

**TRƯỜNG ĐẠI HỌC ĐIỆN LỰC**

**KHOA CÔNG NGHỆ THÔNG TIN**

**BÁO CÁO CHUYÊN ĐỀ HỌC PHẦN**

**NHẬP MÔN HỌC MÁY**

**ĐỀ TÀI:**

**NHẬN DẠNG HOA QUẢ BẰNG THUẬT TOÁN CNN**

|  |  |
| --- | --- |
| **Sinh viên thực hiện** | **: NGUYỄN VIỆT HOÀNG**  **ĐẶNG TUẤN NINH**  **BÙI HẢI ANH** |
| **Giảng viên hướng dẫn** | **: VŨ VĂN ĐỊNH** |
| **Ngành** | **: CÔNG NGHỆ THÔNG TIN** |
| **Chuyên ngành** | **: CÔNG NGHỆ PHẦN MỀM** |
| **Lớp** | **: D16CNPM2** |
| **Khóa** | **: D16** |

***Hà Nội, tháng .. năm ....***1

**PHIẾU CHẤM ĐIỂM**

Sinh viên thực hiện:

|  |  |  |
| --- | --- | --- |
| **Họ và tên** | **Chữ ký** | **Ghi chú** |
| Đặng Tuấn Ninh  Mã sinh viên: 21810310377 |  |  |
| Nguyễn Việt Hoàng Mã sinh viên: 21810310648 |  |  |
| Bùi Hải Anh  Mã sinh viên: |  |  |

Giảng viên chấm:

|  |  |  |
| --- | --- | --- |
| **Họ và tên** | **Chữ ký** | **Ghi chú** |
| Giảng viên chấm 1 : |  |  |
| Giảng viên chấm 2 : |  |  |

**MỤC LỤC**

[LỜI MỞ ĐẦU 4](#_Toc163153220)

[CHƯƠNG 1 : TỔNG QUAN VỀ HỌC MÁY 5](#_Toc163153221)

[1.1 Khái niệm về học máy: 5](#_Toc163153222)

[1.2 Phân nhóm các thuật toán học máy 5](#_Toc163153223)

[1.2.1 Học có giám sát (Supervised Learning) 5](#_Toc163153224)

[1.2.2 Học phi giám sát (Unsupervised Learning)-UL 6](#_Toc163153225)

[1.2.3 Học tăng cường (Reinforcement Learning) 6](#_Toc163153226)

[1.2.4 Học bán giám sát (Semi-Supervised Learning) 7](#_Toc163153227)

[1.3 Ứng dụng của học máy 7](#_Toc163153228)

[CHƯƠNG 2: TỔNG QUAN VỀ CNN 9](#_Toc163153229)

[2.1 CNN (Convolutional Neural Network) là gì? 9](#_Toc163153230)

[2.2 Tìm hiểu Convolutional là gì? 9](#_Toc163153231)

[2.3 Cấu trúc mạng CNN 10](#_Toc163153232)

[2.4 Trường tiếp nhận cục bộ (Local Receptive Field) 11](#_Toc163153233)

[2.5 Trọng số chia sẻ (Shared Weight and Bias) 13](#_Toc163153234)

[2.6 Lớp tổng hợp (Pooling Layer) 13](#_Toc163153235)

[2.7 Cách chọn tham số cho CNN 15](#_Toc163153236)

[CHƯƠNG 3: ỨNG DỤNG CONVOLUTIONAL NEURAL NETWORK CHO BÀI TOÁN NHẬN DIỆN HOA QUẢ 16](#_Toc163153237)

[3.1 Cơ sở dữ liệu huấn luyện 16](#_Toc163153238)

[3.2 Thực hiện CNN trong bài toán nhận diện hoa quả 16](#_Toc163153239)

[KẾT LUẬN 19](#_Toc163153240)

**DANH MỤC HÌNH ẢNH**

[Hình 1: Áp dụng phép tích chập lên hình ảnh 9](#_Toc163154593)

[Hình 2: Quá trình xử lý hình ảnh trong mạng neural tích chập 10](#_Toc163154594)

[Hình 3: Mô hình đơn giản của mạng neural 12](#_Toc163154595)

[Hình 4: Kết nối giữa nhiều neuron đầu vào và một neuron ẩn 12](#_Toc163154596)

