

NGHIÊN CỨU CÁC KỸ THUẬT HỌC MÁY VÀ ỨNG DỤNG XÂY DỰNG DÂY CHUYỀN PHÂN LOẠI HOA QUẢ TỰ ĐỘNG

RESEARCHING MACHINE LEARNING ALGORITHMS AND APPLICATIONS TO CONSTRUCTION
OF AUTOMATIC FRUIT CLASSIFICATION LINE

Nguyễn Thành Dương¹, Vũ Đức Cường¹,
Lê Quảng Nam², Mai Thanh Hồng^{3,*}

TÓM TẮT

Thuật toán học máy được phân thành 4 loại: học có giám sát, học không giám sát, học bán giám sát, học củng cố. Có nhiều thuật toán khác nhau được sử dụng bởi học máy: hồi quy tuyến tính, hồi quy logistic, cây quyết định, mạng nơ-ron, thuật toán SVM (Support Vector Machine), k-láng giềng gần nhất, ... Phân loại trái cây có vai trò quan trọng trong nhiều ứng dụng công nghiệp bao gồm các nhà máy, siêu thị và các lĩnh vực khác. Tầm quan trọng của việc phân loại trái cây cũng có thể được nhìn thấy ở những người có nhu cầu ăn kiêng để giúp họ chọn đúng loại trái cây. Phương pháp phân loại thủ công để phân loại trái cây đã được sử dụng trước đó, tốn nhiều thời gian và luôn cần đến nhân công. Trước đây, nhiều phương pháp học máy để phân loại trái cây đã được đề xuất. Do đó, nhóm tác giả nghiên cứu mô hình hiệu quả để phân loại trái cây bằng cách sử dụng mạng nơ-ron.

Từ khóa: Học máy, trái cây, phân loại, mạng nơ-ron.

ABSTRACT

Machine learning algorithms are classified into 4 types: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. There are many different algorithms used by machine learning: linear regression, logistic regression, decision tree, neural network, SVM (Support Vector Machine) algorithm, k-nearest neighbor, ... Classifier Fruit has an important role in many industrial applications including factories, supermarkets and other fields. The importance of fruit sorting can also be seen in people with dietary needs to help them choose the right fruit. The manual method of sorting fruit has been used before, which is time consuming and always requires labor. In the past, many machine learning methods for fruit classification have been proposed. Therefore, we study an efficient model for fruit classification using a neural network.

Keyword: Machine learning, fruits, classification, neural network.

¹Lớp Kỹ thuật phần mềm 03 - K14, Khoa CNTT, Trường Đại học Công nghiệp Hà Nội

²Lớp kỹ thuật phần mềm 04 - K14, Khoa CNTT, Trường Đại học Công nghiệp Hà Nội

³Khoa CNTT, Trường Đại học Công nghiệp Hà Nội

*Email: maihong1979@gmail.com

1. GIỚI THIỆU

Phân loại quả là một quá trình quan trọng đối với người sản xuất, ảnh hưởng đến việc đánh giá chất lượng quả và thị trường xuất khẩu. Mặc dù con người có thể phân loại theo cách thủ công, nhưng sẽ rất chậm, tốn nhiều công sức

và dễ nhạ. Do đó, cần có một hệ thống phân loại trái cây thông minh. Những năm gần đây, các nhà nghiên cứu đã phát triển nhiều thuật toán để phân loại trái cây bằng nhiều phương pháp khác nhau. Nhóm tác giả đã nghiên cứu một phương pháp sử dụng học máy để phân loại hoa quả. Các đặc điểm về màu sắc, cấu trúc và hình thái là những đặc điểm phổ biến nhất được sử dụng để xác định bệnh tật, độ chín và loại của quả.

Nhóm tác giả đã tập trung nghiên cứu thuật toán CNN (perceptron) để thực hiện việc phân loại. Thuật toán ban đầu được Frank Rosenblatt đề xuất dựa trên ý tưởng của Neural thần kinh, nhanh chóng tạo nên tiếng vang lớn trong lĩnh vực AI. "Sai đầu sửa đấy, cuối cùng sẽ thành công" chính là ý tưởng của thuật toán này. Về cơ bản ta sẽ cho tìm thử 1 nghiệm, nếu như không thỏa mãn, ta tiếp tục vòng lặp để cập nhật nghiệm đến 1 vị trí tốt hơn, cứ như thế cho đến khi tìm được nghiệm thật sự (cách làm này hoàn toàn có thể bởi bài toán này được chứng minh là có hữu hạn số lần thử, do đó luôn luôn tìm được nghiệm thật sự sau 1 khoảng thời gian lặp).

