



## Nhận diện bệnh trên lá của cây cà chua

Lê Đăng Khoa

20146497

Thông tin	Tóm tắt
<p>Trường đại học Sư phạm Kỹ thuật thành phố Hồ Chí Minh, khoa Cơ khí Chế tạo máy.</p> <p>22/05/2023</p>	<p>Lá cây cà chua không chỉ là một phần quan trọng trong cảnh quan vườn trồng hoặc vườn thực phẩm gia đình, mà còn là nguồn cung cấp quan trọng của thực phẩm và ngành công nghiệp nông nghiệp. Tuy nhiên, cây cà chua dễ bị tác động bởi một loạt các bệnh và sâu bọ, gây ảnh hưởng đáng kể đến sức khỏe và năng suất của cây trồng. Trong số những tác nhân gây bệnh, bệnh trên lá cây cà chua đứng đầu với những biểu hiện rõ rệt trên các lá và có thể lan ra trên toàn cây.</p> <p>Hiểu rõ về các bệnh trên lá cây cà chua và các chiến lược quản lý hiệu quả là một yếu tố quan trọng để đảm bảo sự thành công trong việc trồng và chăm sóc cây cà chua. Bằng cách nhìn nhận và phân biệt các triệu chứng bệnh, người trồng có thể xác định chính xác loại bệnh đang ảnh hưởng đến cây cà chua và áp dụng các biện pháp phòng ngừa và điều trị phù hợp.</p>
<p>Từ khóa</p> <p>Tomato plant diseases</p> <p>Tomato leaf diseases</p>	

## 1. Giới thiệu

Cây cà chua (*Solanum lycopersicum*) là một loại cây trồng phổ biến trên toàn thế giới, đặc biệt là trong các vùng nhiệt đới và ôn đới. Tuy nhiên, như bất kỳ loại cây trồng nào khác, cây cà chua cũng có thể mắc phải các bệnh gây hại cho lá, gây suy yếu cho sức khỏe và năng suất của cây. Bệnh trên lá cây cà chua có thể gây ra nhiều tác hại nghiêm trọng, ảnh hưởng đến sức khỏe và năng suất của cây trồng.

Theo số liệu của Tổ chức Nông nghiệp và Lương thực Liên Hiệp Quốc (FAO), mức độ thiệt hại do bệnh trên lá cây cà chua có thể lên tới 40-50% của tổng năng suất. Điều này đồng nghĩa với việc hàng triệu tấn cà chua bị mất đi hàng năm do bệnh tác động. Ngoài ra, bệnh trên lá cây cà chua cũng gây ra mất môi trường, tổn kém nguồn tài nguyên và tăng chi phí sản xuất trong nông nghiệp.

Nghiên cứu cũng chỉ ra rằng bệnh trên lá cây cà chua không chỉ gây thiệt hại về năng suất, mà còn ảnh hưởng đến chất lượng sản phẩm. Các bệnh nấm và vi khuẩn có thể làm giảm giá trị thương mại của cà chua bằng cách gây sự biến dạng, thối rữa và làm mất đi tính thẩm mỹ của quả.

Bệnh trên lá cây cà chua có thể gây tác động kinh tế đáng kể đến ngành nông nghiệp. Thiệt hại năng suất và chất lượng sản phẩm do bệnh tác động có thể làm giảm lợi nhuận của nhà nông và gây thiệt hại cho nền kinh tế địa phương. Ngoài ra, việc sử dụng thuốc bảo vệ thực vật để kiểm soát bệnh có thể gây tác động tiêu cực đến môi trường. Vì vậy, việc phát triển một phương pháp nhận dạng bệnh tự động và chính xác có thể giúp giảm thiểu tác động kinh tế và môi trường.

Việc nhận dạng bệnh trên lá cây cà chua thông qua quan sát trực tiếp và kinh nghiệm của nhà nông có thể gặp khó khăn và chưa đảm bảo độ chính xác cao. Các triệu chứng của bệnh có thể tương đối tương đồng và khó phân biệt, đặc biệt khi quá trình mắc bệnh ở giai đoạn đầu. Việc sử dụng mô hình học máy và mạng nơ-ron tích chập (CNN) có thể cung cấp một phương pháp tự động và chính xác để nhận dạng bệnh trên lá cây cà chua.