[Hình 5: Tạo neuron ẩn thứ 2 13](#_Toc163154597)

[Hình 6: Các lớp ẩn 13](#_Toc163154598)

[Hình 7: Kết nối giữa các neuron đầu vào và lớp ẩn đầu tiên 14](#_Toc163154599)

[Hình 8: Kiến trúc của mạng neural tích chập 15](#_Toc163154600)

[Hình 9: Đoạn code phần huấn luyện 17](#_Toc163154601)

[Hình 10: Xây dựng và huấn luyện mô hình CNN để phân loại hình ảnh 18](#_Toc163154602)

[Hình 11: Ứng dụng dự đoán trên webcam sử dụng mô hình đã huấn luyện 18](#_Toc163154603)

# **LỜI MỞ ĐẦU**

Trong thời đại hiện nay, công nghệ ngày càng trở nên phổ biến và không ai có thể phủ nhận tầm quan trọng và những lợi ích mà nó mang lại cho cuộc sống của chúng ta. Cùng với sự phát triển đó, học máy (machine learning) đã trở thành một trong những đề tài thu hút sự quan tâm của nhiều người. Học máy không chỉ đơn giản là một công nghệ mới mẻ, mà còn là một công cụ mạnh mẽ giúp con người làm việc và hoàn thành nhiệm vụ một cách hiệu quả hơn. Với khả năng "học hỏi" từ dữ liệu và tự động tạo ra các mô hình, machine learning đã mở ra một thế giới mới của tiềm năng và cơ hội trong mọi lĩnh vực.

Tuy không phải là một lĩnh vực mới, nhưng machine learning đang ngày càng thu hút sự quan tâm và đầu tư, chứng tỏ sức mạnh và tiềm năng của nó trong tương lai. Đặc biệt, sự quan tâm này càng được nâng cao khi thấy được những lợi ích rõ ràng mà machine learning mang lại trong công việc và cuộc sống hàng ngày.

Nhận thấy được sự hữu dụng của học máy trong công việc và các lợi ích vượt bậc mà nó đem lại cho chúng ta nên chúng em đã chọn đề tài “ Nhận dạng hoa quả bằng thuật toán CNN” để làm báo cáo

Chúng em xin gửi lời cảm ơn tới các thầy cô giáo trong Trường Đại học Điện Lực nói chung và các thầy cô giáo trong Khoa Công nghệ thông tin nói riêng đã tận tình giảng dạy, truyền đạt cho chúng em những kiến thức cũng như kinh nghiệm quý báu trong suốt quá trình học. Đặc biệt, em gửi lời cảm ơn đến thầy Vũ Văn Định đã tận tình theo sát giúp đỡ, trực tiếp chỉ bảo, hướng dẫn trong suốt quá trình nghiên cứu và học tập của chúng em.

Chúng em xin chân thành cảm ơn!

# **CHƯƠNG 1 : TỔNG QUAN VỀ HỌC MÁY**

## **1.1 Khái niệm về học máy:**

Học Máy (Machine Learning) là một lĩnh vực con của Trí Tuệ Nhân Tạo (Artificial Intelligence), tập trung vào việc phát triển các thuật toán và mô hình để máy tính có thể học từ dữ liệu một cách tự động và cải thiện hiệu suất của mình theo thời gian. Thay vì được lập trình trực tiếp, máy tính được cung cấp với các cơ chế tự học và tự điều chỉnh từ dữ liệu, giúp chúng thực hiện các nhiệm vụ và ra các dự đoán mà không cần sự can thiệp trực tiếp từ con người.

Mặc dù Học Máy có thể tự học, nhưng vẫn cần sự can thiệp của con người trong việc hiểu và xử lý dữ liệu, lựa chọn các thuật toán và kiểm tra hiệu suất của các mô hình đã huấn luyện. Trước khi sử dụng, dữ liệu cần được làm sạch và không chứa sai sót hoặc dữ liệu giả

Để đạt được kết quả tốt trong Học Máy, một lượng lớn dữ liệu là cần thiết để huấn luyện và đánh giá các mô hình. Trong quá khứ, việc thiếu hụt dữ liệu đã làm hạn chế hiệu suất của các thuật toán Học Máy. Tuy nhiên, với sự phát triển của dữ liệu lớn (big data), nguồn cung cấp dữ liệu đủ lớn đã giúp cải thiện đáng kể độ chính xác và khả năng dự đoán của các mô hình Học Máy.

