**TRƯỜNG ĐẠI HỌC PHENIKAA**

**KHOA CÔNG NGHÊ THÔNG TIN**

**⸎⸎⸎⸎⸎**

Logo, company name

Description automatically generated

**BÁO CÁO MÔN THỊ GIÁC MÁY TÍNH**

**NHẬN DIỆN CÁC BỆNH TRÊN LÁ CÂY BẰNG MẠNG NƠ RƠN TÍCH CHẬP (CNN)**

**Danh sách thành viên:**

1. **Bùi Huy Bách**
2. **Lê Toàn Đức**
3. **Nguyễn Minh Hiếu**
4. **Phạm Thành Trung**
5. **Nguyễn Hữu Tài**

**Phần 1: Đặt vấn đề**

* 1. Giới thiệu

Bệnh học thực vật là một lĩnh vực chuyên nghiên cứu về nguyên nhân, triệu chứng, bệnh trạng của thực vật, dựa vào đó đưa ra các biện pháp phòng trừ hiệu quả nhất.

Thực vật bị bệnh là hiện tượng khi thực vật không đủ khả năng thích ứng với những biến đổi của môi trường hoặc với những kích thích của những sinh vật khác, làm đảo lộn hoạt động sinh lý, gây bất lợi đến sinh trưởng hoặc làm thực vật bị chết, nếu trên diện lớn có thể dẫn đến tổn thất về kinh tế và sinh thái.

Tùy thuộc vào cách phân loại các loại cây trồng theo mục đích mà người ta phân thành một số lĩnh vực nhỏ trong bệnh học thực vật:

* Bệnh cây nông nghiệp: Những bệnh liên quan, ảnh hưởng đến sinh trưởng và phát triển của các cây trồng sản xuất chính trong nông nghiệp, làm tổn thất về kinh tế [nông nghiệp](https://vi.wikipedia.org/wiki/N%C3%B4ng_nghi%E1%BB%87p).
* Bệnh cây lâm nghiệp:Bệnh hại cây trồng chủ yếu trong sản xuất [lâm nghiệp](https://vi.wikipedia.org/wiki/L%C3%A2m_nghi%E1%BB%87p).

Theo các bộ phận bị gây bệnh mà phân thành:

* Bệnh hại lá: Gây tổn thương đến bộ phận dinh dưỡng lá cây.
* Bệnh hại ngọn: Ảnh hưởng đến đỉnh sinh trưởng ngọn cây, ngọn cành
* Bệnh hại cành: Có thể là các chứng làm khô cành, mục cành,...
* Bệnh hại vỏ cây: Ảnh hưởng đến phần vỏ cây.
* Bệnh hại gỗ: Thường là các dạng nấm, rêu, ký sinh làm hại gỗ.
* Bệnh hại rễ: Tấn công bộ rễ của thực vật.

Theo các nguyên nhân gây bệnh và cách truyền nhiễm mà phân thành:

* Bệnh cây không truyền nhiễm: Nguyên nhân gây bệnh do các yếu tố của điều kiện ngoại cảnh (thời tiết, độ ẩm, thành phần dinh dưỡng)gây nên.
* Bệnh cây truyền nhiễm: Do sinh vật gây nên: cây ký sinh, nấm, sinh vật nhâh nguyên thủy, phytoplasmas, giun tròn,...

Kĩ thuật này ( nhận diện bệnh cây ) được nghiên cứu và ứng dụng rộng rãi vào nhiều lĩnh vực trong nông nghiệp. Trong chọn giống cây nông nghiệp, thực tiễn chứng minh, áp dụng kĩ thuật thị giác máy tính thông qua lựa chọn thông số đặc trưng ngoại hình để tiến hành phân loại và kiểm tra chất lượng hạt giống đạt kết quả cao.

Diagram

Description automatically generatedDiagram

Description automatically generatedDiagram

Description automatically generatedDiagram

Description automatically generatedDiagram

Description automatically generatedDiagram

Description automatically generated

Hệ thống nhận dạng lá cây hoặc phát hiện bệnh cây sẽ được thực hiện qua một số bước.

Bước đầu tiên: phân lớp ảnh, xác định xem ảnh có chứa một lá hợp lệ hay không, để quyết định xem nó có khả dụng trong quá trình xử lý tiếp theo hay không?

Bước thứ hai: thực hiện phân vùng ảnh để thu được ảnh nhị phân

Bước thứ ba: trích chọn những đặc trưng của lá từ ảnh chụp.

Bước cuối cùng so sánh những đặc trưng vừa được trích chọn với những ảnh từ ảnh lá trong cơ sở dữ liệu đã được gán nhãn và trả lại các loài cây (hoặc loại bệnh cây) với sự tương đồng gần nhất.