Độ chính xác cao nhất được tạo ra trong báo cáo này khi nhận dạng bệnh trên lá là 95.38%.

## 2. Phương pháp và dữ liệu

Trong nghiên cứu lần này, mong muốn của em là tạo nên và phát triển một hệ thống phát hiện bệnh trên lá của cây cà chua hiệu quả từ hình ảnh dữ liệu sử dụng thuật toán mạng nơ-ron tích chập CNN.

### 2.1. Datasets

Bộ dữ liệu hình ảnh được sử dụng trong nghiên cứu này được tìm kiếm trên các nền tảng internet và trang dữ liệu. Bộ dữ liệu bao gồm tổng cộng 1150 hình ảnh chụp lá cây cà chua. Bộ dữ liệu được chia thành 3 phân vùng. Bảng 1 cho thấy sự phân bố số lượng của các hình ảnh trong bộ dữ liệu được thu thập. Chất lượng và kích thước của mỗi hình ảnh là khác nhau do quá trình thu thập từ nhiều nguồn khác nhau và được rescale về chung 1 kích thước trước khi bắt đầu quá trình huấn luyện.

**Bảng 1.** Image Data

	Loại hình	Số lượng ảnh	Tổng
<b>train</b>	HEALTHY	65	650
	DISEASES	585	
<b>test</b>	HEALTHY	30	300
	DISEASES	270	
<b>val</b>	HEALTHY	30	300
	DISEASES	270	



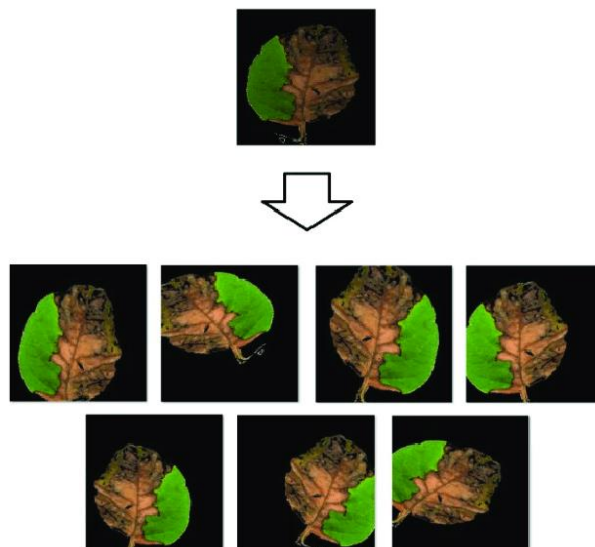
(a): Hình ảnh lá khỏe

(b): Hình ảnh lá bị bệnh

**Hình 1.** Hình ảnh lá khỏe và bệnh từ dữ liệu thu thập

## 2.2. Tăng cường dữ liệu (data augmentation)

Trong mạng neural tích chập (CNN), tăng cường dữ liệu (data augmentation) là một kỹ thuật được sử dụng để mở rộng bộ dữ liệu huấn luyện bằng cách áp dụng các biến đổi đơn giản như xoay, phóng to, thu nhỏ, lật ngang, lật dọc và lật đảo cho các mẫu dữ liệu hiện có.



