

Trí tuệ

tính toán và khoa học thần kinh của Hindawi  
 Tập 2017, ID bài viết 7361042, 6 trang  
<https://doi.org/10.1155/2017/7361042>



## Bài viết nghiên cứu

### Học sâu để nhận dạng thực vật trong môi trường tự nhiên

Yu Sun, Yuan Liu, Guan Wang và Haiyan Zhang

Trường Khoa học và Công nghệ Thông tin, Đại học Lâm nghiệp Bắc Kinh, Bắc Kinh 100083, Trung Quốc

Thư từ nên được gửi tới Haiyan Zhang; zhyzml@bjfu.edu.cn

Nhận ngày 2 tháng 3 năm 2017; Chấp nhận ngày 18 tháng 4 năm 2017; Xuất bản ngày 22 tháng 5 năm 2017

Biên tập viên học thuật: Sergio Solinas

Bản quyền © 2017 Yu Sun và cộng sự. Đây là bài viết truy cập mở được phân phối theo Giấy phép Ghi nhận tác giả Creative Commons, cho phép sử dụng, phân phối và sao chép không hạn chế trong bất kỳ phương tiện nào, miễn là tác phẩm gốc được trích dẫn hợp lý.

Nhận dạng hình ảnh thực vật đã trở thành một trọng tâm liên ngành trong cả phân loại thực vật và thị giác máy tính. Bộ dữ liệu hình ảnh thực vật đầu tiên được thu thập bằng điện thoại di động trong cảnh tự nhiên được trình bày, bao gồm 10.000 hình ảnh của 100 loài cây cảnh trong khuôn viên Đại học Lâm nghiệp Bắc Kinh. Mô hình học sâu 26 lớp bao gồm 8 khối xây dựng còn lại được thiết kế để phân loại thực vật quy mô lớn trong môi trường tự nhiên. Mô hình đề xuất đạt tỷ lệ nhận dạng 91,78% trên bộ dữ liệu BJFU100, chứng tỏ deep learning là một công nghệ đầy hứa hẹn cho lâm nghiệp thông minh.

## 1. Giới thiệu

Nhận dạng hình ảnh thực vật tự động là giải pháp hứa hẹn nhất hướng tới thu hẹp khoảng cách phân loại thực vật, nhận được sự quan tâm đáng kể trong cả cộng đồng thực vật học và máy tính. Khi công nghệ học máy tiến bộ, các mô hình phức tạp đã được đề xuất để nhận dạng nhà máy tự động. Với sự phổ biến của điện thoại thông minh và sự xuất hiện của ứng dụng di động PlantNet [1], hàng triệu bức ảnh thực vật đã được thu thập. Nhận dạng thực vật tự động dựa trên thiết bị di động là điều cần thiết để giám sát sinh thái dựa trên xã hội trong thế giới thực [2], giám sát thực vật ngoại lai xâm lấn [3], phổ biến khoa học sinh thái, v.v.

Việc cải thiện hiệu suất của các mô hình nhận dạng thực vật dựa trên thiết bị di động ngày càng thu hút sự chú ý của các học giả và kỹ sư.