**2. MÔ HÌNH PHÂN LỚP ẢNH CHỤP LÁ CÂY**

Phần này sẽ trình bày mô hình nhận dạng để phân biệt ảnh chụp có phải là lá cây hay không. Trong các hệ thống nhận dạng loài cây hoặc phát hiện/nhận dạng bệnh cây, khâu đầu tiên là xác định xem ảnh đưa vào có chứa lá cây hay không trước khi chuyển tiếp sang giai đoạn nhận dạng lá của cây gì hoặc lá đó bị bệnh gì. Khâu này quan trọng vì nó ảnh hưởng đến độ chính xác nhận dạng của toàn hệ thống. Quá trình phân lớp được thực hiện theo hai bước:

Bước 1: Huấn luyện mô hình phân lớp dữ liệu Chuẩn bị dữ liệu cho quá trình huấn luyện (lựa chọn tập hợp các ảnh chụp lá cây thích hợp; một ảnh đạt yêu cầu là ảnh có chứa lá xác định và nền ảnh màu trắng hoặc đen hoàn toàn hoặc một màu nền đồng nhất làm nổi bật lá). Tập dữ liệu được trích chọn đặc trưng và đưa vào máy huấn luyện phân lớp SVM. Sau khi kết thúc quá trình huấn luyện, hệ thống sẽ lưu lại giá trị các tham số này (các tham số quyết định phân lớp - mô hình sau khi huấn luyện) để phục vụ cho quá trình nhận dạng sau này. Quá trình huấn luyện dữ liệu nhanh hay chậm phụ thuộc vào số lượng mẫu dữ liệu tham gia huấn luyện, thuật toán chọn để huấn luyện dữ liệu. Lựa chọn thuật toán phân lớp SVM.

Bước 2: Phân lớp dữ liệu-thử nghiệm mô hình. Sau khi đã xây dựng xong mô hình huấn luyện dữ liệu (bước 1): tiến hành phân lớp với một mẫu dữ liệu mới x. Dữ liệu x (có thể được thực hiện tiền xử lý hoặc không tùy thuộc vào người dùng ứng dụng quyết định) được trích chọn đặc trưng, sau đó sẽ được đưa vào tính toán thông qua các tham số của hàm quyết định (các tham số của mô hình huấn luyện) để xác định lớp của mẫu dữ liệu x Bước phân lớp dữ liệu này gồm nhiều giai đoạn:

**Giai đoạn tiền xử lý ảnh đầu vào** (dữ liệu nhận dạng/ phân lớp) có thể là ảnh chứa lá hợp lệ, có thể là ảnh không phải là lá cây, … nên dữ liệu này có thể được tiến hành tiền xử lý trước khi nhận dạng (quyết định tiền xử lý phụ thuộc vào người sử dụng ứng dụng).

**Giai đoạn xây dựng model :**

Tại giai đoạn này chúng ta sẽ xử dụng mô hình mạng nơ rơn tích chập (Convolution Neural Network)

Đầu tiên, ta phải hiểu convolution là gì? Là một cửa sổ trượt (Sliding Windows) trên một ma trận như mô tả hình dưới:

A picture containing shape

Description automatically generated

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.

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.



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).

### 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.

Text

Description automatically generated with medium confidence

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****Text Description automatically generated with medium confidence

#### ****Dịch filter qua bên phải một cột sẽ tạo được neuron ẩn thứ 2.****Text Description automatically generated

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.Diagram

Description automatically generated

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.

### 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.

### 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.

Chart

Description automatically generated

Thủ tục pooling phổ biến là max-pooling, thủ tục này chọn giá trị lớn nhất trong vùng đầu vào 2×2.Diagram

Description automatically generated

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.

Diagram

Description automatically generated

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.

**3,Kết quả thực nghiệp**

Dữ liệu hình ảnh gồm 7771 hình ảnh trong 4 class, ứng với 1 file data về lá cây lúc khỏe mạnh và 3 loại bênh của cây táo.

Tập dataset nhóm chia làm 3 phần gồm : Train\_ds(80%) , val\_ds(10%) , loss\_ds(10%).

Tiếp theo đền phần xử lý hình ảnh:

Mỗi bức ảnh đều có giá trị độ dài và rộng là 256. Hàm resize\_and\_rescale để resize những bức ảnh không có cùng kích thước quy định về đúng (25,256) , hàm này còn cho phép sinh mẫu ngẫu nhiên bằng cách biến dạng/gây nhiễu ngẫu nhiên ảnh gốc, tính năng này chưa được dùng đến trong thí dụ này. Tùy chỉnh rescale cho phép chuyển giá trị pixel từ 0:255 thành thang đo 0:1, để thích hợp cho mạng neuron.

Sau đó xây dựng mô hình :

Đây là công đoạn chính của quy trình, nhóm sẽ xây dựng mạng neuron với cấu trúc gồm các lớp như sau:

1. Đầu tiên là 1 lớp Convolutional tiếp nhận dữ liệu tensor với các thông số: img\_width, img\_height, channels đã được khai báo ở trên, hàm kích hoạt là RELU;
2. Lớp ẩn tiếp theo là 5 lớp conv khác
3. Ứng với các lớp comv là 1 lớp pooling để đơn giản hóa thông tin đầu ra để giảm bớt số lượng neuron.
4. Tiếp theo, 1 lớp flatten nối với 1 lớp dense (64 neuron), công dụng của bộ 2 này là chuyển giá trị đầu ra của lớp pooling thành vector feature, sau đó tinh chỉnh
5. Lớp cuối cùng: xuất kết quả , với số neuron đúng bằng số nhãn cần phân loại, với hàm kích hoạt là softmax

Bước tiếp theo huấn luyện mô hình :

Sau khi có cấu trúc mạng CNN, ta bắt đầu thực hiện quy trình huấn luyện cho nó. Quy trình này bắt đầu bằng việc đưa tensor data vào lớp tiếp nhận dữ liệu, mô hình được huấn luyện 10 lượt để tối ưu hóa accuracy.

Do tính chất đơn giản, mô hình được huấn luyện khá nhanh ngay cả khi dùng CPU. Biểu đồ cho thấy sau 50 lượt huấn luyện/kiểm định, giá trị loss giảm dần và accuracy tăng dần,tức là hiệu năng của mô hình đạt đến trạng thái tối ưu