

ĐẠI HỌC QUỐC GIA TP.HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC BÁCH KHOA
KHOA KHOA HỌC VÀ KỸ THUẬT MÁY TÍNH

—o0o—



Luận văn tốt nghiệp

Phát hiện kẹt xe từ camera hành trình

GVHD: TS.Dương Ngọc Hiếu

Nhóm sinh viên thực hiện:

Trương Ngọc Anh 1410141

Lê Nguyễn Minh Trí 1414207

Mục lục

1	Giới thiệu	8
1.1	Tổng quan	8
1.1.1	Tính khả quan:	11
1.1.2	Ưu điểm:	12
1.1.3	Hạn Chế:	12
1.2	Mục tiêu đề tài	12
2	Mạng neural và Mạng neural tích chập (Convolutional Neural Networks CNN	14
2.1	Mạng neural (Neural Network)	14
2.1.1	Mạng neural (Neural Network)	14
2.1.2	Cấu trúc mạng neural	15
2.1.3	Mô hình Feedforward Neural Network	19
2.2	Mạng neural tích chập (Convolutional neural network)	23
2.2.1	Giới thiệu mạng neural tích chập	23
2.2.2	Mô hình mạng neural tích chập	23
2.2.3	Kiến trúc mạng GoogLeNet)	31

3	Apache Hadoop	34
3.1	Tổng quan về hệ thống Apache Hadoop	34
3.2	Hadoop distributed file system - HDFS	35
3.2.1	Thiết kế	35
3.2.2	Các khái niệm trong HDFS	35
3.2.3	HIVE	37
3.3	Tại sao sử dụng Apache Hadoop trong đề tài	39
4	Mô hình loại ảnh giao thông kết hợp Apache Hadoop	40
4.1	Tổng quan các bước	41
4.2	Mô tả tập dữ liệu	41
4.3	Phương pháp transfer learning	41
4.4	Các bước chuẩn bị	42
4.4.1	Chuẩn bị môi trường hiện thực	42
4.5	Các bước hiện thực	44
4.5.1	Cài đặt Apache Hadoop	44
4.5.2	Cài đặt HIVE	52
4.5.3	Xử lý dữ liệu	54
4.5.4	Tạo các bottlenecks	55
4.5.5	Huấn luyện	55
5	Kết quả huấn luyện và đánh giá	57
5.1	Kết quả huấn luyện	57
5.2	Sử dụng mô hình phân loại ảnh mới	59
6	Kết luận	64
6.1	Các kết quả	64

6.2	Hướng phát triển	65
-----	----------------------------	----

Danh sách hình vẽ

2.1	Tế bào thần kinh neuron sinh học	15
2.2	Cấu trúc cơ bản mạng neuron	16
2.3	Đồ thị hàm step	18
2.4	Đồ thị hàm sigmoid	18
2.5	Đồ thị hàm Tanh	19
2.6	Ví dụ kiến trúc mạng feedforward	20
2.7	MLP	21
2.8	Convolutional layer	25
2.9	Convolutional layer	25
2.10	Mảng các giá trị của bộ lọc	26
2.11	ảnh đầu vào	27
2.12	phép toán tích chập	27
2.13	phép toán tích chập	28
2.14	kết quả tầng tích chập	28
2.15	max-pooling 2 x 2	29
2.16	Ví dụ tầng tổng hợp	30
2.17	GoogLeNet	31
2.18	inception module	32

2.19	naive inception module	32
4.1	Mô hình phân loại ảnh giao thông kết hợp Hadoop	40
4.2	Kiểm tra Hadoop service trên master	50
4.3	Kiểm tra Hadoop service trên worker(slave)	51
4.4	Kiểm tra Hadoop service trên worker(slave1)	51
4.5	Liệt kê thư mục dữ liệu trên HDFS	52
5.1	Đồ thị biểu diễn độ chính xác (accuracy)	58
5.2	Đồ thị biểu diễn độ sai số (cross entropy)	59
5.3	Ảnh kiểm thử 1 - kết quả	60
5.4	Ảnh kiểm thử 2 - kết quả	61
5.5	Ảnh kiểm thử 3 - kết quả	62
5.6	Ảnh kiểm thử 4 - kết quả	63

Danh sách bảng

4.1	Các chiều dữ liệu.	43
4.2	Một vài kiểu dữ liệu.	43

Lời cảm ơn

Trước khi vào phần giới thiệu nội dung đề tài luận văn, nhóm sinh viên tụi em xin gửi lời cảm ơn chân thành và sự tri ân sâu sắc đối với thầy **Dương Ngọc Hiếu** đã giúp đỡ, hướng dẫn tận tình cho tụi em về những khái niệm căn bản và phương hướng tiếp cận thông qua việc phân tích ảnh được dùng trên công nghệ Deep Learning, giúp tụi em chỉnh sửa những thiếu sót trong quá trình nghiên cứu và thực hiện đề tài này. Ngoài ra, tụi em cũng xin chân thành cảm ơn anh **Nguyễn Thanh Trông** đã hướng dẫn, giúp đỡ tụi em về việc ứng dụng & kết hợp Deep Learning với Big Data, bổ sung những thiếu sót, sai sót của tụi em trong quá trình thực hiện đề tài. Đồng thời, do những kinh nghiệm thực tiễn và phương hướng tiếp cận về công nghệ Deep Learning & Big Data còn nhiều hạn chế nên bài luận văn không thể tránh khỏi những thiếu sót, tụi em rất mong sẽ nhận được ý kiến đóng góp từ phía Thầy, Cô để tụi em học hỏi thêm được nhiều kinh nghiệm và sẽ hoàn thành tốt đề tài luận văn này hơn nữa !

Tụi em xin chân thành cảm ơn !

Chương 1

Giới thiệu

1.1 Tổng quan

Từ những năm gần đây, tình trạng ùn tắc giao thông ở Thành Phố Hồ Chí Minh đã không dừng lại ở diễn biến phức tạp mà lại còn gia tăng hơn trước. Cụ thể, trên các tuyến đường giờ đây, không chỉ riêng giờ cao điểm, mà các khung giờ khác, tình trạng kẹt xe ngày càng trở phổ biến đối với mỗi người khi tham gia giao thông. Nguyên nhân dẫn đến sự việc trên, một phần ảnh hưởng bởi điều kiện thời tiết, một phần do các công trình cải tạo hạ tầng, khi thi công lấn chiếm mặt đường. Nhưng phần lớn là do mật độ xe cộ ngày một đông dần, dẫn đến việc ùn tắc, ùn ứ, di chuyển chậm,... Từ những nguyên nhân trên, thực trạng giao thông ở Tp. Hồ Chí Minh trở thành một bài toán mà các công ty, các nhà công nghệ hàng đầu Việt Nam quan tâm đến.

Hiện nay, để giải quyết những bài toán như trên, đã có một số ý kiến về việc xây dựng một hệ thống "Smart City", "Hệ thống giao thông thông

minh",... Những hệ thống trên được đưa ra dựa vào sự bùng nổ kỹ thuật số và thông tin. Dữ liệu trở nên một cách đa dạng, không còn là những file sql, text thông thường nữa, mà đã được nâng cấp, cải tiến thành hình ảnh, video,... giúp cho chúng ta có thể nhìn được dữ liệu một cách trực quan, phong phú hơn, nhưng để đáp ứng được nhu cầu thiết yếu ấy, chúng ta cần một lượng tập dữ liệu cực lớn để lưu trữ, và giờ đây nó đã trở thành một thuật ngữ mà khi nhắc đến ai cũng có thể biết đến đó là "Big Data".

"Big Data" có những lợi ích như: cắt giảm chi phí, tăng thời gian phát triển, tối ưu hoá sản phẩm và giảm thời gian. Giúp chúng ta có thể sử dụng, phân phối một lượng dữ liệu lớn một cách linh hoạt, hiệu quả, đạt được hiệu suất cao,... Ngoài ra, "Big Data" còn được ứng dụng vào các dự án công nghệ, thông qua việc thu thập dữ liệu trong suốt nhiều năm để dự đoán về một quá trình nào đó có thể xảy ra hay không,...

Một số công nghệ sử dụng "Big Data": Hadoop, Apache Spark, Data lakes, NQL Databases, In-memory databases,...

Bên cạnh với sự bùng nổ dữ liệu đến từ "Big Data", thì với những năm gần đây, trí tuệ nhân tạo (AI) nổi lên như một bằng chứng của cuộc cách mạng công nghiệp lần thứ tư (1-động cơ hơi nước, 2-năng lượng điện, 3-công nghệ thông tin). Trí tuệ nhân tạo đã và đang trở thành phần cốt lõi trong các hệ thống công nghệ cao. Nó đã len lỏi vào hầu hết các lĩnh vực trong đời sống mà có thể chúng ta không nhận ra. Xe tự hành của Google và Tesla, hệ thống tự tag khuôn mặt trong ảnh của Facebook, trợ lý ảo Siri của Apple,...

Học máy (Machine Learning) là một tập con của Trí Tuệ Nhân Tạo. Nó là một lĩnh vực nhỏ trong khoa học máy tính, có khả năng tự học hỏi dựa trên dữ liệu được đưa vào mà không cần phải được lập trình cụ thể

(Machine Learning is the subfield of computer science, that "gives computers the ability to learn without being explicitly programmed"—Wikipedia).

Những năm gần đây, sự phát triển của các hệ thống tính toán cùng với lượng dữ liệu khổng lồ được thu thập bởi các hãng công nghệ lớn đã giúp machine learning tiến thêm một bước dài. Một lĩnh vực mới được ra đời được gọi là Học sâu (Deep Learning). Deep Learning đã giúp máy tính thực thi những việc tưởng chừng như không thể vào những năm trước đây: phân loại hàng ngàn vật thể khác nhau trong bức ảnh, tự tạo chú thích cho ảnh, bắt chước giọng nói và chữ viết của con người, giao tiếp với con người, chuyển đổi ngôn ngữ,...

Deep Learning được nghiên cứu và được chú tâm nhiều ở các doanh nghiệp vì Deep Learning đã khai thác được Big Data cùng với độ chính xác cao trên tập dữ liệu ảnh và tín hiệu số cao, tạo ra nhiều sự quan tâm đến mọi người nghiên cứu. Việc áp dụng sự kết hợp Deep Learning và Big Data trở nên thành việc khả thi và cần thiết đối với các dữ liệu cần được lưu trữ có dung lượng lớn, nhưng điều đó trở nên khó khăn đối với một số đối tượng nghiên cứu, tại vì chúng ta cần xây dựng một hệ thống Big Data, có thể điều đó tốn khá nhiều chi phí để thực thi, tuy nhiên, nó lại khiến cho kết quả của độ chính xác sẽ đạt được có thể sẽ cao hơn so với dự kiến của chúng ta. Các nhà công nghệ lớn cũng đã áp dụng sự kết hợp đó như Facebook, Google,...

Từ những ưu điểm, nhược điểm, đã nêu trên cùng với mong muốn áp dụng công nghệ Deep Learning vào để giải quyết những khó khăn, hạn chế trong thực tế đời sống. Nhóm sinh viên tại em đã có một ý tưởng về việc xác định việc kẹt xe thông qua tấm hình được cắt ra từ camera hành trình của xe bus. Trên một chiếc xe bus thông thường sẽ có 4 camera: 2 camera trong

xe, 1 camera trước xe, và 1 camera sau xe, đối với mỗi camera trước và sau, khi xe bus di chuyển từ trạm A đến một trạm B. Trên tuyến đường đó, xe bus sẽ quay được hình ảnh giao thông trên đường và xuất ra thành một file video. Sau đó, chúng ta sẽ cắt từng bức ảnh từ video đó theo một thời gian cố định và đưa vào công nghệ Deep Learning xử lý. Kết quả sau cùng sẽ là xác định được con đường này đang "KẸT XE" hay là "THÔNG THOÁNG".