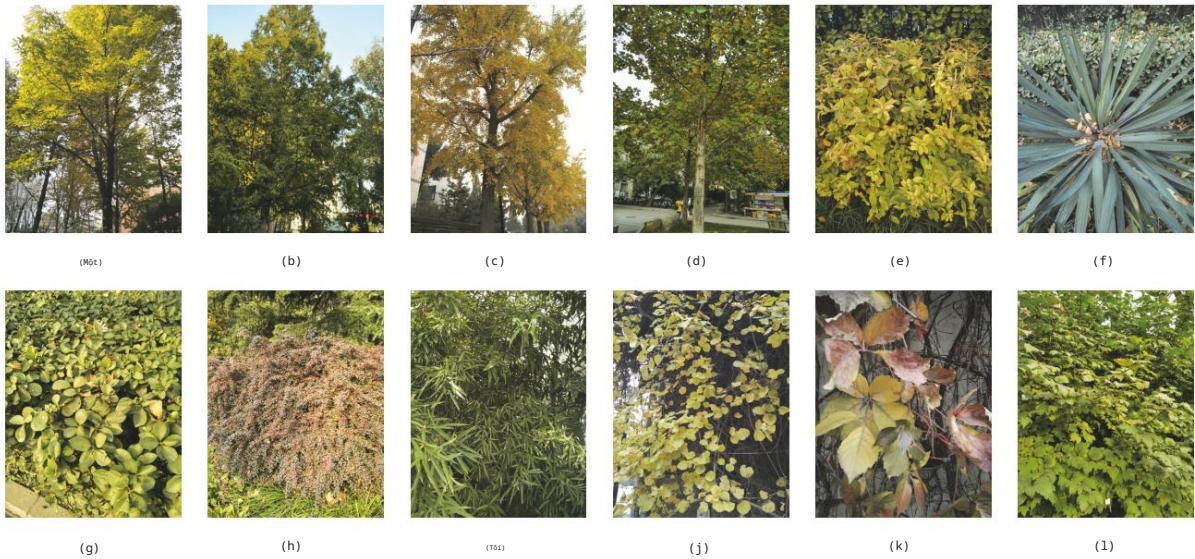
Ngày nay, nhiều nỗ lực đã được tiến hành trong việc trích xuất các đặc tính địa phương của lá, hoa hoặc quả. Hầu hết các nhà nghiên cứu sử dụng các biến thể về đặc tính của lá làm công cụ so sánh để nghiên cứu thực vật và một số bộ dữ liệu về lá bao gồm bộ dữ liệu lá Thụy Điển, bộ dữ liệu Flavia và bộ dữ liệu ICL là điểm chuẩn tiêu chuẩn. Trong [4], Soderkvist đã trích xuất các đặc điểm hình dạng và đặc điểm mômen của lá và phân tích 15 lớp cây Thụy Điển khác nhau bằng cách sử dụng phương pháp lan truyền ngược cho mạng nơ-ron truyền thẳng. Trong [5], Fu et al. đã chọn độ tương phản cục bộ và các tham số khác để mô tả đặc điểm của các pixel xung quanh của tinh mạch. Mạng lưới thần kinh nhân tạo được sử dụng để phân đoạn các tinh mạch và

những chiếc lá khác. Thí nghiệm cho thấy mạng lưới thần kinh hiệu quả hơn trong việc xác định hình ảnh tinh mạch. Li và cộng sự. [6] đã đề xuất một phương pháp trích xuất gân lá hiệu quả bằng cách kết hợp kỹ thuật rần với mạng lưới thần kinh tế bào, thu được kết quả khả quan về phân đoạn lá.

Ông và Huang đã sử dụng mạng lưới thần kinh xác suất làm công cụ phân loại để xác định hình ảnh lá cây, có độ chính xác nhận dạng tốt hơn so với mạng lưới thần kinh BP [7].

Vào năm 2013, ý tưởng nhận dạng lá dựa trên tự nhiên đã được đề xuất và phương pháp thuật toán phân đoạn đường viền dựa trên mô hình lá đa giác đã được sử dụng để thu được hình ảnh đường viền [8]. Với việc học sâu đang trở thành điểm nóng trong lĩnh vực nhận dạng hình ảnh, Liu và Kan đã đề xuất các đặc điểm kết cấu kết hợp với đặc điểm hình dạng, sử dụng kiến trúc mạng niềm tin sâu làm bộ phân loại [9]. Zhang và cộng sự. đã thiết kế một hệ thống deep learning bao gồm tám lớp Mạng thần kinh Convolution để xác định hình ảnh chiếc lá và đạt được tỷ lệ nhận dạng cao hơn. Một số nhà nghiên cứu tập trung vào những bông hoa.

Nilsback và Zisserman đã đề xuất một phương pháp túi từ ngữ trực quan để mô tả màu sắc, hình dạng, đặc điểm kết cấu và các đặc điểm khác [10]. Trong [11], Zhang và cộng sự. kết hợp các đặc trưng Harr với các đặc trưng SIFT của ảnh bông hoa, mã hóa chúng bằng phương pháp mã hóa thừa thớt không âm và phân loại chúng theo phương pháp k-lân cận gần nhất. Trong [12], họ đã đưa ra phương pháp nhận dạng hoa hồng đang hái bằng cách tích hợp mạng nơ-ron BP. Các nghiên cứu nhận dạng thực vật bằng quả còn tương đối hiếm. Li và cộng sự. đề xuất phương pháp tích hợp đa tính năng sử dụng ưu tiên Ainet làm



Hình 1: Hình ảnh ví dụ về bộ dữ liệu BJFU100. (a) Buckeye Trung Quốc, (b) metasequoia, (c) Ginkgo biloba, (d) cây hoa tulip lai, (e) Weigela cv florida hoàng tử đỏ, (f) Yucca gloriosa, (g) Euonymus kiautschovicus Loes, (h) Berberis thunbergii var. atropurpurea, (i) tre đỏ, (j) Celastrus orbiculatus, (k) Parthenocissus quinquefolia, và (l) Viburnum opulus.

thuật toán nhận dạng [13]. Sau bao nhiêu năm vẫn tiếp tục thăm dò công nghệ nhận dạng thực vật, chuyên dụng các ứng dụng di động như LeafSnap [14], Pl@ntNet [1], hoặc ứng dụng Nhận dạng hoa của Microsoft Garage [15] có thể thuận tiện sử dụng để xác định thực vật.

