**BỘ GIÁO DỤC VÀ ĐÀO TẠO**

**TRƯỜNG ĐẠI HỌC CẦN THƠ**

**KHOA CÔNG NGHỆ THÔNG TIN & TRUYỀN THÔNG**

**🙞 🕮 🙜**



**BÁO CÁO BÀI TẬP**

**MÁY HỌC NÂNG CAO**

**Đề tài**

**NHẬN DIỆN GƯƠNG MẶT NGƯỜI**

|  |  |
| --- | --- |
| **SV Thực Hiện**:  Tân  Thành  Tuyền  Nguyễn Hồng Phát (B1509942) | **GV Hướng Dẫn**:  PGS. Phạm Nguyên Khang |

**NHẬN XÉT CỦA GIẢNG VIÊN**

**----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------**

**----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------**

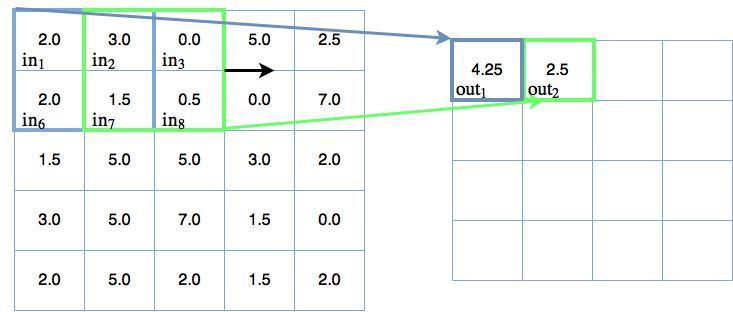
**Phần I: Giới Thiệu**

1. **Bài Toán**:

Các mô hình mạng nơ-ron của Học sâu (Deep learning) đang được áp dụng rộng rãi vào nhiều mặt của cuộc sống, từ những mô hình phức tạp được áp dụng vào các sản phẩm chỉ được nghiên cứu và phát triển tại các công ty công nghệ lớn: Alex Assistant (Amazon), Alpha Go (Google),... cho đến những ứng dụng đơn giản như nhận dạng và phân loại mà chúng ta vẫn tận dụng hàng ngày. Để bắt kiệp với tốc độ phát triển đó, trong bài báo cáo này chúng em sẽ ứng dụng mạng nơ-ron tích chập (CNN) để giải quyết một bài toán cơ sở về nhận dạng gương mặt người.

1. **Mạng Nơ-ron tích chập**:

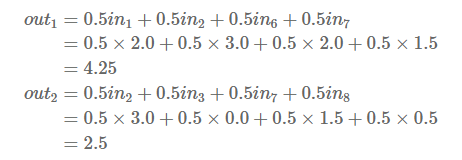
Mạng nơ-ron đa tầng (MPL) có thể cho được kết quả tốt trong việc dự đoán chữ số từ tập dữ liệu MNIST. Nhưng thực tế, dữ liệu từ tập MNIST là đơn giản (ảnh xám, độ phân giải không cao – 28x28… Do đó, khi tiến hành phân loại ảnh trên các tập dữ liệu phức tạp hơn (ảnh màu, ảnh chứa các đối tượng phức tạp hơn chữ số: xe, con vật…) thì phải cần đến mô hình mạng nơ-ron tích chập (hay **CNN**).



*Hình 1*. Tích chập của mặt nạ 2x2.

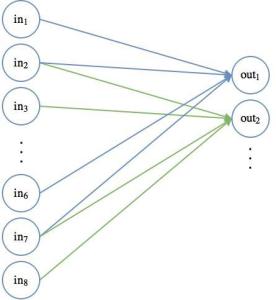
Trái: dữ liệu ảnh ban đầu. Phải dữ liệu ảnh kết quả - của 2 bước tích chập.

Ý tưởng chính của CNNs là dịch chuyển (hay tích chập) một mặt nạ - filter trên dữ liệu ảnh đầu vào – một ma trận 2D (xem Hình 1), mỗi lần tích chập mặt nạ sẽ áp dụng một trọng số (weight) lên dữ liệu để tính toán kết quả. Cụ thể các tính toán dưới đây (xem Hình 2) đã được thực hiện để có được kết quả ở Hình 1.



*Hình 2. Các tính toán của 2 bước tích chập mặt nạ.*

Để tính toán tất cả kết quả, các bước tích chập sẽ được thực hiện trên tất cả các điểm tọa độ (x, y) có thể của ma trận dữ liệu ảnh đầu vào.