## **1.2 Phân nhóm các thuật toán học máy**

### **1.2.1 Học có giám sát (Supervised Learning)**

Là phương pháp sử dụng những dữ liệu đã được gán nhãn từ trước để suy luận ra quan hệ giữa đầu vào và đầu ra. Các dữ liệu này được gọi là dữ liệu huấn luyện và chúng là cặp các đầu vào-đầu ra. Học có giám sát sẽ xem xét các tập huấn luyện này để từ đó có thể đưa ra dự đoán đầu ra cho 1 đầu vào mới chưa gặp bao giờ. Ví dụ dự đoán giá nhà, phân loại email. Các mô hình như mạng Nơ-ron, SVM,CNN,…

Supervised Learning là thuât toán dự đoán đầu ra (output) của một dữ liệu mới (new input) dựa trên cặp (input, output) đã biết từ trước. Cặp dữ liệu này còn đươc gọi là data, label tức dữ liệu, nhãn. Supervised Learning là nhóm phổ biến nhất trong các thuật toán Machine Learning.

### **1.2.2 Học phi giám sát (Unsupervised Learning)-UL**

Khác với học có giám sát, học phi giám sát sử dụng những dữ liệu chưa được gán nhãn từ trước để suy luận. Phương pháp này thường được sử dụng để tìm cấu trúc của tập dữ liệu. Tuy nhiên lại không có phương pháp đánh giá được cấu trúc tìm ra được là đúng hay sai. Ví dụ như phân cụm dữ liệu, triết xuất thành phần chính của một chất nào đó. Ứng dụng phổ biến nhất của học không giám sát là gom cụm (cluster).

Trong thuật toán này, chúng ta không biết được dữ liệu đầu ra hay nhãn mà chỉ có dữ liệu đầu vào. Thuật toán Học không giám sát dựa vào cấu trúc của dữ liệu để thực hiện một công việc nào đó, ví dụ như phân nhóm hoặc giảm số chiều của dữ liệu để thuận tiện trong việc lưu trữ và tính toán.

Một cách toán học, Học không giám sát là khi chúng ta chỉ có dữ liệu vào X mà không biết nhãn Y tương ứng.

Sự khác nhau giữa học có giám sát và học không giám sát:

Học có giám sát: Là cách huấn luyện một mô hình trong đó dữ liệu học có đầu vào và đầu ra tương ứng đầu vào đó. Mô hình được huấn luyện bằng cách giảm thiểu sai số lỗi (loss) của các dự đoán tại các vòng lặp huấn luyện. Sau quá trình huấn luyện. mô hình sẽ có khả năng đưa ra dự đoán về đầu ra với một đầu vào mới gặp (không có trong dữ liệu học). Nếu không gian đầu ra được biểu diễn dứới dạng rời rạc, ta gọi đó là bài toán phân loại (classification). Nếu không gian đầu ra được biểu diễn dưới dạng liên tục, ta gọi đó là bài toán hồi quy (regression).

Học không giám sát: Là cách huấn luyện một mô hình trong đó dữ liệu học chỉ bao gồm đầu vào mà không có đầu ra. Mô hình sẽ được huấn luyện cách để tìm cấu trúc hoặc mối quan hệ giữa các đầu vào. Một trong những phương pháp học không giám sát quan trọng nhất là phân cụm (clustering): Tạo các cụm khác nhau với mỗi cụm biểu diễn một đặc trưng nào đó của dữ liệu và phân các đầu vào mới vào các cụm theo các đặc trưng của đầu vào đó. Các phương pháp học không giám sát khác có thể kể đến như: phát hiện điểm bất thường (Anomaly Detection), Singular-value Decomposition,…

### **1.2.3 Học tăng cường (Reinforcement Learning)**

Phương pháp học tăng cường tập trung vào việc làm sao để cho 1 tác tử trong môi trường có thế hành động sao cho lấy được phần thưởng nhiều nhất có thể. Khác với học có giám sát nó không có cặp dữ liệu gán nhãn trước làm đầu vào và cũng không có đánh giá các hành động là đúng hay sai.