Từ ý tưởng ban đầu, cơ bản chỉ là phân tích và nêu lên kết quả trên một bức ảnh, tụi em đã mở rộng ra bằng việc dùng "Big Data" để lưu trữ dữ liệu, bao gồm: các file SQL và video lấy về từ xe bus. Đối với file SQL, có các thuộc tính như: Kinh độ, Vĩ độ, Thời gian,... và video từ camera hành trình được lấy từ xe bus. Từ đó, tụi em sẽ xác định được ở khoảng thời gian nào, ở vị trí nào, con đường đó có kẹt xe hay không và hiển thị trên bản đồ giao thông ở Tp. Hồ Chí Minh, giúp giảm tải việc kẹt xe ở các tuyến đường vào những giờ cao điểm, góp một phần tìm ra lời giải cho bài toán giao thông ở thành phố ta.

1.1.1 Tính khả quan:

Để đánh giá về tính khả quan trong đề tài này, chúng ta cần phân tích một số đặc điểm. Thứ nhất là về tập ảnh giao thông, ở đây, tập ảnh giao thông tốt hay không còn tùy thuộc vào nhiều yếu tố như nguồn gốc tập ảnh hay cơ cấu cơ sở hạ tầng camera giao thông được đầu tư như thế nào. Đối với thành phố Hồ Chí Minh riêng, chúng ta vẫn còn đang xây dựng hệ thống cơ sở hạ tầng giám sát nên chưa thể dựa vào nguồn camera giám sát để lấy dữ liệu. Vì thế, chúng em sẽ sử dụng tập ảnh từ camera hành trình trên các xe buýt để thực hiện xây dựng mô hình. Khi đã có tập dữ liệu đủ

tốt, chúng ta có thể sử dụng mô hình được xây dựng trên đề tài này và có thể áp dụng vào hệ thống thực tiễn.

1.1.2 Ưu điểm:

- Hệ thống file được lưu trữ bằng cách dùng công nghệ Big Data - Hadoop.

1.1.3 Hạn Chế:

- Ảnh lấy từ Camera hành trình còn có nhiều hạn chế.
- Dòng Stream lấy từ Video chưa thực hiện một cách triệt để.

1.2 Mục tiêu đề tài

Đối với đề tài này, tìm hiểu các kiến thức về học sâu, mạng neuron và các khái niệm cơ bản, xây dựng được mô hình có khả năng phân loại hình ảnh giao thông cũng như xây dựng một mô hình HDFS sử dụng Apache Hadoop để kết hợp với việc huấn luyện bài toán phân loại ảnh.

Các công việc thực hiện:

- Xây dựng được bộ dữ liệu phục vụ cho việc huấn luyện.
- Xây dựng một mô hình HDFS lưu trữ dữ liệu video và dữ liệu ảnh huấn luyện.
- Cấu hình, sử dụng kiến trúc mạng có sẵn để huấn luyện.
- Phân tích được tình trạng giao thông qua một bức ảnh.

- Xác định được thời gian, kinh độ, vĩ độ, gắn với bức ảnh được trích ra từ camera.

Chương 2

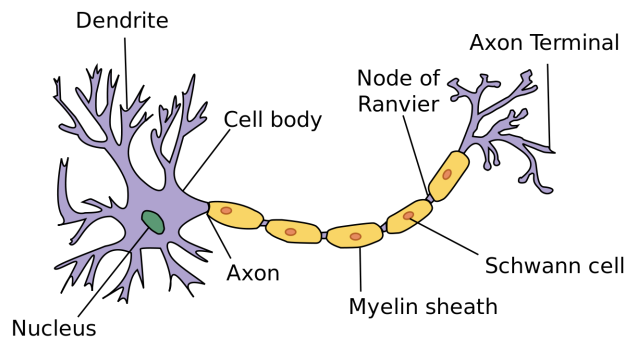
Mạng neural và Mạng neural tích chập (Convolutional Neural Networks CNN

2.1 Mạng neural (Neural Network)

2.1.1 Mạng neural (Neural Network)

Theo khái niệm về sinh học, mạng neural là sự kết nối giữa các tế bào thần kinh neural lại với nhau. Trong lĩnh vực trí tuệ nhân tạo, mạng neural còn được gọi là Artificial Neural Network (ANN) - mạng neural nhân tạo, đây là mô hình xử lý dữ liệu, mô phỏng lại chức năng và cách hoạt động của hệ thống neural sinh học ở con người. Hình [2.1](#) minh họa cấu trúc của tế bào thần kinh neuron.

Mạng neural có gồm nhiều đơn vị kết nối, làm việc như một thể thống



Hình 2.1: Tế bào thần kinh neuron sinh học

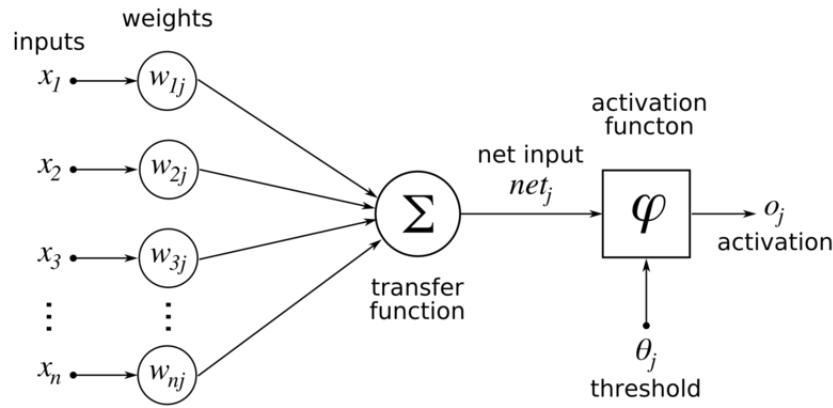
nhất thông qua việc trao đổi thông tin nhờ các liên kết.

2.1.2 Cấu trúc mạng neural

Như đã trình bày, các neuron trong một mạng làm việc như một thể thống nhất bằng việc trao đổi thông tin. Thực tế, đây là quá trình điều chỉnh các trọng số được truyền từ input ban đầu kết hợp với các hàm tính toán để có được các thông số trọng số phù hợp nhất. Quá trình này còn được gọi là quá trình học hay huấn luyện. Hình 2.2 mô tả cấu trúc đơn giản nhất của một mạng neuron.

Cấu trúc mạng neuron

- **Tập các node:** bao gồm nhiều node, mỗi node là đơn vị nhỏ nhất giữ chức năng xử lý thông tin của mạng.
- **Các tầng:** Các node trên được xếp thành các tầng, các node chung một tầng không thể kết nối nhau. Trong đó tầng input và tầng output là 2 tầng thiết yếu. Tùy vào một số mạng cụ thể có thể có thêm một hay nhiều tầng nằm ở giữa được gọi là tầng ẩn (hidden layer).



Hình 2.2: Cấu trúc cơ bản mạng neuron

- **Tầng input - input layer:** các nối ở tầng này nhận dữ liệu đầu vào và truyền tới các node ở các tầng kế tiếp, trong một số trường hợp còn có chức năng xử lý thông tin.
- **Tầng ẩn - hidden layer:** một số mạng có thể có thêm tầng ẩn, số lượng tầng ẩn trong một mạng có thể nhiều hơn 1. Có chức năng nhận các giá trị từ từng input hoặc tầng ẩn trước nó, tính toán các giá trị và gửi đến các node ở các tầng ẩn hoặc tầng output tiếp theo đó tùy theo từng mạng cụ thể.
- **Tầng output - output layer:** nhận giá trị từ tầng trước đó (tầng ẩn hoặc tầng input) để tính toán các giá trị ngõ ra.
- **Các liên kết:** mỗi node trong một tầng truyền thông tin qua các node ở các tầng khác thông qua các liên kết. Các giá trị mà các liên kết này được gán sẽ được gọi là trọng số liên kết (weight). Giá trị trọng số được kết nối vào neuron j với neuron k là w_{kj} .
- **Hàm truyền - transfer function:** dùng để tính tổng các tích input

với trọng số liên kết của nó.

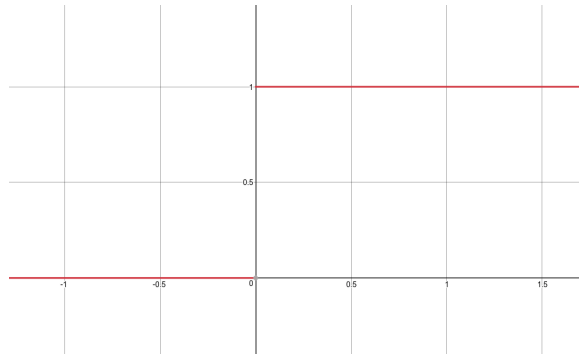
$$\sum input_j * w_{ij}$$

- **Activation function:** dùng để tính toán giá trị input sang giá trị output. Tùy vào mục đích và cụ thể từng loại mạng mà có nhiều loại activation function khác nhau.

Trong các bài toán khác nhau, người có những loại hàm activation như sau.

– *Step function:*

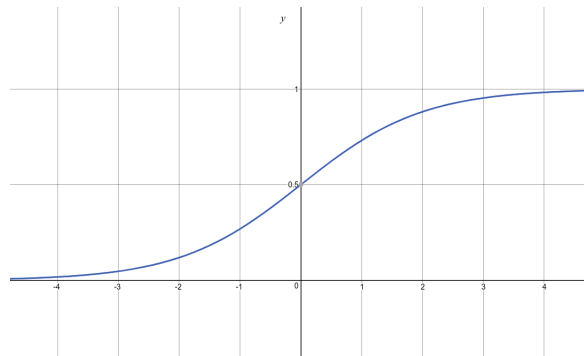
$$\begin{cases} 0 & x < 0 \\ 1 & x > 0 \end{cases}$$



Hình 2.3: Đồ thị hàm step

– *sigmoid function:*

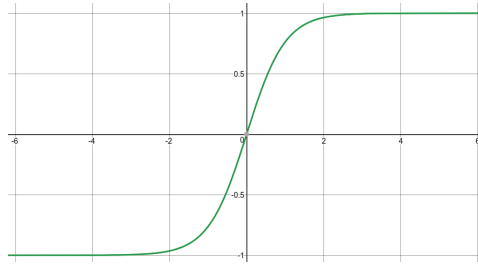
$$A(x) = \frac{1}{1 + e^{-x}}$$



Hình 2.4: Đồ thị hàm sigmoid

– *Tanh function:*

$$\text{Tanh}(x) = \frac{2}{1 + e^{-2x}} - 1$$



Hình 2.5: Đồ thị hàm Tanh

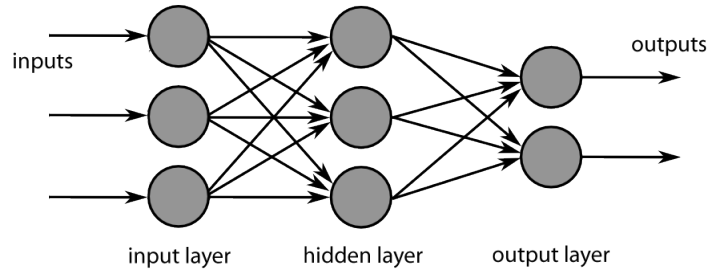
2.1.3 Mô hình Feedforward Neural Network

Thời điểm hiện nay, chúng ta có rất nhiều loại mô hình mạng neuron do sự khác nhau về sự kết hợp cũng như về mặt kiến trúc và thuật toán mà mạng đó áp dụng. Trong phần này, chúng ta sẽ tìm hiểu về mô hình mạng Feedforward Neural Network (FFNN), đây là kiến trúc mạng neuron được sử dụng phổ biến trong các bài toán dự báo. Mô hình gồm hai thành phần chính đó là kiến trúc feedforward - mạng truyền thẳng và giải thuật Backpropagation được áp dụng trong mạng.

2.1.3.1 Kiến trúc Feedforward

Đối với mạng feedforward, cấu trúc gồm một tầng input, một tầng output và có thể có nhiều hơn một tầng ẩn nằm giữa hai tầng input và output.