*Hình 3. CNN dưới gốc độ MLP.*

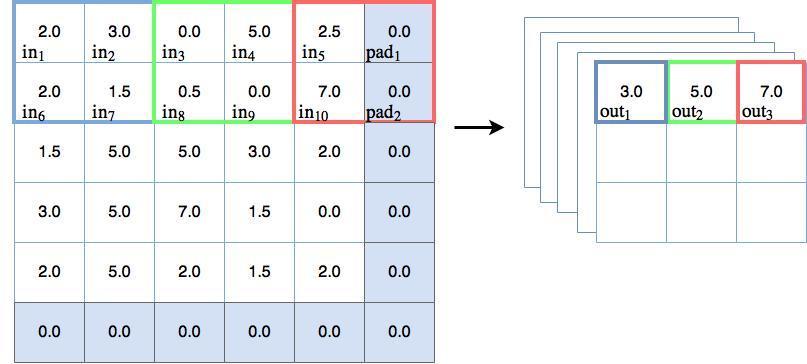
CNNs cũng có thể được biểu diễn ở gốc độ của một mạng MLP (xem Hình 3), tuy nhiên CNNs có các đặc trưng sau so với một MLP:

* Không giống với mạng MLP, không phải tất cả các nút (node) ở tầng đầu vào (input layer) được kết nối với tầng đầu ra (output layer).
* Trọng số (weights) của mặt nạ (filter) là các hằng số không đổi trong lúc tiến hành tích chập trên cùng một ảnh.

**Pooling** là một khái niệm quan rọng được sử dụng trong CNNs. Hai mục tiêu của Pooling là:

* Giảm số lượng trọng số phải huấn luyện của mạng (parameters).
* Cải thiện khả năng phát hiện đặc trưng (feature detection) bằng cách giảm tác động từ kích thước (scale) và hướng (orientation) của dữ liệu ảnh.

Về nguyên lý, Pooling cũng là một kĩ thuật “cửa sổ trược” (sliding window) giống như tích chập mặt nạ. Nhưng thay vì áp dụng một trọng số (weight) để tính toán kết quả, Pooling áp dụng một phép toán thống kê (max, mean, L2,…).

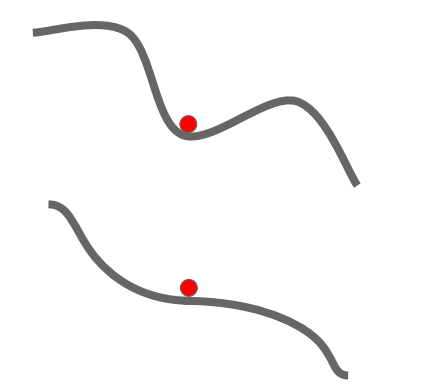


*Hình 4. Max Pooling của Stride 2..*

Trong ví dụ của hình trên, để tránh hiện tượng tính toán lại trên các khu vực đã được trượt qua của một ảnh (như ở phần tích chập mặt nạ), trong Max Pooling có một khái niệm gọi là Stride và giá trị của Stride được đặt sao cho không xảy ra hiện tượng lặp (overlap) vừa nêu (xem Hình 2). Việc đặt giá trị của Stride như vậy cũng đồng thời giúp giảm chiều của ma trận kết quả do đó giảm được số lượng trọng số của mạng.

**Tầng liên kết đầy** (fully connected layer) thường là tầng cuối trong một mô hình mạng nơ-ron có vài trò phân loại các đối tượng quan tâm của bài toán.

Thuật toán cập nhật trọng số **Adam** (Adam optimizer) là một sự kết hợp của hai giải thuật Momentum optimization và RMSProp. Adam giúp cho quá trình huấn luyện thoát được nơi có gradient rất bé như local optima và saddle point (xem Hình 5) nhờ vào ưu điểm của Moemtum optimization.



*Hình 5. Local optima (trên) và Saddle Point (dưới).*

Ngoài ra Adam là một giải thuật Adaptive Learning Rate như ưu điểm của RMSProp, điều này có nghĩa mỗi trọng số (weights) trong mô hình mạng sẽ có một tốc độ học (learning rate) tương ứng, khác với giải thuật SGD truyền thống chỉ có một tốc độ học cho tất cả các trọng số. Đồng thời, tốc độ học cũng được điều chỉnh lại tương ứng với các trọng số (các trọng số có giá trị lớn sẽ có tốc độ học nhỏ hơn), điều này giúp thúc đẩy quá trình huấn luyện (đi đúng hướng) để tìm được đến điểm tối ưu nhanh hơn (global optima).

**Phần II: Cài Đặt Và Huấn Luyện**

1. **Cài đặt.**
2. **Huấn luyện.**

**Phần II: Kết Luận**