Để có thể lấy được thông số như độ lớn, độ rộng, độ chín của trái cây, ta cần một công cụ khác đó là OpenCV. OpenCV (OpenSource Computer Vision) là một thư viện mã nguồn mở. OpenCV được thiết kế để tính toán hiệu quả và với sự tập trung nhiều vào các ứng dụng thời gian thực. OpenCV được sử dụng cho đa dạng nhiều mục đích và ứng dụng khác nhau bao gồm: Hình ảnh street view, kiểm tra và giám sát tự động, ... Đặc biệt, nó giúp người lập trình xử lý và hiển thị hình ảnh/ Video/ I/O (core, imgproc, highgui) từ tệp dữ liệu có sẵn.

2. PHƯƠNG PHÁP VÀ TƯ LIỆU

2.1. Thuật toán perceptron

Một mạng lưới thần kinh xoắn (CNN) là một loại hình cụ thể của mạng lưới thần kinh nhân tạo sử dụng perceptron, một đơn vị thuật toán máy học, cho học có giám sát, phân tích dữ liệu. CNNs áp dụng cho xử lý hình ảnh, xử lý ngôn ngữ tự nhiên và các loại nhiệm vụ nhận thức.

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.

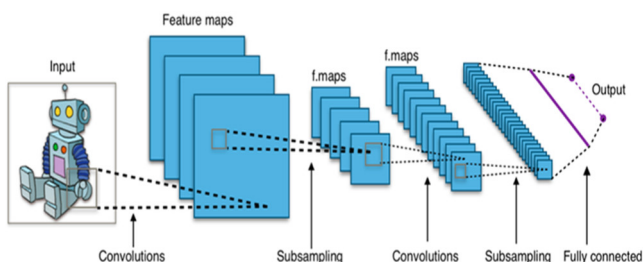
Mỗi một lớp sau khi thông qua các hàm kích hoạt sẽ tạo ra các thông tin trừu tượng hơn cho các lớp tiếp theo. Trong mô hình mạng truyền ngược (feedforward neural network) thì mỗi neural đầu vào (input node) cho mỗi neural đầu ra trong các lớp tiếp theo.

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 (training) 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 1. Mạng CNN

Trong mô hình CNN có hai 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 góc 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.

2.2. OpenCV

OpenCV (OpenSource Computer Vision) là một thư viện mã nguồn mở. OpenCV được phát hành theo giấy phép BSD, do đó nó hoàn toàn miễn phí cho cả học thuật và

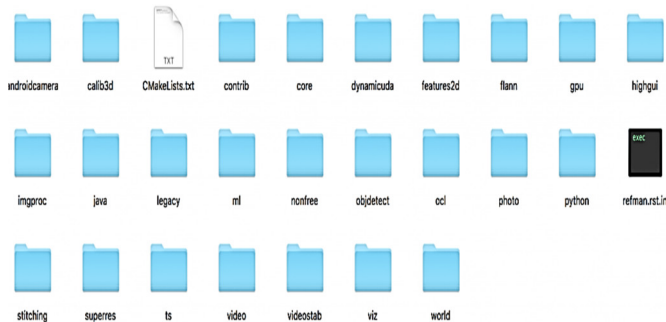
thương mại. Nó có các interface C++, C, Python, Java và hỗ trợ Windows, Linux, Mac OS, iOS và Android. OpenCV được thiết kế để tính toán hiệu quả và với sự tập trung nhiều vào các ứng dụng thời gian thực. Được viết bằng tối ưu hóa C/C++, thư viện có thể tận dụng lợi thế của xử lý đa lõi. Được sử dụng trên khắp thế giới, OpenCV có cộng đồng hơn 47 nghìn người dùng và số lượng download vượt quá 6 triệu lần. Phạm vi sử dụng từ nghệ thuật tương tác, cho đến lĩnh vực khai thác mỏ, bản đồ trên web hoặc công nghệ robot.

Trang chủ chính thức của OpenCV: <http://opencv.org/>

Trang download các phiên bản của OpenCV: <http://opencv.org/downloads>

Các module của OpenCV:

Folder modules chứa tất các modules của thư viện OpenCV và mỗi module này sẽ được build dưới dạng thư viện liên kết tĩnh và thư viện liên kết động mà chúng ta sẽ sử dụng trong project. Mỗi module sẽ có một chức năng riêng biệt.



Hình 2. Các module của OpenCV

Core

Đây là module chứa các cấu trúc, class cơ bản mà OpenCV sẽ sử dụng trong việc lưu trữ và xử lý hình ảnh như Mat, Scale, Point, Vec... và các phương thức cơ bản sử dụng cho các module khác

Imgproc

Đây là một module xử lý hình ảnh của OpenCV bao gồm các bộ lọc (filter) linear và non-linear và các phép biến đổi hình học (tranformation) như resize hoặc các phép biến đổi affine, chuyển đổi hệ màu và các thuật toán liên quan đến histogram (biểu đồ) của hình ảnh.