Như hình 2.6, một mạng Feedforward, trong tầng input và output thì số lượng neuron tại mỗi hai tầng này sẽ là cố định tùy theo đặc tính của dữ liệu. Đối với tầng ẩn, số lượng tầng ẩn cũng như số lượng neuron trong mỗi tầng tùy thuộc vào cá nhân thiết kế.



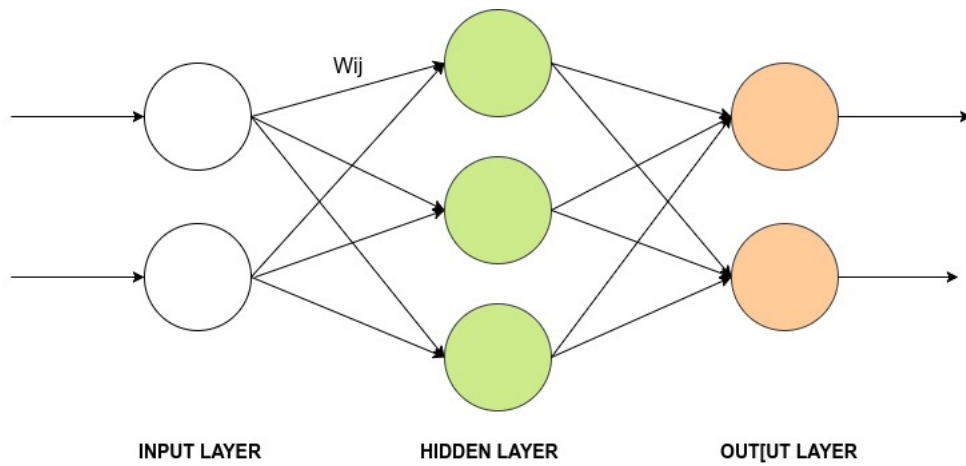
Hình 2.6: Ví dụ kiến trúc mạng feedforward

2.1.3.2 Giải thuật Backpropagation

Ở nội dung này, chúng ta sẽ không đi sâu vào kỹ thuật xử lý các phép tính[2] đạo hàm mà chỉ trình bày giải thuật lan truyền ngược một cách đơn giản nhất.

Các ký hiệu và hàm được dùng trong trình bày giải thuật:

- W_{ij} : là trọng số nối node thứ i tới node j ở layer kế tiếp.
- I_j : là đầu vào tại node thứ j .
- O_j : là kết quả xuất tại node thứ j .
- θ_j : bias tại node j .
- l : tốc độ học của mạng (learning rate).
- Err_j : giá trị lỗi tại node thứ j .
- Activation function được dùng trong nội dung này là hàm Sigmoid như mục trên



Hình 2.7: MLP

Nội dung thuật toán.

Input:

- Mạng feed forward với n input, m node ở tầng ẩn, và p output.
- Hệ số học hay tốc độ học l .
- Tập dữ liệu huấn luyện D .
- Sai số học ϵ .

Output: Vector các trọng số mới sau khi huấn luyện.

Nội dung thuật toán:

- **Bước 1:** Khởi tạo ngẫu nhiên các giá trị trọng số W_{ij} .
- **Bước 2:** Tính toán các giá trị đầu vào I_j và đầu ra O_j .
 - Tại node i ở tầng input:

$$I_i = x_i, O_i = I_i$$

- Tại node j ở tầng khác:

$$I_j = \sum_i W_{ij} O_i + \theta_j$$

$$O_j = \frac{1}{1 + e^{-I_j}}$$

- **Bước 3:** Tính toán lỗi trung bình và đánh giá.

- Tại node thuộc tầng output:

$$Err_j = O_j(1 - O_j)(T_j - O_j)$$

- Tại node thuộc tầng ẩn:

$$Err_j = O_j(1 - O_j) \sum_k Err_k W_{jk}$$

Với Err_k, W_{jk} là giá trị lỗi tại node k ở tầng tiếp theo và giá trị trọng số của node j đến k .

Thuật toán sẽ dừng lại khi $Err_k \leq \epsilon$

- **Bước 4:** Cập nhật các trọng số và độ lệch

$$W_{ij} = W_{ij} + (l)Err_j O_i$$

$$\theta_j = \theta_j + (l)Err_j$$

Thuật toán sẽ tiếp tục lặp lại bước 2 cho đến khi thỏa điều kiện dừng và cho ra các tập trọng số và độ lệch tốt nhất.

2.2 Mạng neural tích chập (Convolutional neural network)

2.2.1 Giới thiệu mạng neural tích chập

Trong vài năm trở lại đây, chúng ta thấy được sự nở rộ của các hệ thống thông minh từ các công ty công nghệ lớn trên thế giới. Các chức năng nhận dạng, phân loại hay dự đoán được áp dụng rộng rãi vào các lĩnh vực thương mại, vận tải..v..v.

Mô hình Deep learning được sử dụng phổ biến và phát triển giúp các hệ thống thông minh có độ chính xác cao ngày nay chính là Convolutional Neural Networks(CNN) - mạng neuron tích chập. Trong các nội dung tới, chúng ta sẽ tìm hiểu các khái niệm, kiến trúc, cũng như ứng dụng của CNN trong lĩnh vực phân loại ảnh.

2.2.2 Mô hình mạng neural tích chập

2.2.2.1 Input và output

Phân loại ảnh là công đoạn chuyển hóa từ một đầu là một hình ảnh và kết quả là một nhãn ứng với hình ảnh đó hoặc là các xác suất mà hệ thống dự đoán dựa trên đặc điểm của ảnh. Với con người, công việc nhận diện này được hình thành từ khi mới sinh ra, chúng ta có thể đưa ra kết quả của một hình ảnh bất kỳ mà không chút khó khăn. Nhưng máy tính thì không đơn giản như vậy, đầu vào và kết quả phải được đưa về dạng kỹ thuật số mà máy có thể hiểu được.

Khi một máy tính nhận vào một ảnh, nó sẽ thấy một mảng các giá trị

pixel tùy thuộc vào kích thước và độ phân giải của ảnh[3]. Ví dụ, một ảnh màu có kích thước 224 x 224 pixel thì máy tính sẽ thấy hình ảnh này dưới dạng một mảng có kích thước 224 x 224 x 3, giá trị 3 do thuộc tính ảnh màu(RGB) mà có được, giá trị này sẽ là 1 nếu đây là ảnh trắng đen.

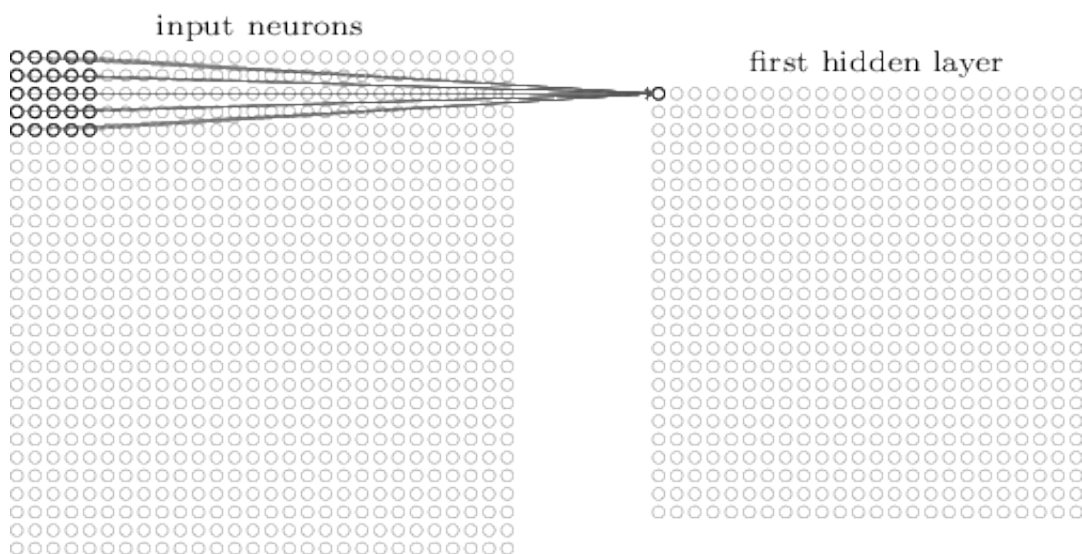
Đối với output, đây là một mảng các giá trị xác suất, mảng giá trị này cũng tùy thuộc vào số lượng nhãn(lớp) cần dự đoán. Ví dụ, (0.90 cho xe ô tô, 0.1 cho xe tải).

2.2.2.2 Convolutional layer

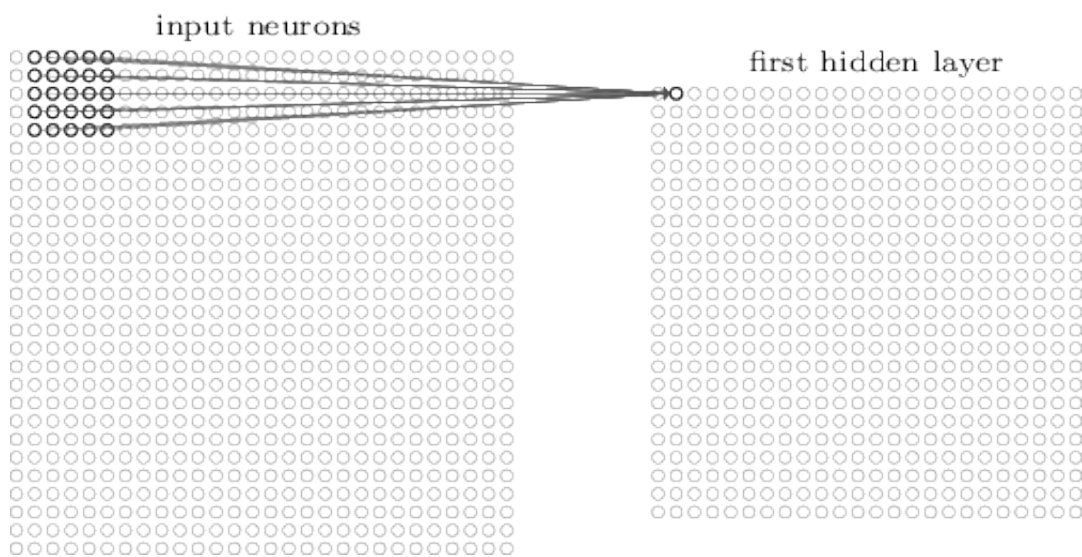
Tầng đầu tiên của một mạng CNN luôn luôn là tầng tích chập (convolutional layer)[4]. Như đã biết đầu vào (input) là một mảng các giá trị pixel. Trong trường hợp cụ thể để dễ hình dung ta chọn input là mảng các giá trị pixel có kích thước 32 x 32 x 3 với 32 x 32 là chiều dài và chiều rộng của tấm hình và 3 là giá trị RGB khi là ảnh màu. Đối với input trên có nghĩa là sẽ có một ma trận có kích thước 32 x 32 pixel mỗi pixel sẽ chứa 3 giá trị mà mỗi giá trị đó lần lượt biểu diễn cho giá trị của 3 màu sắc trên máy tính là đỏ(RED), lục(GREEN) và lam(BLUE).

Tạm thời bỏ qua giá trị RGB để đi vào cách hoạt động của tầng tích chập này. Cách đơn giản để giải thích cách hoạt động của tầng tích chập là tưởng tượng sẽ có một khuôn sẽ trượt từ phía trên bên trái cho đến hết tấm ảnh[5]. Với kích thước ảnh là 32 x 32 như trên, chọn kích thước ô trượt ví dụ là 5 x 5. Ô trượt có kích thước 5 x 5 sẽ trượt lần lượt qua cả input ảnh, ô trượt này được gọi là kernel hay filter(bộ lọc). Bộ lọc là một mảng các giá trị trọng số. Một điểm ghi chú là chiều sâu của bộ lọc sẽ bằng với chiều sâu của ảnh, với input 32 x 32 x 3 thì bộ lọc cũng sẽ có 5 x 5 x 3. Hình 2.8 và

2.9 minh họa cách kernel trượt trên input.