Mặc dù nghiên cứu về phân loại thực vật tự động đã mang lại kết quả tốt đẹp, người ta phải lưu ý rằng những mô hình đó vẫn còn xa so với yêu cầu của một hệ sinh thái hoàn toàn tự động kịch bản giám sát [3]. Các bộ dữ liệu nói trên thiếu hình ảnh thực vật dựa trên thiết bị di động thu được trong cảnh tự nhiên khác nhau rất nhiều về người đóng góp, máy ảnh, khu vực, thời kỳ của năm, từng cây riêng lẻ, v.v. Các mô hình phân loại truyền thống phụ thuộc nhiều vào tiền xử lý để loại bỏ nền phức tạp và tăng cường các tính năng mong muốn. Cái gì hơn nữa, kỹ thuật tính năng thủ công không có khả năng xử lý các tập dữ liệu quy mô lớn bao gồm không bị ràng buộc hình ảnh.

Để vượt qua những thách thức nói trên và lấy cảm hứng từ bước đột phá của deep learning trong nhận dạng hình ảnh, chúng tôi có được bộ dữ liệu BJFU100 bằng điện thoại di động một cách tự nhiên môi trường. Bộ dữ liệu được đề xuất chứa 10.000 hình ảnh của 100 loài cây cảnh tại Đại học Lâm nghiệp Bắc Kinh khuôn viên đại học. Mô hình học sâu 26 lớp bao gồm 8 khối xây dựng còn lại được thiết kế cho nhà máy không được kiểm soát nhận biết. Mô hình đề xuất đạt được sự công nhận tỷ lệ 91,78% trên tập dữ liệu BJFU100.

2. Bộ dữ liệu BJFU100 được đề xuất và dữ liệu sâu  
Mô hình học tập

Kiến trúc học sâu được hình thành bởi nhiều tuyến tính và các phép biến đổi phi tuyến của dữ liệu đầu vào, với mục tiêu mang lại những biểu diễn trừu tượng và có tính phân biệt hơn [16]. Những phương pháp này đã cải thiện đáng kể công nghệ nhận dạng giọng nói, nhận dạng đối tượng trực quan,

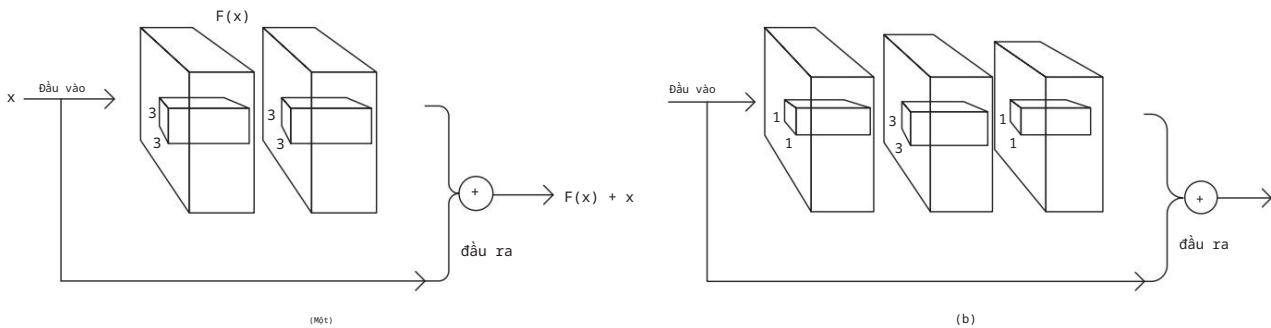
phát hiện đối tượng và nhiều lĩnh vực khác như thuốc khám phá và bộ gen [17]. Dây thần kinh tích chập sâu các mạng được đề xuất trong [18] đã thể hiện hiệu suất vượt trội trong nhiệm vụ phân loại hình ảnh quy mô lớn của ILSVRC 2012 [19]. Mô hình đã được đào tạo trên hơn một triệu

hình ảnh và đã đạt được tỷ lệ lỗi kiểm tra top 5 là 15,3% trên 1.000 lớp. Nó gần như giảm một nửa tỷ lệ lỗi của phương pháp cạnh tranh tốt nhất. Sự thành công này đã mang lại một cuộc cách mạng về thị giác máy tính [17]. Những tiến bộ gần đây trong lĩnh vực đã nâng cao tính khả thi của các ứng dụng học sâu để giải quyết các vấn đề phức tạp trong thế giới thực [20].