Highgui

Đây là một module cho phép tương tác với người dùng trên UI (User Interface) như hiển thị hình ảnh, video capturing.

Features2d

Module tìm các đặc trưng (feature) của hình ảnh. Trong module có implement các thuật toán rút trích đặc trưng như PCA...

Calib3d

Hiệu chuẩn máy ảnh và xây dựng lại 3D.

Objdetect

Module cho việc phát hiện các đối tượng như khuôn mặt, đôi mắt, cốc, người, xe hơi, vv trong hình ảnh. Các thuật toán được sử dụng trong module này là Haar-like Features.

ML

Module này chứa các thuật toán về Machine Learning phục vụ cho các bài toán phân lớp (Classification) và bài toán gom cụm (Clustering). Ví dụ như thuật toán SVM (Support Vector Machine), ANN...

Video

Module phân tích video gồm ước lượng chuyển động, trừ nền, và các thuật toán theo dõi đối tượng (object tracking).

3. KẾT QUẢ THẢO LUẬN

3.1. Thu thập tập dữ liệu

Các loại mẫu hình ảnh dữ liệu: quả nhỏ xanh, quả nhỏ chín, quả to xanh, quả to chín.



Hình 3. Hình ảnh minh họa về dữ liệu

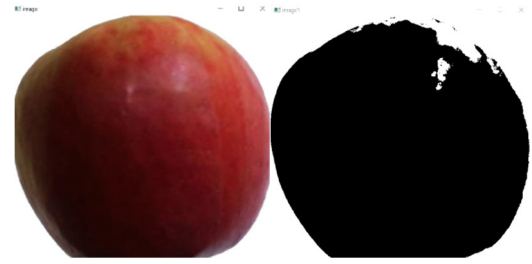
Sau khi đã thu thập được bộ dữ liệu, ta sử dụng opencv để xử lý và trích xuất các trường thông tin mà ta cần lấy của từng hình ảnh (độ lớn, độ rộng, độ chín) để tiến hành training vào trong file .csv.

Kết quả sau khi training tập dataset:

	A	B	C	D	E
1	dai	rong	chín	loại	loại1
2	285	147	4.6	1	1
3	291	159	4.7	1	1
4	281	125	3.4	1	1
5	295	166	4.9	1	1
6	298	166	5	1	1
7	304	166	4.9	1	1
8	305	162	4.9	1	1
9	311	160	4.9	1	1
10	307	158	4.9	1	1
11	305	157	4.8	1	1
12	301	158	4.8	1	1
13	295	157	4.8	1	1
14	289	152	4.8	1	1
15	283	126	3.4	1	1
16	284	154	4.7	1	1
17	277	154	4.6	1	1
18	274	153	4.5	1	1
19	269	152	4.4	1	1
20	266	153	4.3	1	1
21	263	156	4.3	1	1
22	264	157	4.2	1	1
23	268	158	4.3	1	1
24	271	153	4.4	1	1
25	271	152	4.4	1	1

Hình 4. Dữ liệu các thông số sau khi đọc toàn bộ tập data training

3.2. Kết quả sau khi cho input



Hình 5. Dữ liệu sau khi tiến xử lý

Kết quả phân loại theo ảnh đầu vào:

Gia trị dau vào: `[[233. 232. 9.6]]`

Ket qua: `[[1 2]]`

Qua tao thuoc loai nho chin

Hình 6. Hình ảnh kết quả sau khi đưa input vào

4. KẾT LUẬN

Bài báo đã trình bày nghiên cứu các kỹ thuật học máy và ứng dụng xây dựng dây chuyền phân loại hoa quả tự động để có thể tự động hóa trong chế biến thực phẩm phát huy tác dụng nâng cao năng suất. Không còn phải phụ thuộc vào nhân công nhưng vẫn có thể thực hiện năng suất và đạt kết quả mong muốn và nhanh chóng. Đây sẽ là công nghệ của tương lai, khi mà nó có thể giúp máy tính thực thi những việc tưởng chừng như không thể vào 10 năm trước: phân loại cả ngàn vật thể khác nhau trong các bức ảnh, tự tạo chú thích cho ảnh, bắt chước giọng nói và chữ viết của con người, giao tiếp với con người, hay thậm chí cả sáng tác văn hay âm nhạc.

TÀI LIỆU THAM KHẢO

- [1]. OpenCV Python Tutorial - GeeksforGeeks
- [2]. Detect Ripe Fruit in 5 Minutes with OpenCV | by James Thesken | Medium