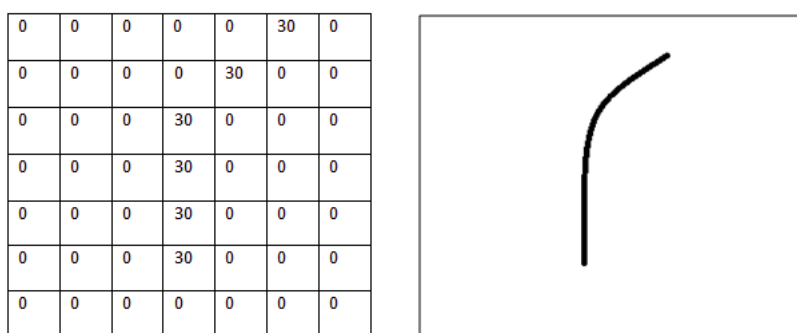


Hình 2.8: Convolutional layer



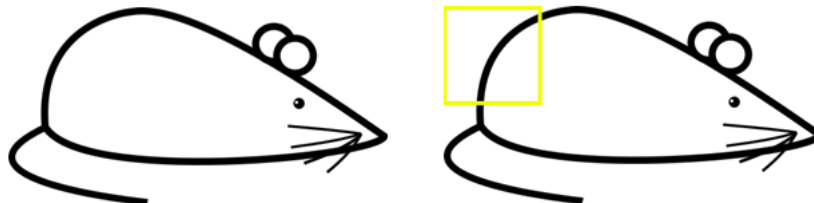
Hình 2.9: Convolutional layer

Bây giờ ta sẽ đi vào việc mà tăng tích chập thực sự làm với những phép tính. Như đã trình bày rằng mỗi bộ lọc sẽ là một mảng các giá trị pixel, công dụng của mảng giá trị này nhằm mục đích phát hiện các đặc tính của mỗi vùng input mà filter trượt qua. Các đặc tính ở đây có thể là đường thẳng, đường cong, màu đơn giản. Ví dụ ta có một filter có kích thước là $7 \times 7 \times 3$ dùng để phát hiện một dạng đường cong như hình 2.10.

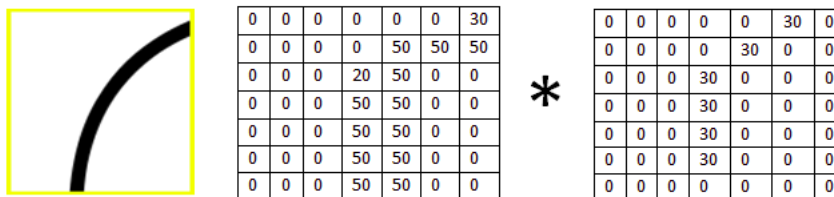


Hình 2.10: Mảng các giá trị của bộ lọc

Khi bộ lọc trên trượt đến vùng được đánh dấu vàng có dạng đường cong giống với bộ lọc như hình 2.11. Lúc đó phép toán tích chập sẽ được thực hiện như hình 2.12 với ma trận bên trái chính là giá trị pixel của vùng được đánh dấu trên ảnh mà bộ lọc trượt tới, ma trận bên phải chính là bộ lọc được sử dụng hiện tại.



Hình 2.11: ảnh đầu vào



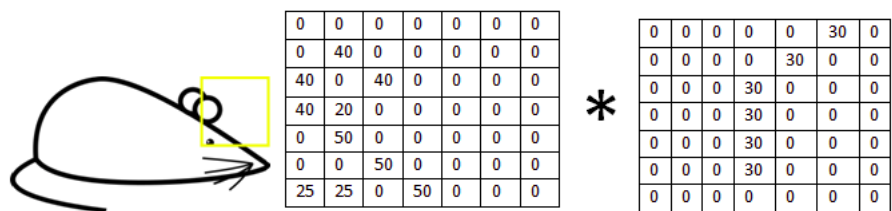
Hình 2.12: phép toán tích chập

Kết quả của phép tính trên sẽ có kết quả như sau:

$$(50 * 30) + (50 * 30) + (50 * 30) + (20 * 30) + (50 * 30) = 6600$$

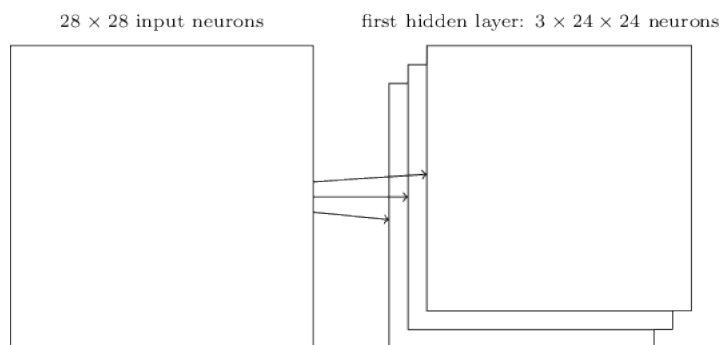
Đây là một con số rất lớn, thông thường nếu bộ lọc trượt tới một vùng mà vùng đó có hình dạng tương tự như bộ lọc thì kết quả khi thực hiện phép tính là một con số rất lớn. Ngược lại, kết quả sẽ ra rất nhỏ hoặc bằng 0. Ví dụ là hình ảnh [2.13](#) kết quả sẽ là 0 khi thực hiện phép tính

Trên đây là mô tả đơn giản về tầng tích chập, trên thực tế, chúng ta có thể có nhiều bộ lọc để phát hiện các khuôn mẫu, hình dạng khác nhau trong ảnh đầu vào. Kết quả xuất của tầng tích chập thứ nhất còn được gọi là bản



Hình 2.13: phép toán tích chập

đồ đặc tính (feature map) và có thể có nhiều feature map cho một input sau khi hoàn thành tầng tích chập. Như 2.14 biểu diễn một input kích thước như hình sau khi qua tầng tích chập và sử dụng tập các bộ lọc 5 x 5 cho ra tập 3 feature map mà mỗi cái nhận diện được một khuôn dạng khác nhau xuất hiện trong input.

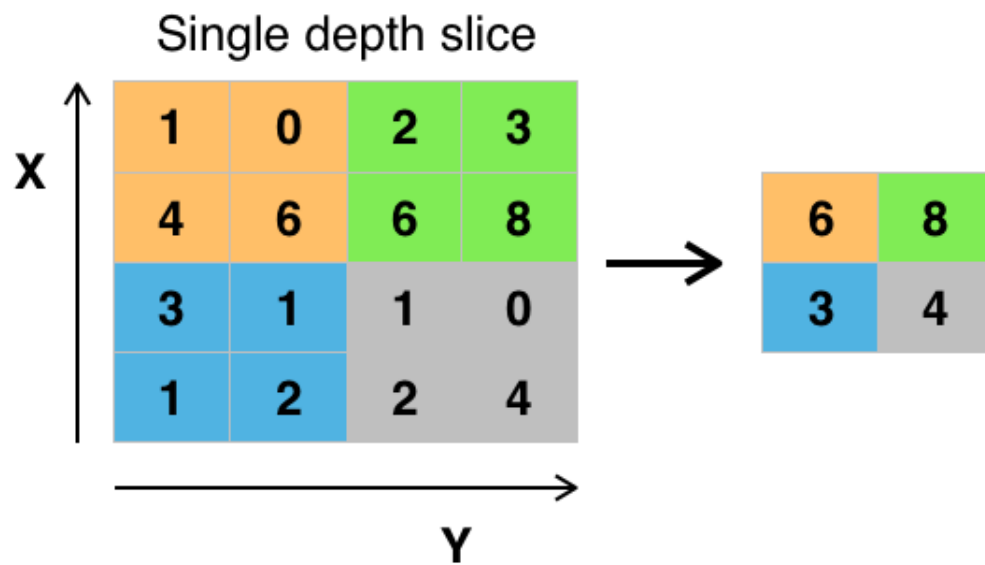


Hình 2.14: kết quả tầng tích chập

2.2.2.3 Pooling Layer

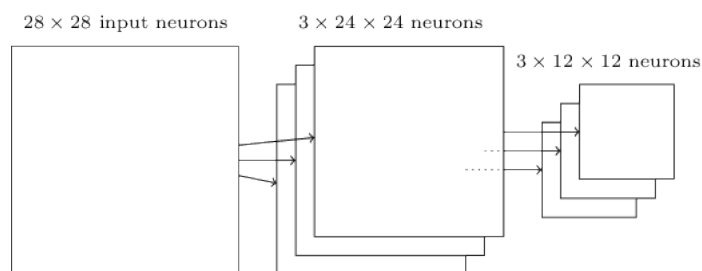
Sau khi qua một hoặc vài tầng tích chập, CNN sẽ chứa các tầng tổng hợp (Pooling Layer) ngay sau đó. Ở đây, tầng tổng hợp sẽ đơn giản hóa thông tin được lấy từ đầu ra từ tầng tích chập trước đó. Có nhiều kiểu tổng hợp

khác nhau đối với tầng này, nhưng max-pooling là phép tổng hợp phổ biến được sử dụng. Hình 2.15 biểu diễn một ví dụ phép max-pooling với một bộ lọc có kích thước là 2×2 . Giá trị lớn nhất trong mỗi vùng được trượt qua sẽ được chọn làm kết quả xuất ra.



Hình 2.15: max-pooling 2×2

Công dụng của phép pooling này giúp giảm đi kích thước của tập miêu tả đặc trưng từ đó cũng làm cho số lượng tham số và tính toán giảm theo. Và do chúng ta có thể có nhiều feature map từ tầng tích chập nên phép tổng hợp cũng sẽ được áp dụng độc lập cho mỗi feature map. Nếu có 3 feature map thì sẽ có 3 phép tổng hợp trong trường hợp này là max-pooling.



Hình 2.16: Ví dụ tầng tổng hợp

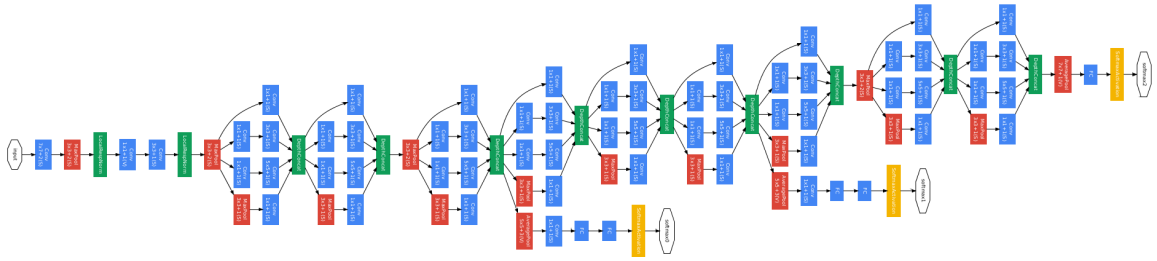
2.2.2.4 Fully Connected Layer

Sau khi thông tin input đã qua các tầng tích chập và tổng hợp, các giá trị dữ liệu sẽ đi đến tầng fully connected để xuất ra kết quả. Kết quả tại tầng fully connected là một vector có kích thước bằng với số lớp(nhãn) mà bài toán cần dự đoán. Như bài toán phân loại ảnh giao thông ùn tắc và thông thoáng thì lúc này số nhãn cần dự đoán và kích thước vector tại tầng này sẽ bằng 2. Giá trị của các phần tử trong vector sẽ là giá trị xác suất của mỗi nhãn mà mạng dự đoán, $[0.8, 0.2]$ sẽ biểu diễn 80% ảnh này thuộc lớp 1 và 20% ảnh thuộc lớp thứ 2. Về cơ bản, cách kết nối ở tầng này giống như cách kết nối neuron giữa các tầng với nhau ở mạng neuron ở mục trước. Khi đó, tất cả neuron ở tầng pooling sẽ kết nối với từng neuron trong tầng cuối.