### **1.2.4 Học bán giám sát (Semi-Supervised Learning)**

Các bài toán khi chúng ta có một lượng lớn dữ liệu X nhưng chỉ một phần trong chúng được gán nhãn được gọi là Semi-Supervised Learning. Những bài toán thuộc nhóm này nằm giữa hai nhóm được nêu bên trên. Một ví dụ điển hình của nhóm này là chỉ có một phần ảnh hoặc văn bản được gán nhãn (ví dụ bức ảnh về người, động vật hoặc các văn bản khoa học, chính trị) và phần lớn các bức ảnh/văn bản khác chưa được gán nhãn được thu thập từ internet Thực tế cho thấy rất nhiều các bài toán Machine Learning thuộc vào nhóm này vì việc thu thập dữ liệu có nhãn tốn rất nhiều thời gian và có chi phí cao. Rất nhiều loại dữ liệu thậm chí cần phải có chuyên gia mới gán nhãn được (ảnh y học chẳng hạn). Ngược lại, dữ liệu chưa có nhãn có thể được thu thập với chi phí thấp từ internet.

## **1.3 Ứng dụng của học máy**

Nhiều hoạt động hàng ngày của chúng ta được trợ giúp bởi các thuật toán Machine Learning, bao gồm:

* Trong y tế: được sử dụng để giúp trong việc chuẩn đoán bệnh thông qua phân loại hình ảnh y tế và dự đoán kết quả điều trị dựa trên dữ liệu bệnh nhân.
* Trong lĩnh vực ngân hàng: phân tích rủi ro tín dụng, phát hiện gian lận tài khoản, cung cấp hỗ trợ tự động cho khách hàng qua chatbot, và dự đoán xu hướng thị trường để quản lý rủi ro đầu tư.
* Trong giáo dục: cá nhân hóa quá trình học, đưa ra đề xuất giáo trình và tài liệu phù hợp với nhu cầu của từng học sinh. Nó cũng được áp dụng để đánh giá hiệu suất học tập của học sinh dựa trên dữ liệu từ bài kiểm tra và bài tập, giúp giáo viên hiểu rõ hơn về tiến độ học tập của học sinh và điều chỉnh phương pháp giảng dạy một cách hiệu quả.
* Trong thương mại điện tử: tăng cường trải nghiệm mua sắm của khách hàng thông qua việc cá nhân hóa gợi ý sản phẩm, dự đoán xu hướng mua hàng, và tối ưu hóa quá trình tìm kiếm sản phẩm. Nó cũng giúp trong việc phát hiện gian lận trong các giao dịch trực tuyến và tối ưu hóa quy trình phân phối và quản lý tồn kho.
* Trong kinh tế nói chung: dự đoán xu hướng thị trường, tối ưu hóa quy trình sản xuất và chuỗi cung ứng, phân tích rủi ro tài chính và quản lý danh mục đầu tư, cũng như trong quản lý rủi ro để đưa ra các biện pháp phòng ngừa kịp thời.

# **CHƯƠNG 2: TỔNG QUAN VỀ CNN**

## **2.1 CNN (Convolutional Neural Network) là gì?**

Convolutional Neural Network (CNNs) là một loại mô hình Deep Learning được sử dụng rộng rãi trong việc nhận dạng hình ảnh và các vấn đề thị giác máy tính khác. CNN tự động học và trích xuất các đặc trưng từ dữ liệu hình ảnh, giúp nâng cao độ chính xác và hiệu suất của các ứng dụng liên quan đến xử lý hình ảnh.

## **2.2 Tìm hiểu Convolutional là gì?**

Là một cửa sổ trượt (Sliding Windows) trên một ma trận như mô tả hình dưới:

**A colorful squares with black text

Description automatically generated**

Hình 1: Áp dụng phép tích chập lên hình ảnh

Các convolutional layer có các parameter(kernel) đã được học để tự điều chỉnh lấy ra những thông tin chính xác nhất mà không cần chọn các feature.

Trong hình ảnh ví dụ trên, ma trận bên trái là một hình ảnh trắng đen được số hóa. Ma trận có kích thước 5×5 và mỗi điểm ảnh có giá trị 1 hoặc 0 là giao điểm của dòng và cột.