**Hình 2.** Hình ảnh lá sau khi tiến hành tăng cường

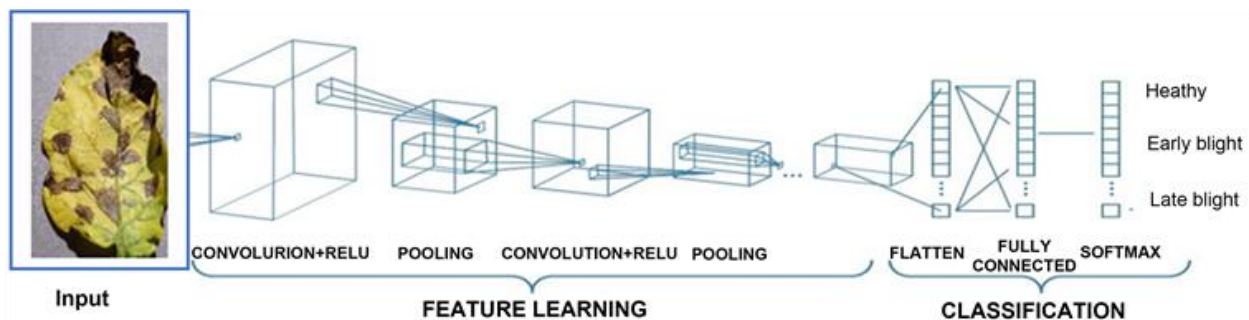
### 2.3. Thuật toán CNN - Convolutional Neural Network

CNN (Convolutional Neural Network) là một loại mạng nơ-ron nhân tạo được sử dụng rộng rãi trong lĩnh vực thị giác máy tính và xử lý ảnh. Được giới thiệu bởi Yann LeCun và đồng nghiệp vào những năm 1990, CNN đã trở thành một trong những thuật toán quan trọng nhất trong việc nhận dạng và phân loại hình ảnh. 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.

Mô hình CNN bao gồm một chuỗi các lớp, trong đó mỗi lớp có một nhiệm vụ cụ thể và thực hiện một phép biến đổi trên dữ liệu đầu vào. Dưới đây là một giới thiệu chi tiết về các lớp quan trọng trong mô hình CNN:

- Lớp đầu vào ( input layer ):
  - Lớp đầu tiên của mô hình, nhận dữ liệu đầu vào dưới dạng tensor (thường là hình ảnh).
  - Kích thước của lớp đầu vào phụ thuộc vào kích thước hình ảnh và số lượng kênh màu
- Lớp tích chập (Convolutional Layer):
  - Lớp tích chập là trung tâm của mô hình CNN và thực hiện việc trích xuất đặc trưng từ dữ liệu đầu vào.
  - Mỗi lớp tích chập sử dụng một tập hợp các bộ lọc (filters) để thực hiện phép tích chập trên đầu vào.
  - Kết quả của phép tích chập là các bản đồ đặc trưng biểu diễn thông tin cấu trúc trong dữ liệu.
- Lớp kích hoạt (Activation Layer):
  - Sau mỗi lớp tích chập, một lớp kích hoạt thường được áp dụng để đưa ra một phản hồi phi tuyến tính cho mạng.
  - Hàm kích hoạt phổ biến nhất là ReLU (Rectified Linear Unit), nhưng cũng có thể sử dụng các hàm khác như Sigmoid hoặc Tanh.