2.2.3 Kiến trúc mạng GoogLeNet)

2.2.3.1 Ý tưởng

Đây là kiến trúc mạng tích chập với 22 tầng. GoogLeNet còn là quán quân của ILSVRC 2014 [9]. Mạng googLeNet có cấu trúc mạng nằm trong mạng, có 9 tầng mà mỗi tầng là một inception module. Theo tài liệu cho biết, việc áp dụng inception module giúp làm giảm đáng kể số lượng tham số tính toán giúp giải quyết vấn đề về tài nguyên. 2.17 minh họa cho cấu trúc của mạng googLeNet.

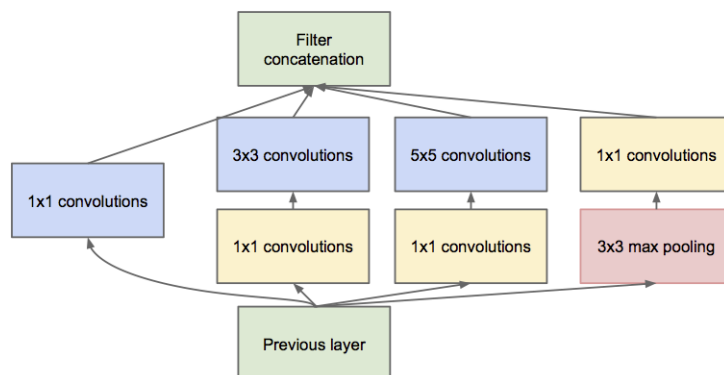


Hình 2.17: GoogLeNet

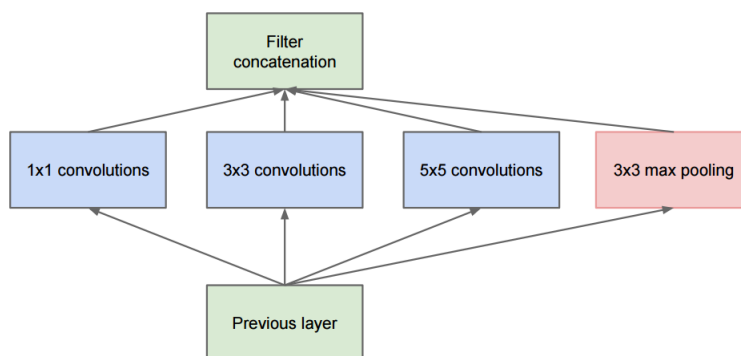
2.2.3.2 Inception module

Với minh họa kiến trúc của mạng googLeNet, ta sẽ thấy các layer là một khối mạng nhỏ nằm bên trong. Đây là các inception module. Đối với các mạng tích chập thông thường, khi một tập dữ liệu bắt đầu đi vào một tầng thì sẽ chỉ có hai sự lựa chọn đó chính là tầng tích chập hoặc tầng tổng hợp, nhưng với googLeNet sẽ có tập input sẽ đi vào một lớp module tại đó sẽ các

phương thức tích chập và pooling sẽ được tính toán một cách song song và độc lập với nhau [9].



Hình 2.18: inception module



Hình 2.19: naive inception module

Hình 2.18 miêu tả cấu trúc của một module trong mạng. Trong khi đó 2.19 là một ý tưởng ban đầu mà tác giả đã nghĩ tới. Ở 2.18 chúng ta thấy trước khi thực hiện các phép tích chập với các filter 3 x 3 và 5 x 5, input đều được xử lý qua phép tích chập với filter 1 x 1. Các bộ lọc 1 x 1 có tác dụng làm giảm đi chiều của các input[8], điều này giúp cho khối lượng tham số

phải tính toán ở phép toán tích chập với bộ lọc 3×3 và 5×5 sẽ được giảm đi một cách đáng kể.

Chương 3

Apache Hadoop

3.1 Tổng quan về hệ thống Apache Hadoop

Apache Hadoop là một dự án phát triển phần mềm nhằm cung cấp một nền tảng phân tán, có thể mở rộng linh hoạt và có độ tin cậy cao.

Ngoài ra Hadoop còn được xem là một thư viện hay framework cho phép xử lý phân tán khối lượng lớn dữ liệu trên nhiều cụm máy tính bằng các mô hình lập trình. Framework này được thiết kế với mục đích có khả năng mở rộng từ một máy chủ đơn lẻ lên đến rất nhiều trạm làm việc mà mỗi máy trạm có khả năng tính toán và lưu trữ cục bộ.

3.2 Hadoop distributed file system - HDFS

3.2.1 Thiết kế

HDFS là một hệ thống file nhằm lưu trữ một lượng rất rất lớn (Lớn ở đây theo nghĩa có thể là hàng trăm megabytes, gigabytes, hoặc terabytes) với cơ chế streaming data access trên những thiết bị phổ thông ¹. Đây là cơ chế phân luồng dữ liệu trong HDFS với mục đích ghi một lần chạy nhiều lần. Điển hình là dữ liệu sẽ được sinh và sao chép từ nguồn và sau đó có rất nhiều tiến trình phân tích khác nhau thực thi dữ liệu trên. Mỗi hoạt động phân tích sẽ thực thi liên quan đến một phần nào đó khác nhau trên cả một tập dữ liệu trên.

Từ các thiết bị phổ thông ở đây là những những thiết bị phần cứng máy tính, HDFS không yêu cầu một phần cứng đắt tiền hay độ tin cậy cao. Mà nó được thiết kế để chạy trên những cụm máy tính từ nhiều nhà cung ứng khác nhau. Vì thế mà xác suất để một node (đơn vị phần cứng) gặp lỗi và thất bại là rất lớn, đặc biệt là những cụm có hàng ngàn máy trạm. Với đặc tính đó, HDFS được thiết kế sao cho không có sự gián đoạn nào được phát hiện ở người dùng khi mà việc một số lượng node gặp lỗi giữa chừng.

3.2.2 Các khái niệm trong HDFS

- **Block(sector):** trong các đĩa cứng thông thường, đây là các phần nhỏ cuối cùng được chia ra để chứa dữ liệu, và đây cũng là đơn vị nhỏ nhất mà một lần đọc và ghi dữ liệu được hiện thực. Thông thường, block ở các đĩa cứng có dung lượng nhỏ chỉ tầm 512 byte. Ở HDFS,

¹các máy tính hoặc máy trạm

block thường có dung lượng lớn hơn rất nhiều - mặc định là 128MB. Do dung lượng của một block trên đĩa lớn như vậy nên các file tập tin trong HDFS nhiều khi có dung lượng và được lưu trữ nằm trong cả một block (trong HDFS). Nguyên nhân của việc block trong HDFS có dung lượng lớn như vậy là nhằm việc tối thiểu hóa chi phí tìm kiếm [12]

- **Namenodes và Datanodes:** một HDFS gồm nhiều node, trong đó có hai loại chính đó là một master node (Namenode) - máy chủ và nhiều datanodes (những thiết bị nô lệ). Namenode là chương trình chỉ đạo các datanodes thực hiện các nhiệm vụ I/O ở mức thấp. Còn datanodes được sự điều phối của Namenode, lưu trữ và thường xuyên báo cáo với master các khối(block) mà nó hiện đang lưu trữ. Ngoài ra, các datanodes còn giao tiếp với nhau để sao lưu các khối dữ liệu dự phòng.
- **jobTracker:** Trình nền JobTracker là một liên lạc giữa ứng dụng của bạn và Hadoop. Một khi bạn gửi mã nguồn của bạn tới các cụm (cluster), JobTracker sẽ quyết định kế hoạch thực hiện bằng cách xác định những tập tin nào sẽ xử lý, các nút được giao các nhiệm vụ khác nhau, và theo dõi tất cả các nhiệm vụ khi chúng đang chạy. Nếu một nhiệm vụ (task) thất bại (fail), JobTracker sẽ tự động chạy lại nhiệm vụ đó, có thể trên một node khác, cho đến một giới hạn nào đó được định sẵn của việc thử lại này. Chỉ có một JobTracker trên một cụm Hadoop. Nó thường chạy trên một máy chủ như là một nút master của cluster.
- **TaskTraker:** Như với các trình nền lưu trữ, các trình nền tính toán cũng phải tuân theo kiến trúc master/slave: JobTracker là giám sát

tổng việc thực hiện chung của một công việc MapReduce và các task-Tracker quản lý việc thực hiện các nhiệm vụ riêng trên mỗi node slave. Mỗi TaskTracker chịu trách nhiệm thực hiện các task riêng mà các JobTracker giao cho. Mặc dù có một TaskTracker duy nhất cho một node slave, mỗi TaskTracker có thể sinh ra nhiều JVM để xử lý các nhiệm vụ Map hoặc Reduce song song. Một trong những trách nhiệm của các TaskTracker là liên tục liên lạc với JobTracker. Nếu JobTracker không nhận được nhịp đập từ một TaskTracker trong vòng một lượng thời gian đã quy định, nó sẽ cho rằng TaskTracker đã bị treo (cached) và sẽ gửi lại nhiệm vụ tương ứng cho các nút khác trong cluster. Cấu trúc liên kết này có một node Master là trình nền NameNode và JobTracker và một node đơn với SNN trong trường hợp node Master bị lỗi. Đối với các cụm nhỏ, thì SNN có thể thường chú trong một node slave. Mặt khác, đối với các cụm lớn, phân tách NameNode và JobTracker thành hai máy riêng. Các máy slave, mỗi máy chỉ lưu trữ một DataNode và TaskTracker, để chạy các nhiệm vụ trên cùng một node nơi lưu dữ liệu của chúng

3.2.3 HIVE

Hive là hạ tầng kho dữ liệu cho Hadoop. Nhiệm vụ chính là cung cấp sự tổng hợp dữ liệu, truy vấn và phân tích. Nó hỗ trợ phân tích các tập dữ liệu lớn được lưu trong HDFS của Hadoop cũng như trên Amazon S3. Điểm hay của HIVE là hỗ trợ truy xuất giống SQL đến dữ liệu có cấu trúc, được biết với tên HiveSQL (hoặc HQL) cũng như phân tích big data với MapReduce. Hive không được xây dựng để hồi đáp nhanh các câu truy vấn nhưng nó được

xây dựng cho các ứng dụng khai thác dữ liệu (data mining). Các ứng dụng khai thác dữ liệu có thể mất nhiều phút đến nhiều giờ để phân tích dữ liệu và HIVE được dùng chủ yếu.

3.2.3.1 Cách tổ chức của HIVE

Dữ liệu được tổ chức thành 3 định dạng trong HIVE.

Tables: Chúng rất tương tự như bảng (tables) trong RDBMS và chứa các dòng (rows). Hive chỉ được xếp lớp trên HDFS, do đó tables được ánh xạ trực tiếp vào các thư mục của hệ thống tập tin. Nó cũng hỗ trợ các tables được lưu trên các hệ thống tập tin khác.

Partitions: Hive tables có thể có nhiều hơn 1 partition. Chúng được ánh xạ với các thư mục con và các hệ thống tập tin.

Buckets: Trong Hive, dữ liệu có thể được chia thành các buckets. Buckets được lưu trữ như các tập tin trong partition trong hệ thống tập tin.

Hive cũng có metastore để lưu tất cả metadata. Nó là cơ sở dữ liệu quan hệ chứa thông tin khác nhau liên quan đến Hive Schema (column types, owners, key-value data, statistics,...). Chúng ta có thể dùng MySQL cho việc này.

3.2.3.2 HIVESQL

Ngôn ngữ truy vấn Hive cung cấp các toán tử cơ bản giống SQL. Đây là một số tác vụ mà HQL có thể làm dễ dàng.

- Tạo và quản lý tables và partitions.
- Hỗ trợ các toán tử Relational, Arithmetic và Logical khác nhau.