Convolution hay tích chập là nhân từng phần tử trong ma trận 3. Sliding Window hay còn gọi là kernel, filter hoặc feature detect là một ma trận có kích thước nhỏ như trong ví dụ trên là 3×3.

Convolution hay tích chập là nhân từng phần tử bên trong ma trận 3×3 với ma trận bên trái. Kết quả được một ma trận gọi là Convoled feature được sinh ra từ việc nhận ma trận Filter với ma trận ảnh 5×5 bên trái.

## **2.3 Cấu trúc mạng CNN**

Mạng CNN là một tập hợp các lớp Convolution chồng lên nhau và sử dụng các hàm nonlinear activation như ReLU và tanh để kích hoạt các trọng số trong các node. Mỗi một lớp sau khi thông qua các hàm kích hoạt sẽ tạo ra các thông tin trừu tượng hơn cho các lớp tiếp theo.

Mỗi một lớp sau khi thông qua các hàm kích hoạt sẽ tạo ra các thông tin trừu tượng hơn cho các lớp tiếp theo. Trong mô hình mạng truyền ngược (feedforward neural network) thì mỗi neural đầu vào (input node) cho mỗi neural đầu ra trong các lớp tiếp theo.

Mô hình này gọi là mạng kết nối đầy đủ (fully connected layer) hay mạng toàn vẹn (affine layer). Còn trong mô hình CNNs thì ngược lại. Các layer liên kết được với nhau thông qua cơ chế convolution.

Layer tiếp theo là kết quả convolution từ layer trước đó, nhờ vậy mà ta có được các kết nối cục bộ. Như vậy mỗi neuron ở lớp kế tiếp sinh ra từ kết quả của filter áp đặt lên một vùng ảnh cục bộ của neuron trước đó.

Mỗi một lớp được sử dụng các filter khác nhau thông thường có hàng trăm hàng nghìn filter như vậy và kết hợp kết quả của chúng lại. Ngoài ra có một số layer khác như pooling/subsampling layer dùng để chắt lọc lại các thông tin hữu ích hơn (loại bỏ các thông tin nhiễu).

Trong quá trình huấn luyện mạng (traning) CNN tự động học các giá trị qua các lớp filter dựa vào cách thức mà bạn thực hiện. Ví dụ trong tác vụ phân lớp ảnh, CNNs sẽ cố gắng tìm ra thông số tối ưu cho các filter tương ứng theo thứ tự raw pixel > edges > shapes > facial > high-level features. Layer cuối cùng được dùng để phân lớp ảnh.



Hình 2: Quá trình xử lý hình ảnh trong mạng neural tích chập

Trong mô hình CNN có 2 khía cạnh cần quan tâm là tính bất biến (Location Invariance) và tính kết hợp (Compositionality). Với cùng một đối tượng, nếu đối tượng này được chiếu theo các gốc độ khác nhau (translation, rotation, scaling) thì độ chính xác của thuật toán sẽ bị ảnh hưởng đáng kể.

Pooling layer sẽ cho bạn tính bất biến đối với phép dịch chuyển (translation), phép quay (rotation) và phép co giãn (scaling). Tính kết hợp cục bộ cho ta các cấp độ biểu diễn thông tin từ mức độ thấp đến mức độ cao và trừu tượng hơn thông qua convolution từ các filter.

Đó là lý do tại sao CNNs cho ra mô hình với độ chính xác rất cao. Cũng giống như cách con người nhận biết các vật thể trong tự nhiên.

Mạng CNN sử dụng 3 ý tưởng cơ bản:

* Các trường tiếp nhận cục bộ (local receptive field)
* Trọng số chia sẻ (shared weights)
* Tổng hợp (pooling).

## **2.4 Trường tiếp nhận cục bộ (Local Receptive Field)**

Đầu vào của mạng CNN là một ảnh. Ví dụ như ảnh có kích thước 28×28 thì tương ứng đầu vào là một ma trận có 28×28 và giá trị mỗi điểm ảnh là một ô trong ma trận. Trong mô hình mạng ANN truyền thống thì chúng ta sẽ kết nối các neuron đầu vào vào tầng ảnh.

Tuy nhiên trong CNN chúng ta không làm như vậy mà chúng ta chỉ kết nối trong một vùng nhỏ của các neuron đầu vào như một filter có kích thước 5×5 tương ứng (28- 5 + 1) 24 điểm ảnh đầu vào. Mỗi một kết nối sẽ học một trọng số và mỗi neuron ẩn sẽ học một bias. Mỗi một vùng 5×5 đấy gọi là một trường tiếp nhận cục bộ.