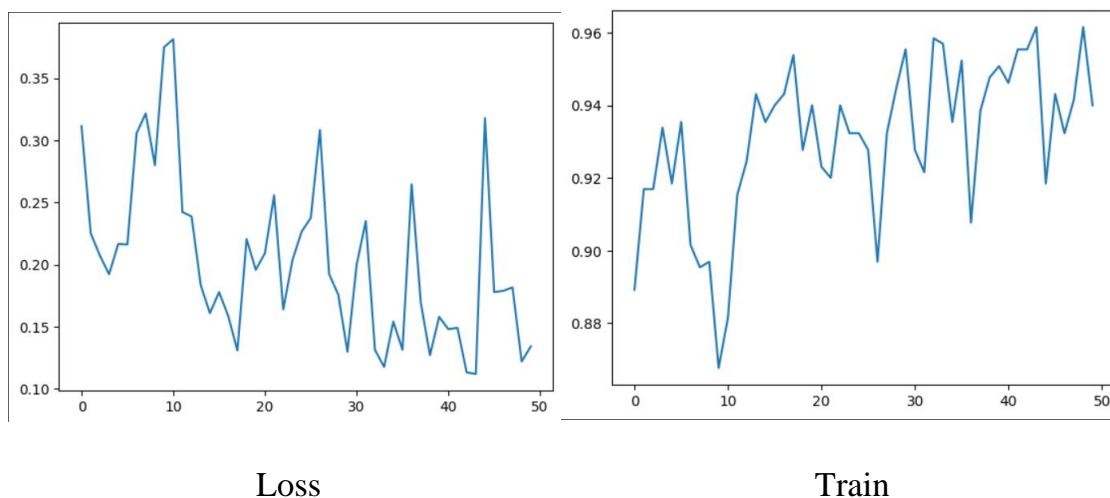
- Lớp gộp (Pooling Layer):
  - Lớp gộp được sử dụng để giảm kích thước không gian của các bản đồ đặc trưng và giảm độ phức tạp tính toán.
  - Một kỹ thuật phổ biến trong lớp gộp là lớp gộp tối đa (max pooling), nơi giá trị lớn nhất trong một cửa sổ được chọn để trích xuất và làm đại diện cho khu vực tương ứng.
- Lớp đầy đủ kết nối (Fully Connected Layer):
  - Lớp đầy đủ kết nối sử dụng các nơ-ron để kết nối các bản đồ đặc trưng đã được trích xuất với lớp đầu ra.
  - Mỗi nơ-ron trong lớp đầy đủ kết nối kết nối với tất cả các nơ-ron trong lớp trước đó và truyền tải thông tin qua mạng.
- Lớp đầu ra (Output Layer):
  - Lớp đầu ra cuối cùng của mô hình chứa các nơ-ron đưa ra dự đoán cho các lớp đầu vào
  - Số lượng nơ-ron trong lớp đầu ra phụ thuộc vào số lượng lớp phân loại hoặc số lượng giá trị đầu ra mong muốn.



**Hình 2.** Cấu trúc CNN

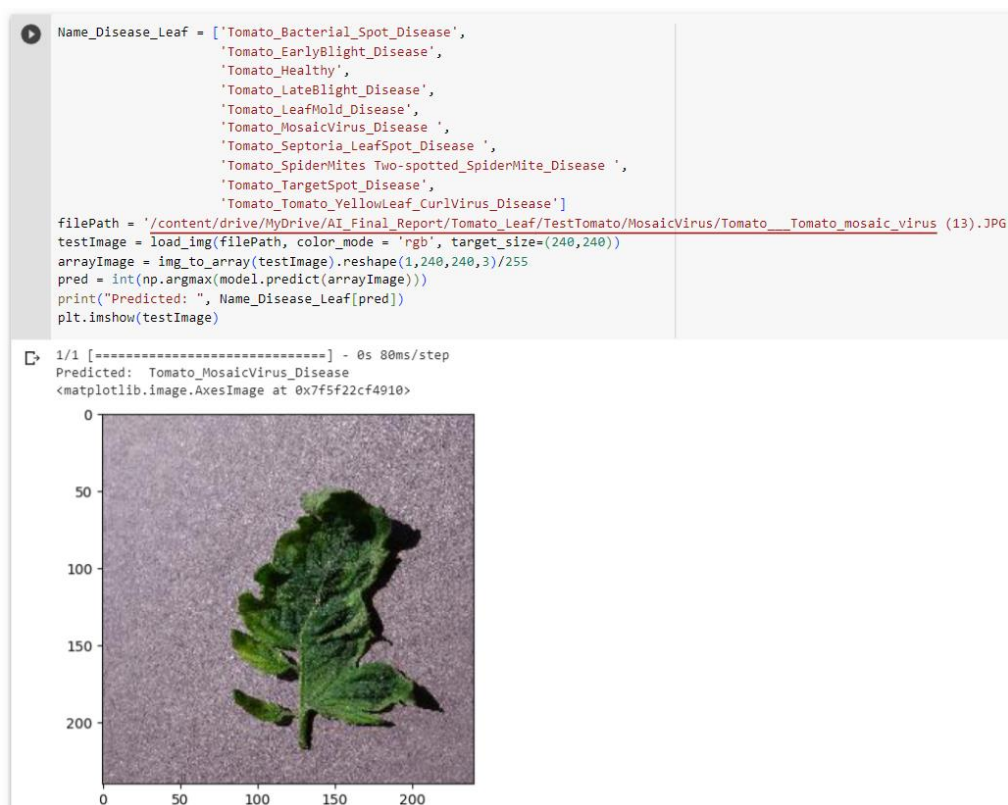
### 3. Kết quả

Mô hình CNN chuẩn đoán bệnh trên lá cây cà chua với độ chính xác là 95%, đưa ra kết quả nhận dạng đúng khá cao, phân biệt được giữa lá khỏe mạnh và lá mắc các loại bệnh. Hình 4 biểu thị độ chính xác của mô hình với số lần học là 50.