- Evaluate functions
- Tải về nội dung 1 table từ thư mục cục bộ hoặc kết quả của câu truy vấn đến thư mục HDFS.
- Đây là ví dụ truy vấn HQL:

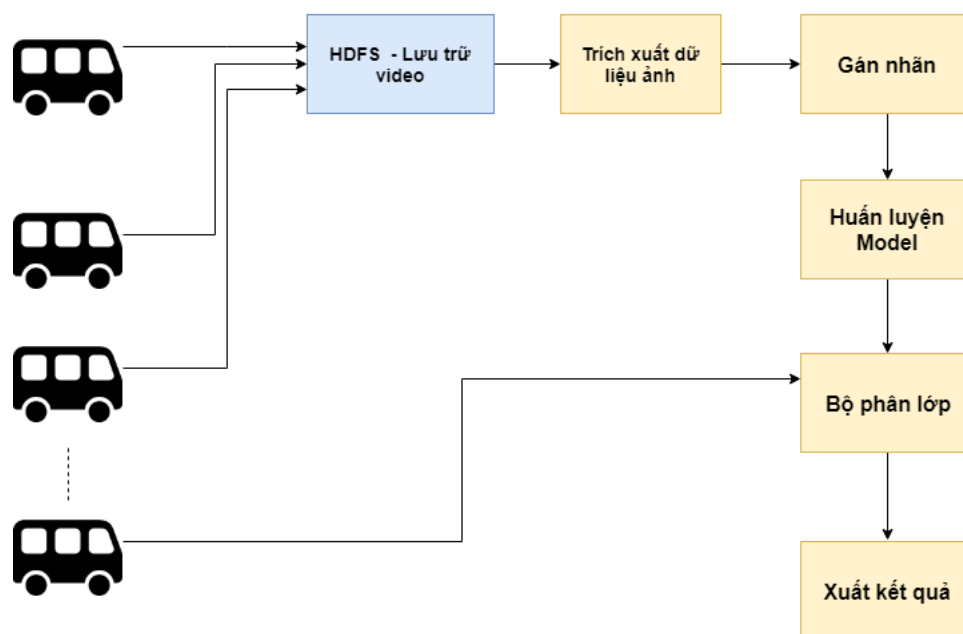
```
SELECT upper(name), salesprice
FROM sales ;
SELECT category , count(1)
FROM products
GROUP BY category ;
```

3.3 Tại sao sử dụng Apache Hadoop trong đề tài

Hiện nay trên các e buýt đều được trang bị camera hành trình, các video này ghi lại hình ảnh trên các con đường. Các hình ảnh này được gửi về hệ thống một cách liên tục hàng ngày và hàng giờ. Do kích thước dữ liệu tăng lên đáng kể cần một hệ thống lưu trữ có khả năng mở rộng linh động, vì thế hệ thống HDFS do Apache Hadoop hỗ trợ là công cụ hợp lý. Ngoài ra hệ thống HDFS còn có thể kết hợp với các thư viện hỗ trợ học sâu như Tensorflow nên ta có thể áp dụng các mô hình trong lĩnh vực machine learning nói chung và deep learning nói riêng để xây dựng một hệ thống phân loại ảnh. Khi kết hợp các thành phần trên, ta có thể tinh chỉnh tập dữ liệu sao cho phù hợp với bài toán của thể của từng trường hợp cũng như các lĩnh vực khác mà không nhất thiết là lĩnh vực giao thông.

Chương 4

Mô hình loại ảnh giao thông kết hợp Apache Hadoop



Hình 4.1: Mô hình phân loại ảnh giao thông kết hợp Hadoop

4.1 Tổng quan các bước

Hệ thống phân loại ảnh kết hợp Apache Hadoop và mô hình học sâu Deeplearning sẽ được thực hiện dựa trên sơ đồ trên. Hiện nay, các xe buýt đều được trang bị camera hành trình và hệ thống GPS. Các video được ghi lại sẽ được lưu tại một nơi. Do số lượng của các video được gửi về liên tục và hàng ngày và giờ, nên việc sử dụng hệ thống lưu trữ HDFS của Apache Hadoop là lựa chọn hợp lý. Tiếp đến là việc truy xuất ảnh từ video và gán nhãn ảnh phục vụ cho việc huấn luyện mô hình học sâu. Đây cũng là công đoạn tốn kém chi phí nhất. Sau cùng là việc huấn luyện tập dữ liệu đã xây dựng.

4.2 Mô tả tập dữ liệu

4.3 Phương pháp transfer learning

Ngày nay, với sự phát triển của Deep learning do nguồn dữ liệu to lớn và các máy tính ngày càng cải tiến về khả năng tính toán khiến cho kết quả chính xác của các bài toán phân loại ngày càng cao. Như mô hình mạng googLeNet có rất nhiều tầng khiến cho việc huấn luyện tốn kém thời gian. Thay vì như vậy chúng ta áp dụng phương pháp transfer learning.

Transfer learning là công đoạn lấy model đã được huấn luyện từ một tập dữ liệu khác [6] và tích hợp lại với tập dữ liệu có đang có. Việc này là khả thi do mô hình đã huấn luyện trước đó sử dụng tập ảnh khổng lồ giúp cho mô hình học được loạt những đặc tính thường thấy trong cùng một ảnh. Có thể thấy tất cả các ảnh đều có những đặc tính cơ bản giống nhau, khi muốn

huấn luyện lại cho tập ảnh của chính mình chúng ta chỉ cần thay tầng cuối cùng của mạng bằng tập dữ liệu của mình. Mô hình sẽ tự điều chỉnh lại các trọng số cũng như độ lệch từ các giá trị từ mô hình đã huấn luyện với tập dữ liệu mới mà không cần làm lại từ đầu.

4.4 Các bước chuẩn bị

4.4.1 Chuẩn bị môi trường hiện thực

Bộ phân loại ảnh giao hông được huấn luyện trên nền tảng hệ điều hành Linux, ngôn ngữ Python phiên bản 3.6 kết hợp với thư viện Tensorflow mã nguồn mở chuyên được sử dụng cho những mô hình học sâu.

Tensorflow[11] là một thư viện học sâu mã nguồn mở được Google phát triển. Thư viện này đã thu hút được sự chú ý lớn từ cộng đồng Deep-learning. Tensorflow cho phép chạy các thuật toán machine learning trên nhiều GPU, có nhiều module được dựng sẵn giúp cho việc xây dựng và thực thi mô hình đơn giản hơn.

Các khái niệm:

- **Tensor:** đây là cấu trúc dữ liệu được sử dụng hoàn toàn trong Tensorflow. Hay nói cách khác, tất cả dữ liệu đều biểu diễn dưới dạng tensor. Đơn giản, tensor là một mảng gồm n chiều hay list kèm theo một số thuộc tính khác.
- **Rank:** còn được gọi là số chiều của dữ liệu
- **Shape:** biểu diễn chiều của tensor. Ví dụ, $t = [[1, 2, 3], [4, 5, 6], [7, 8, 9]]$ có shape là $[3, 3]$, $t = [[[1], [2], [3]], [[4], [5], [6]], [[7], [8], [9]]]$ có shape là

Rank	Đơn vị số	Ví dụ
0	Scalar	$s = 123$
1	Vector	$s = [0.8, 0.1, 0.1]$
2	Matrix	$s = [[1,2,3], [4,5,6], [7,8,9]]$
3	3-Tensor	$s = [[[1], [2], [3]], [[4], [5], [6]], [[7], [8], [9]]]$
n	n-Tensor	n chiều dữ liệu...

Bảng 4.1: Các chiều dữ liệu.

$[1, 3, 3], \dots$

- **Type:** là các kiểu dữ liệu được sử dụng trong Tensorflow. Một vài kiểu dữ liệu cơ bản như.

Data type	Python code	Mô tả
<i>DT-FLOAT</i>	tf.float32	32 bits floating point.
<i>DT-DOUBLE</i>	tf.float64	64 bits floating point.
<i>DT-INT16</i>	tf.int16	16 bits signed integer.
<i>DT-INT32</i>	tf.int32	32 bits signed integer.
<i>DT-INT64</i>	tf.int64	64 bits signed integer.
...

Bảng 4.2: Một vài kiểu dữ liệu.

4.5 Các bước hiện thực

4.5.1 Cài đặt Apache Hadoop

Trong phần hướng dẫn này chúng ta sẽ biết được cách cài đặt một cụm gồm nhiều máy Hadoop. Thông qua nhiều bước,ta sẽ biết cách cấu hình cũng như khởi tạo và kết thúc hoạt động của một cụm máy Hadoop.

4.5.1.1 Cấu hình cần thiết để cài đặt một cụm 3 máy hadoop

Một cụm Hadoop có thể rất nhiều máy, nhưng để có được cái nhìn tổng quan và vì thời gian của đề tài nên chúng ta sẽ tiến hành cài Hadoop trên 3 máy tính. Việc bổ sung thêm máy trạm vào cụm sẽ được hướng dẫn ở nội dung cuối phần này.

- **Số lượng máy tính:** 3 máy tính.
- **Hệ điều hành:** sử dụng hệ điều hành Ubuntu phiên bản 14.04 hoặc 16.04 hoặc cao hơn.
- **Phiên bản Apache Hadoop:** Gói cài đặt [Hadoop phiên bản 3.1.0](#)

4.5.1.2 Cài đặt Hadoop trên máy Master(Namenode)

Mục này được thực hiện chỉ riêng trên máy được chọn làm Master.

1. Các vấn đề trước khi cài đặt Hadoop:

- Thêm địa chỉ IP đầu vào vào file */etc/hosts*

Listing 4.1: Nội dung file */etc/hosts*

127.0.0.1	localhost
127.0.1.1	master
192.168.1.50	master
192.168.1.61	slave
192.168.1.73	slave1

Chúng ta bỏ qua 2 dòng đầu của file và chú ý tới 3 dòng tiếp theo. bên cột trái, đó chính là lần lượt địa chỉ IP của 3 máy sẽ cài Hadoop. 192.168.1.50 là máy được chọn làm master và từ *master* bên cột phải là hostname của máy đó. Tương tự 2 dòng cuối, đó là IP của 2 máy được chọn làm datanode và hostname lần lượt là *slave*, *slave1*.

- **Cài đặt Java:** hiện nay, Java được cài ở đây sẽ có phiên bản là Java 8. Thực hiện lần lượt các lệnh sau trên terminal để cài đặt Java 8.

Listing 4.2: Cài đặt Java 8

```
sudo add-apt-repository ppa:webupd8team/java
sudo apt-get update
sudo apt-get install oracle-java8-installer
```

- **Cấu hình SSH:** Thực hiện lần lượt các lệnh sau để cài cũng như cấu hình kết nối SSH máy master tới các máy datanodes.

Listing 4.3: Cài đặt SSH

```
sudo apt-get install openssh-server openssh-client  
ssh-keygen -t rsa -P ""
```

Sao chép nội dung file `.ssh/id_rsa.pub` của máy master đến file `.ssh/authorized_keys` của máy master cũng như các máy datanodes. Sau đó kiểm tra kết nối đã thành công hay chưa bằng lệnh.

Listing 4.4: Kết nối SSH

```
ssh slave  
ssh slave1
```

2. Cài đặt Hadoop

- **Tải gói cài đặt:** Tải gói cài đặt Hadoop theo link phía dưới
<http://hadoop.apache.org/releases.html>
- **Giải nén gói cài đặt:** giải nén đến thư mục `/usr/local` bằng câu lệnh

```
tar xzf hadoop-3.1.0.tar.gz -C /usr/local/
```
- **Chỉnh sửa `/.bashrc`:** thêm các biến môi trường cài đặt Hadoop vào file `/.bashrc` như nội dung sau.

```

export HADOOP_PREFIX="/usr/local/hadoop-3.1.0"
export PATH=$PATH:$HADOOP_PREFIX/bin
export PATH=$PATH:$HADOOP_PREFIX/sbin
export HADOOP_MAPRED_HOME=${HADOOP_PREFIX}
export HADOOP_COMMON_HOME=${HADOOP_PREFIX}
export HADOOP_HDFS_HOME=${HADOOP_PREFIX}
export YARN_HOME=${HADOOP_PREFIX}

```

Sau khi lưu nội dung chỉnh sửa, thực thi `source ~/.bashrc` để các biến hệ thống thiết lập các biến môi trường.