A diagram of a brain function

Description automatically generated

Hình 3: Mô hình đơn giản của mạng neural

Một cách tổng quan, ta có thể tóm tắt các bước tạo ra 1 hidden layer bằng các cách sau:

* Tạo ra neuron ẩn đầu tiên trong lớp ẩn 1

A diagram of a person's brain

Description automatically generated with medium confidence

Hình 4: Kết nối giữa nhiều neuron đầu vào và một neuron ẩn

* Dịch filter qua bên phải một cột sẽ tạo được neuron ẩn thứ 2.

A diagram of a number of objects

Description automatically generated with medium confidence

Hình 5: Tạo neuron ẩn thứ 2

Với bài toán nhận dạng ảnh người ta thường gọi ma trận lớp đầu vào là feature map, trọng số xác định các đặc trương là shared weight và độ lệch xác định một feature map là shared bias. Như vậy đơn giản nhất là qua các bước trên chúng ta chỉ có 1 feature map. Tuy nhiên trong nhận dạng ảnh chúng ta cần nhiều hơn một feature map.

A diagram of a machine learning

Description automatically generated

Hình 6: Các lớp ẩn

Như vậy, local receptive field thích hợp cho việc phân tách dữ liệu ảnh, giúp chọn ra những vùng ảnh có giá trị nhất cho việc đánh giá phân lớp.

## **2.5 Trọng số chia sẻ (Shared Weight and Bias)**

Đầu tiên, các trọng số cho mỗi filter (kernel) phải giống nhau. Tất cả các nơ-ron trong lớp ẩn đầu sẽ phát hiện chính xác feature tương tự chỉ ở các vị trí khác nhau trong hình ảnh đầu vào. Chúng ta gọi việc map từ input layer sang hidden layer là một feature map. Vậy mối quan hệ giữa số lượng Feature map với số lượng tham số là gì?

Tóm lại, một convolutional layer bao gồm các feature map khác nhau. Mỗi một feature map giúp detect một vài feature trong bức ảnh. Lợi ích lớn nhất của trọng số chia sẻ là giảm tối đa số lượng tham số trong mạng CNN.

## **2.6 Lớp tổng hợp (Pooling Layer)**

Lớp pooling thường được sử dụng ngay sau lớp convulational để đơn giản hóa thông tin đầu ra để giảm bớt số lượng neuron. Thủ tục pooling phổ biến là max-pooling, thủ tục này chọn giá trị lớn nhất trong vùng đầu vào 2×2.

A diagram of a graph

Description automatically generated

Hình 7: Kết nối giữa các neuron đầu vào và lớp ẩn đầu tiên

Như vậy qua lớp Max Pooling thì số lượng neuron giảm đi phân nửa. Trong một mạng CNN có nhiều Feature Map nên mỗi Feature Map chúng ta sẽ cho mỗi Max Pooling khác nhau. Chúng ta có thể thấy rằng Max Pooling là cách hỏi xem trong các đặc trưng này thì đặc trưng nào là đặc trưng nhất. Ngoài Max Pooling còn có L2 Pooling.

Cuối cùng ta đặt tất cả các lớp lại với nhau thành một CNN với đầu ra gồm các neuron với số lượng tùy bài toán.

A diagram of a block diagram

Description automatically generated

Hình 8: Kiến trúc của mạng neural tích chập

2 lớp cuối cùng của các kết nối trong mạng là một lớp đầy đủ kết nối (fully connected layer) . Lớp này nối mọi nơron từ lớp max pooled tới mọi nơron của tầng ra.

## **2.7 Cách chọn tham số cho CNN**

* Số các convolution layer: càng nhiều các convolution layer performance càng được cải thiện. Sau khoảng 3 hoặc 4 layer, các tác động được giảm một cách đáng kể
* Filter size: thường filter theo size 5×5 hoặc 3×3
* Pooling size: thường là 2×2 hoặc 4×4 cho ảnh đầu vào lớn
* Cách cuối cùng là thực hiện nhiều lần việc train test để chọn ra được param tốt nhất.

