) => Resize và stack image thành (48x48x3) để kích thước phù hợp với VGG16 và đồng thời để kích thước các tensor qua từng lớp Convolutional đủ lớn.

Tiếp theo, ta cần đóng băng (freeze) mô hình VGG16, cắt bỏ các layer Fully-Connected (hình bên dưới) và feed các bức ảnh đã được chỉnh lại phù hợp với VGG16 và tiến hành feature extraction.

Diagram

Description automatically generated

Sau khi đã thực hiện feature extraction, ta lại xây dựng một mạng Neural Network đơn giản để nhận tensor 1x1x512 vào và cho vào Fully-Connected với 10 neurons để thực hiện classification. Hình bên dưới thể hiện Neural Network đơn giản đó.

A group of yellow and red rectangles with black text

Description automatically generated with low confidence

Vậy là ta đã xong, bước tiếp theo chỉ cần huấn luyện mô hình và đánh giá kết quả.

1. **ĐÁNH GIÁ KẾT QUẢ**

Như biểu đồ ở dưới ta có thể thấy, chỉ cần qua 5 epoch ta đã đạt được hội tựu (convergence) và accuracy rất cao với thời gian huấn luyện rất ngắn.

Qua đây ta có thể thấy được ưu điểm của phương pháp này là ta có thể không cần phải huấn lụyện lại mô hình một cách hoàn toàn mà chỉ cần huấn luyện một phần nên thời gian huấn luyện cũng giảm đi rất nhiều.

Chart, line chart

Description automatically generated

1. **KẾT QUẢ TRÊN BẢNG XẾP HẠNG:**

Đứng 32/163 với độ chính xác là 0. 97952Graphical user interface, application

Description automatically generated