- **Cấu hình file `hadoop-env.sh`:** cấu hình biến `JAVA_HOME`

```
export JAVA_HOME=/usr/lib/jvm/java-8-oracle/
```

Thay đổi đường dẫn `JAVA_HOME` nếu cài đặt Java 8 ở một nơi khác.

- **Cấu hình file `core-site.xml`:** thêm các dòng sau

```

<configuration>
<property>
    <name>fs.defaultFS</name>
    <value>hdfs://master:9000</value>
</property>
<property>
    <name>hadoop.tmp.dir</name>
    <value>/home/'your\_user'/hdata</value>
</property>
</configuration>

```


- **Cấu hình file hdfs-site.xml:** thêm các nội dung như sau

```
<configuration>
<property>
    <name>dfs.replication</name>
    <value>2</value>
</property>
</configuration>
```

- **Cấu hình file mapred-site.xml:** thêm các nội dung như sau

```
<configuration>
<property>
    <name>mapreduce.framework.name</name>
    <value>yarn</value>
</property>
</configuration>
```

- **Cấu hình yarn-set.xml:** thêm các nội dung như sau

```
<configuration>
<property>
    <name>yarn.nodemanager.aux-services</name>
    <value>mapreduce_shuffle</value>
</property>
```

```
<property>
    <name>yarn.resourcemanager.resource-tracker.address</name>
    <value>master:8025</value>
```

```

</property>
<property>
    <name>yarn.resourcemanager.scheduler.address</name>
    <value>master:8030</value>
</property>
<property>
    <name>yarn.resourcemanager.address</name>
    <value>master:8040</value>
</property>
</configuration>

```

- **Cấu hình file workers:** thêm hostname của 3 máy trạm mà đã đặt ở file */etc/hosts*

```

master
slave
slave1

```

4.5.1.3 Cài đặt Hadoop trên các datanodes

1. Các cấu hình chuẩn bị:

- **Thêm địa chỉ IP vào file */etc/hosts*:** thực hiện giống như đã làm ở máy master
- **Cài đặt Java 8**

2. Cài đặt Hadoop cho tất cả máy datanodes: thao tác này sẽ được thực hiện trên terminal của máy master.

- **sử dụng SSH để gửi file cài đặt:** thực hiện lần lượt các lệnh sau để gửi file cài đặt hadoop tới các máy datanodes

```
scp /usr/local/hadoop-3.1.0 slave:/usr/local/hadoop-3.1.0
scp /usr/local/hadoop-3.1.0 slave1:/usr/local/hadoop-3.1.0
```

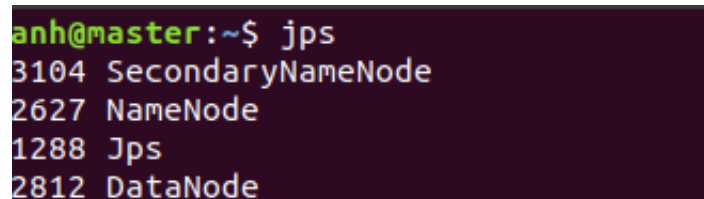
Sau các lệnh này được thực thi xong thì các máy datanodes đã được cài đặt hadoop thành công

- **Cấu hình file `/.bashrc`:** chúng ta có thể copy nội dung từ file `/.bashrc` của master đến các máy datanodes.

3. Khởi động cụm máy Hadoop: chạy các lệnh sau để khởi động Hadoop

- `bin/hdfs namenode -format`
- `sbin/start-dfs.sh`

4. Kiểm tra xem Hadoop service: lần lượt thực thi lệnh `jps` từ terminal của các máy master và datanode để kiểm tra, nếu kết quả như hình dưới đây thì các service hadoop đang hoạt động.



```
anh@master:~$ jps
3104 SecondaryNameNode
2627 NameNode
1288 Jps
2812 DataNode
```

Hình 4.2: Kiểm tra Hadoop service trên master

```
anh@slave:~$ jps
20394 Jps
9743 DataNode
```

Hình 4.3: Kiểm tra Hadoop service trên worker(slave)

```
anh@slave1:~$ jps
31668 Jps
1677 DataNode
```

Hình 4.4: Kiểm tra Hadoop service trên worker(slave1)

4.5.1.4 Thêm một datanodes vào một cụm Hadoop đang hoạt động

Trong quá trình sử dụng, việc muốn mở rộng cụm máy Hadoop là điều có thể xảy ra. Chúng ta có thể dễ dàng thêm một datanodes và một cụm máy Hadoop đang hoạt động theo các bước sau.

1. **Cài đặt Java 8:** đảm bảo rằng máy được thêm vào cụm phải được cài đặt Java 8 như đã trình bày ở phía trên.
2. **Thêm địa chỉ IP vào /etc/hosts:** thêm địa chỉ IP của máy datanode mới vào file /etc/hosts của máy master và các máy datanodes đã có cũng như datanode mới.
3. **Cấu hình SSH:** copy nội dung file /.ssh/id_rsa.pub vào máy datanodes mới giống như đã làm với các máy datanode đã trước.
4. **Cài đặt Hadoop:** chúng ta cũng sử dụng SSH để gửi thư mục cài Hadoop từ máy master sang máy datanode mới như đã thực hiện.

5. **Cấu hình file workers:** thêm hostname của máy datanode mới vào file workers trên máy master.
6. **Khởi động datanode:** thực hiện file *start-dfs.sh* trong thư mục sbin trên máy master để tiến hành khởi động datanode mới vào cụm máy.

4.5.1.5 Lưu trữ dữ liệu ảnh trên HDFS

Để đưa dữ liệu lên HDFS, ta sử dụng câu lệnh sau.

```
hdfs dfs -put ~/Document/traffic/ hdfs://master:9000/
```

Với */Document/traffic* là đường dẫn đến thư mục lưu dữ liệu ảnh, và *hdfs://master:9000/* là đường dẫn đến vị trí lưu của HDFS. Để xem kết quả put ta thực hiện lệnh sau để liệt kê các thư mục dữ liệu đã được lưu trên HDFS.

```
anh@master:/usr/local/hadoop/sbin$ hdfs dfs -ls /traffic
2018-06-02 14:00:28,706 WARN util.NativeCodeLoader: Unable to load native-hadoop
library for your platform... using builtin-java classes where applicable
Found 2 items
drwxr-xr-x - anh supergroup          0 2018-05-07 09:27 /traffic/ket
drwxr-xr-x - anh supergroup          0 2018-05-07 09:52 /traffic/thong
```

Hình 4.5: Liệt kê thư mục dữ liệu trên HDFS

4.5.2 Cài đặt HIVE

Để cài đặt HIVE và lưu trữ dữ liệu GPS, ta thực hiện các bước sau đây.

4.5.2.1 Tải gói cài đặt HIVE

Tải gói cài đặt HIVE ở với đường dẫn sau <http://mirror.downloadvn.com/apache/hive/>, ta sẽ sử dụng phiên bản 2.2.0.

4.5.2.2 Cài đặt HIVE

1. **Giải nén:** giải nén gói cài đặt và copy thư mục có được vào vị trí thư mục muốn cài chẳng hạn như */usr/local/hive*.
2. **Cấu hình biến môi trường:** thêm vào file */.bashrc* nội dung như sau

```
export HIVE_HOME=/usr/local/hive
export PATH=$PATH:$HIVE_HOME/bin
export CLASSPATH=$CLASSPATH:/usr/local/Hadoop/lib/*:.
export CLASSPATH=$CLASSPATH:/usr/local/hive/lib/*:.
```

Thực hiện lệnh *source ~/.bashrc* để thực thi cấu hình.

4.5.2.3 Tải và cài đặt Apache Derby

Làm theo các lệnh hướng dẫn để tải à cài đặt Apache Derby

1. **Tải Apache Derby:** vào và tải ở địa chỉ sau *http://archive.apache.org/dist/db/derby/* với phiên bản 10.14.2.0
2. **Giải nén:** giải nén vào thư mục */usr/local/derby*
3. **Cấu hình biến môi trường cho Derby:** Thêm nội dung file */.bashrc* như sau.

```
export DERBY_HOME=/usr/local/derby
export PATH=$PATH:$DERBY_HOME/bin
```

4.5.2.4 dữ liệu GPS

Dữ liệu GPS là tập dữ liệu lưu vị trí GPS của một xe buýt trong một thời điểm nhất định. Dữ liệu này được lưu bằng bảng trong HIVE. Cấu trúc của bảng dữ liệu bao gồm độ, vĩ độ, thời điểm theo định dạng DD/MM/YYYY hh:mm:ss.

4.5.3 Xử lý dữ liệu

Để xây dựng mô hình phân loại hình ảnh, chúng ta cần phải có một tập huấn luyện đủ tốt. Ở đây, các hình ảnh được trích xuất từ camera hành trình từ các tuyến xe buýt.

Cấu trúc tổ chức tập dữ liệu gồm 2 thư mục chính:

- Thư mục **ket**: chứa các hình ảnh được cho là giao thông trong tình trạng ùn tắc. Bao gồm khoảng 7500 hình ảnh định dạng JPG.
- Thư mục **thong**: chứa các hình ảnh được cho là giao thông trong tình trạng thông thoáng. Bao gồm 7700 hình ảnh định dạng JPG.

Các hình ảnh được trình bày và lưu trữ vào các thư mục có chứa các tên mô tả cho đặc tính của những hình ảnh đó. Với bộ dữ liệu trên, chương trình tạo một kiểu dữ liệu dictionary với khóa chính là giá trị biểu diễn cho tên thư mục và cũng là tên class cần phân loại, value chính là đường dẫn các file ảnh tương ứng.

Để sử dụng được trong mô hình mạng, các hình ảnh sẽ được mã hóa sang một định dạng mới nhờ các phương thức hỗ trợ có sẵn trong thư viện Tensorflow. Sau khi được mã hóa, kết quả chính là các tensor có thông số

shape như sau [299, 299, 3], với hai vị trí đầu tiên chính là kích thước của hình ảnh cũng như của tensor, giá trị 3 biểu diễn cho độ sâu (ảnh màu).

Sau cùng, bộ dữ liệu đã được mã hóa được chia thành 3 tập con sử dụng với 3 mục đích khác nhau: tập huấn luyện(Training set), tập validation để tránh vấn đề overrrfit trong quá trình huấn luyện và tập kiểm thử dùng để kiểm tra độ chính xác của mô hình sau khi huấn luyện hoàn tất. Riêng tập huấn luyện sẽ được dùng để tạo ra các bottlenecks

4.5.4 Tạo các bottlenecks

Bottlenecks[10] là một từ được dùng để chỉ tầng (layer) nằm ngay trước fully-connected layer. Với kiến trúc mạng googLeNet, tầng này đã được huấn luyện tập dữ liệu trước đó nên có được kết quả đủ tốt để phân biệt được đặc tính của mỗi lớp(class) yêu cầu. Có nghĩa ở bước này chúng ta sẽ tạo ra một bản tóm tắt các giá trị trọng số đủ tốt cho mỗi ảnh input. Tầng cuối cùng của kiến trúc mạng sẽ sử dụng các giá trị bottlenecks này để huấn luyện và điều chỉnh để phân loại các lớp mới. Điều này nhờ vào việc mạng đã được huấn luyện bởi tập dữ liệu gồm 1000 lớp khác nhau của ImageNet giúp cho việc phát hiện các mẫu đặc tính trở nên dễ dàng hơn.

4.5.5 Huấn luyện

Sau khi hoàn tất tạo các giá trị bottleneck, việc thực hiện thực hiện cấu hình mạng và huấn luyện bắt đầu. Tổng số bước huấn luyện sẽ được cài đặt mặc định là 4000 bước, tuy nhiên có thể thay đổi lại tùy theo tình huống.