# **CHƯƠNG 3: ỨNG DỤNG CONVOLUTIONAL NEURAL NETWORK CHO BÀI TOÁN NHẬN DIỆN HOA QUẢ**

## **3.1 Cơ sở dữ liệu huấn luyện**

- Dữ liệu huấn luyện gồm 586 ảnh hoa quả trong đó gồm có:

* 175 ảnh quả chuối
* 92 ảnh quả dâu tây
* 74 ảnh quả dứa
* 88 ảnh quả khế
* 80 ảnh quả măng cụt
* 77 ảnh quả xoài

- Dữ liệu test gồm 586 ảnh hoa quả trong đó gồm có:

* 175 ảnh quả chuối
* 92 ảnh quả dâu tây
* 74 ảnh quả dứa
* 88 ảnh quả khế
* 80 ảnh quả măng cụt
* 77 ảnh quả xoài

## **3.2 Thực hiện CNN trong bài toán nhận diện hoa quả**

Bước đầu chúng ta sẽ thực hiện viết code để xây dựng và huấn luyện một mô hình để phân loại hình ảnh thành 6 lớp khác nhau.

A computer screen shot of a program

Description automatically generated

Hình 9: Đoạn code phần huấn luyện

A screen shot of a computer code

Description automatically generated

Hình 10: Xây dựng và huấn luyện mô hình CNN để phân loại hình ảnh

Sau dó chúng ta sử dụng mô hình đã được huấn luyện để phân loại các đối tượng trong hình ảnh từ webcam.

A computer screen shot of a program code

Description automatically generated

Hình 11: Ứng dụng dự đoán trên webcam sử dụng mô hình đã huấn luyện

# **KẾT LUẬN**

Với sự tiến bộ ngày càng nhanh chóng của trí tuệ nhân tạo, việc ứng dụng học máy vào các công việc hàng ngày trở thành một lựa chọn hết sức thích hợp. Thông qua quá trình tìm hiểu và khảo sát, em đã lựa chọn và triển khai báo cáo “Nhận dạng hoa quả bằng thuật toán CNN”.

Tích lũy kiến thức cơ bản cùng với sự hướng dẫn tận tình từ thầy Vũ Văn Định, em đã hoàn thành đề tài và hiểu rõ hơn về quy trình và cách thức áp dụng thuật toán CNN vào nhận dạng hoa quả.Mặc dù hệ thống đã được hoàn thành, nhưng vẫn tồn tại nhiều hạn chế và vấn đề chưa được giải quyết. Cụ thể là nhận dạng hoa quả vẫn còn nhầm, chưa được chuẩn hóa và chưa đạt đến mức hoàn thiện mong muốn. Em rất mong nhận được sự đóng góp từ thầy cô và các bạn để trong thời gian sắp tới, chúng em sẽ nỗ lực khắc phục những hạn chế và vấn đề tồn đọng, đồng thời bổ sung thêm các chức năng mới cho hệ thống.

Em xin chân thành đến thầy cô và các bạn đã hỗ trợ và đóng góp ý kiến. Đồng thời, em sẽ không ngừng phát triển và cải tiến hệ thống để đáp ứng tốt nhất nhu cầu và mong muốn của người sử dụng.

**TÀI LIỆU THAM KHẢO**

- Nhận dạng chữ số viết tay sử dụng kỹ thuật học sâu (deep learning) – Huỳnh Văn Nhựt – chuyên ngành : khoa học máy tính – Tóm tắt luận văn thạc sĩ – Trường đại học Bách khoa Đà Nằng

<http://tainguyenso.dut.udn.vn/bitstream/DUT/485/1/HuynhVanNhut.TT.pdf.pdf>

- Tìm hiểu về Convolutional Neural Network và làm một ví dụ nhỏ về phân loại ảnh – Trần Đức Trung – Viblo

<https://viblo.asia/p/tim-hieu-ve-convolutional-neural-network-va-lam-mot-vi-du-nho-ve-phan-loai-anh-aWj53WXo56m>

- Nhận dạng ảnh chó mèo bằng Neural Network – Võ Hùng Vĩ - Nhận dạng hình ảnh – Giải thuật – Home – Thị giác máy tính

<https://thigiacmaytinh.com/nhan-dang-anh-cho-meo-bang-neural-network/>