**Hình 3.** Đồ thị độ chính xác của mô hình và giá trị mất mát với 50 lần học

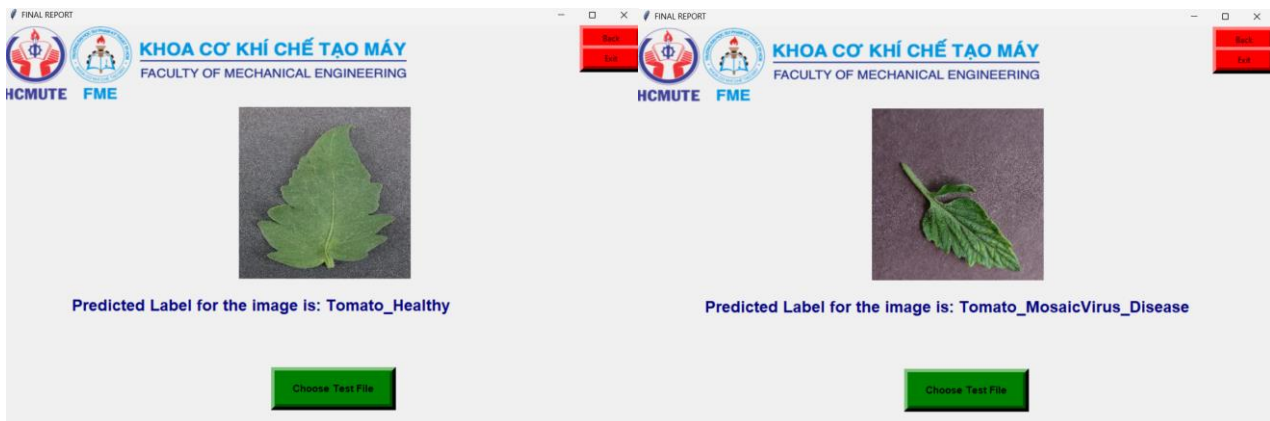
Kết quả dự đoán bằng hình ảnh đưa qua từ Google Drive:



**Hình 4.** Kết quả dự đoán trên Google Colab



Để mô hình được ứng dụng rộng rãi, bất cứ ai cũng có thể sử dụng một giao diện được tạo ra với yêu cầu đơn giản và dễ sử dụng ở hình 5.



**Hình 5.** Kết quả dự đoán giao diện

#### 4. Kết luận

Bệnh trên lá cây cà chua có thể gây tác động kinh tế đáng kể đến ngành nông nghiệp. Thiệt hại năng suất và chất lượng sản phẩm do bệnh tác động có thể làm giảm lợi nhuận của nhà nông và gây thiệt hại cho nền kinh tế địa phương. Ở bài báo này em tiến hành xây dựng 1 mô hình bằng mô hình CNN với độ chính xác là 95.38% để nhận dạng được các loại bệnh. Cùng với đó xây dựng 1 giao diện người dùng đơn giản, thuận tiện khi sử dụng.



## Tài liệu tham khảo

[1]. Tomato Detection using Python

Link truy cập: [\(141\) Plant Leaf Disease Prediction using Deep learning - YouTube](#)

[2]. 16 loại bệnh thường gặp trên cây cà chua

Link truy cập: [16 loại bệnh thường gặp trên cây cà chua và cách trị tận gốc \(lisado.vn\)](#)

[3]. TOPDev, Thuật toán CNN – Convolutional Neural Network

Link truy cập: <https://topdev.vn/blog/thuat-toan-cnn-convolutional-neural-network/>