Mỗi bước huấn luyện sẽ chọn ra 100 dữ liệu ngẫu nhiên ¹ để đưa vào tầng cuối cùng ² để dự đoán lớp, lớp dự đoán sẽ được so sánh với các lớp thực tế để mạng điều chỉnh và cập nhật các giá trị trọng số thông qua cơ chế lan truyền ngược như đã trình bày ở chương trước. Do phép toán được thực hiện trên tập huấn luyện nên sẽ gây ra vấn đề overfit, vì thế mà tập validation sẽ được sử dụng để đo lại giá trị sai lệch và độ chính xác. Nếu độ chính xác tại tập huấn luyện cao nhưng tại tập validation không thay đổi hoặc thấp thì chúng ta mô hình mạng gặp phải vấn đề overfit và việc huấn luyện tiếp tục không còn có ích.

¹Tập dữ liệu lúc này là những bottlenecks

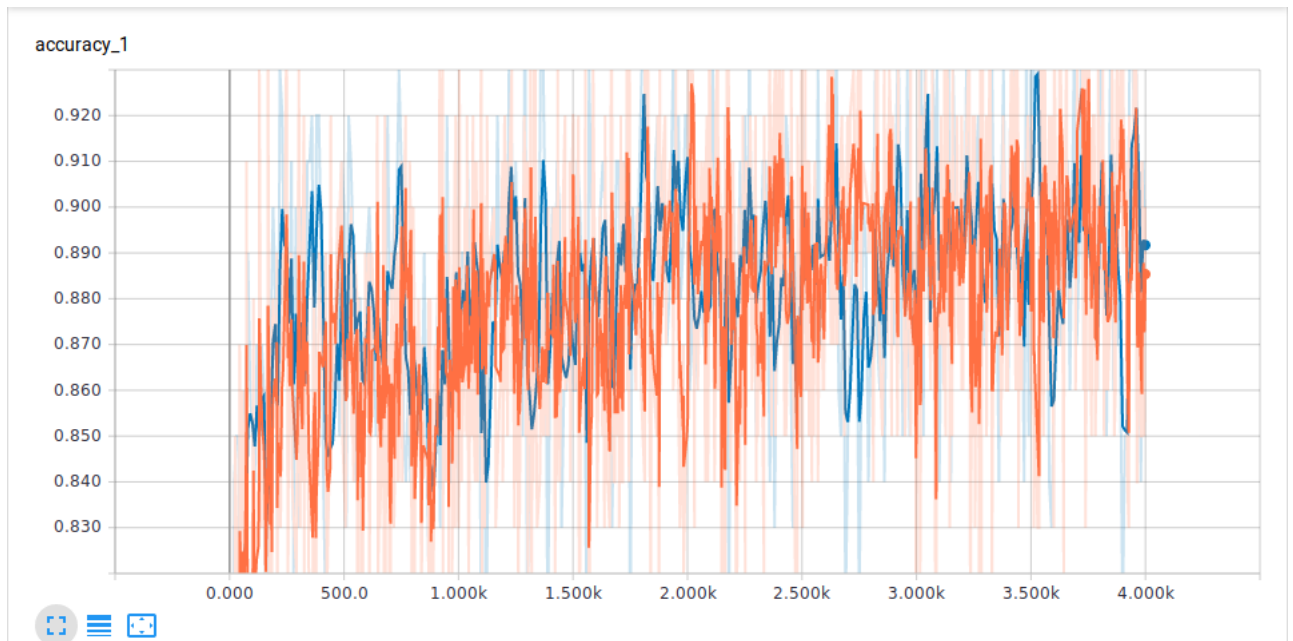
²Tầng fully-connected với softmax là activation function

Chương 5

Kết quả huấn luyện và đánh giá

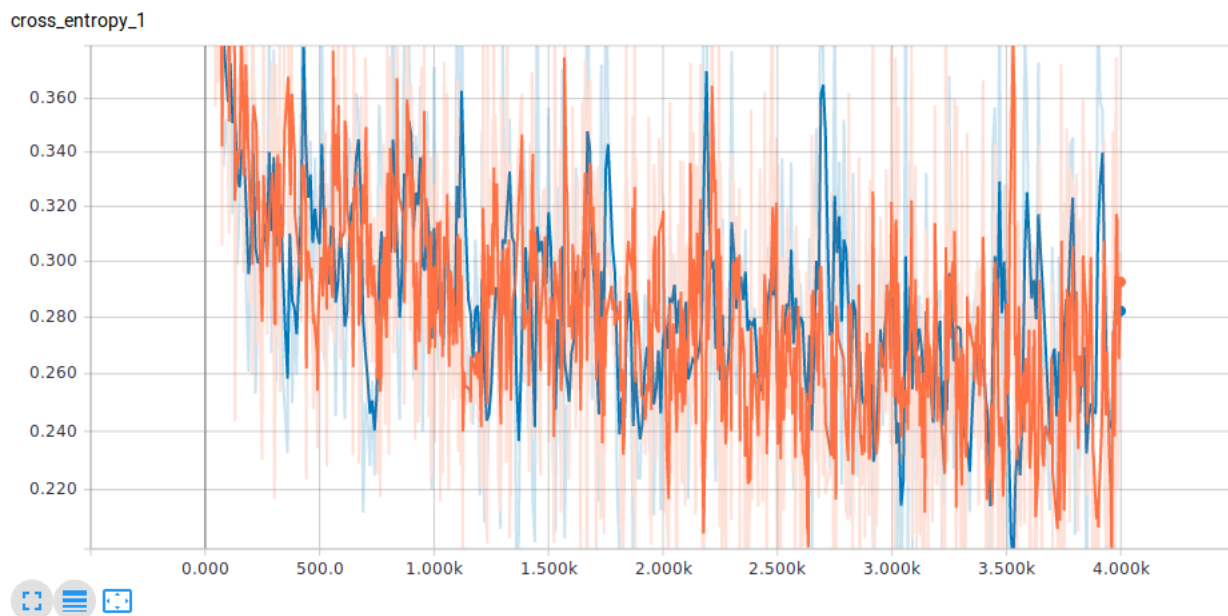
5.1 Kết quả huấn luyện

Với các thông số cà mạng đã cấu hình ở chương trước, ?? cho thấy kết quả huấn sau khi huấn luyện mạng có kết quả tương đối tốt với giá trị $cross_entropy = 0.29$ và độ chính xác ở tập kiểm thử là 88.7%. Không những thế việc áp dụng kỹ thuật transfer learning giúp tiết kiệm thời gian huấn luyện gấp nhiều lần. Cụ thể, việc xây dựng mô hình thủ công và huấn luyện lại từ đầu đối với máy tính không hỗ trợ card GPU sẽ mất thời gian là 8 tiếng nhưng đối với việc áp dụng mô hình inception model với kỹ thuật transfer learning thì chỉ mất từ 45 đến 60 phút để hoàn thành.



Hình 5.1: Đồ thị biểu diễn độ chính xác (accuracy)

Đồ thị 5.1 biểu diễn chi tiết độ chính xác trong quá trình huấn luyện với đường màu cam biểu diễn cho quá trình huấn luyện và đường màu xanh lam biểu diễn cho kiểm tra trên tập Validation. Tương tự với 5.2.



Hình 5.2: Đồ thị biểu diễn độ sai số (cross entropy)

5.2 Sử dụng mô hình phân loại ảnh mới

Sử dụng mô hình để phân loại một số hình ảnh khác chưa được phân loại. 5.3 và 5.4 có thể nhận biết bằng mắt thường đây là hình ảnh của những con đường đang trong tình trạng ùn tắc. Kết quả dự đoán cho lớp cả hai hình ảnh này là chính xác theo lớp kẹt xe.



Hình 5.3: Ảnh kiểm thử 1 - kết quả



Hình 5.4: Ảnh kiểm thử 2 - kết quả

Đối với hai hình ảnh tiếp theo được phân biệt vào loại đường thông thoáng một cách dễ dàng. Mô hình dự đoán xác suất hai hình ảnh dưới đây thuộc lớp thông thoáng lần lượt là.



Hình 5.5: Ảnh kiểm thử 3 - kết quả



Hình 5.6: Ảnh kiểm thử 4 - kết quả

Như vậy, kết quả kiểm thử đối với một số hình ảnh chưa được phân loại của mô hình đã huấn luyện cho kết quả khá chính xác. Tuy nhiên, đối với vấn đề phân loại giao thông thì không chỉ có hai trường hợp ùn tắc hay thông thoáng mà còn tồn tại nhiều trường hợp hơn. Như tình huống đông xe nhưng di chuyển chậm, trên thực tế đây là tình huống không phải ùn tắc nhưng ảnh chụp gần giống với ảnh ùn tắc. Cần chú ý vấn đề này khi lựa chọn, phân loại ảnh huấn luyện, hoặc để giải quyết tốt hơn cần phải tăng số lượng lớp(nhãn) cần phân loại lên thành 3 hay lớn hơn thay vì 2 như ban đầu.

Chương 6

Kết luận

6.1 Các kết quả

Luận văn này giúp nghiên cứu, tìm hiểu về mạng học sâu cũng như việc phân loại ảnh giao thông. Các kết quả đã đạt được đáp ứng được các mục tiêu đã đặt ra.

- Xây dựng được mô hình HDFS sử dụng Apache Hadoop có thể lưu trữ dữ liệu video tập dữ liệu ảnh
- Xây dựng được bộ dữ liệu phục vụ cho việc huấn luyện mô hình phân biệt 2 loại ảnh giao thông. Số lượng ảnh thu thập được trung bình mỗi lớp khoảng 1000 cho đến 2000 ảnh.
- Huấn luyện thành công mô hình học sâu GoogleNet sử dụng tập dữ liệu được lưu trữ từ hệ thống HDFS.
- Xác định được vị trí và cho kết quả tình hình giao thông.

6.2 Hướng phát triển

Sau khi đạt được kết quả huấn luyện khá tốt, hướng phát triển của đề tài trong tương lai như sau:

1. Tiếp tục xem xét việc xác định các lớp ảnh giả thông trong tương lai giúp phát hiện nhiều loại hình như ùn tắc, thông thoáng, đông xe di chuyển chậm,.v.v. giúp cụ thể hóa tình trạng giao thông.
2. Xây dựng trang web giúp nhận biết kẹt xe.
3. Có thể sử dụng nền tảng đã xây dựng để áp dụng cho nhiều bài toán phân loại ảnh khác.

Hệ thống tiềm năng trên có thể giúp dân cư sinh sống ở các khu đô thị cũng như ban quản lý nắm bắt tình hình giao thông để phân luồng di chuyển và khắc phục một cách nhanh nhất.

Tài liệu tham khảo

- [1] Rémi Cadène, Nicolas Thome, and Matthieu Cord. Master's Thesis : Deep Learning for Visual Recognition. 2016.
- [2] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep Learning. Chapter 6:Deep Feedforward Networks.
- [3] Andrej Karpathy. CS231n: Convolutional Neural Networks for Visual Recognition. CNN overview, 2016.
- [4] Andrej Karpathy. CS231n: Convolutional Neural Networks for Visual Recognition. convolutional layers, 2016.
- [5] Andrej Karpathy. CS231n: Convolutional Neural Networks for Visual Recognition. CNN architecture, 2016.
- [6] Andrej Karpathy. CS231n: Convolutional Neural Networks for Visual Recognition. transfer learning, 2016.
- [7] Andrej Karpathy. CS231n: Convolutional Neural Networks for Visual Recognition. module 1 - Neural Networks and module 2 - Convolutional Neural Network, 2016.

- [8] Min Lin, Qiang Chen, and Shuicheng Yan. Network in network. 2013.
- [9] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed and Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. 2014.
- [10] tensorflow.org. Image retraining.
- [11] tensorflow.org. Tensorflow programmer guide.
- [12] Tom White. *Hadoop: The Definitive Guide*. oreilly.
- [13] Matthew D. Zeiler and Rob Fergus. Visualizing and understanding convolutional networks. 2013.