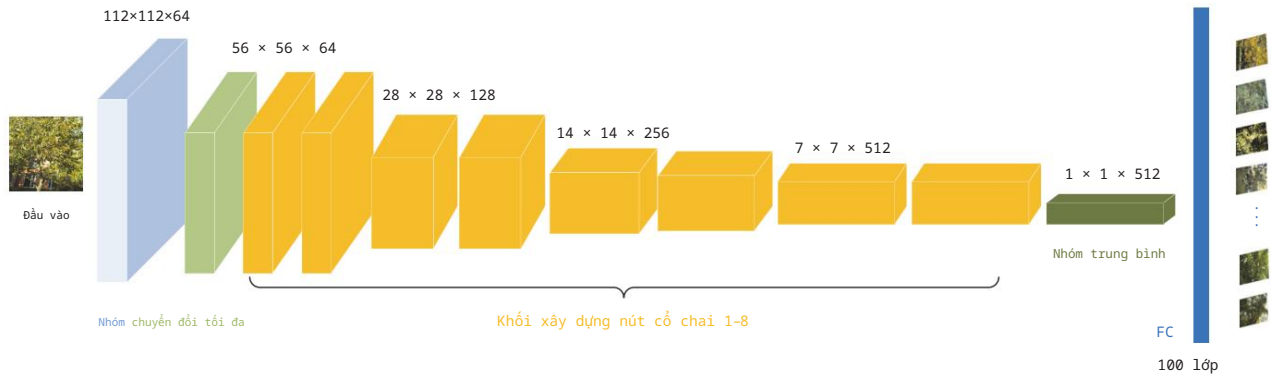
2.1. Bộ dữ liệu BJFU100. Bộ dữ liệu BJFU100 được thu thập từ cảnh thiên nhiên bằng thiết bị di động. Nó bao gồm 100 loài trồng cây cảnh tại khuôn viên trường Đại học Lâm nghiệp Bắc Kinh. Mỗi danh mục chứa một trăm bức ảnh khác nhau được điện thoại thông minh mua lại trong môi trường tự nhiên. Điện thoại thông minh được trang bị ống kính một tiêu cự tương đương 28 mm tiêu cự và cảm biến RGB có độ phân giải 3120 × 4208.

Đối với các trụ cao, hình ảnh được chụp từ góc thấp ở mặt đất như trong Hình 1(a)–1(d). Cây bụi thấp bị bắn từ một góc cao, như trong Hình 1(e)–1(h). Khác cây cảnh được chụp từ một góc ngang. Đối tượng có thể có kích thước khác nhau theo thứ tự độ lớn (nghĩa là một số hình ảnh cho thấy chỉ chiếc lá, những cái khác là toàn bộ cây từ xa), như được hiển thị trong Hình 1(i)–1(l).

2.2. Mạng lưới dư sâu. Với độ sâu mạng ngày càng tăng, các phương pháp truyền thống không được như mong đợi cải thiện độ chính xác nhưng gây ra các vấn đề như biến mất độ dốc và suy thoái. Mạng dư thừa, nghĩa là, ResNet, giới thiệu các kết nối bỏ qua cho phép thông tin (từ đầu vào hoặc thông tin đã học ở các lớp trước) được truyền đi nhiều hơn vào các lớp sâu hơn [23, 24]. Với độ sâu ngày càng tăng, ResNets cung cấp khả năng xấp xỉ hàm tốt hơn như



Hình 2: (a) Một khối xây dựng cơ bản. (b) Một khối xây dựng “nút thắt cổ chai” của các mạng dư sâu.



Hình 3: Kiến trúc của mô hình ResNet 26 lớp để nhận dạng thực vật.

họ đạt được nhiều thông số hơn và đóng góp thành công vào giải quyết các vấn đề độ dốc biến mất và suy thoái. Sâu mạng dư với các đơn vị dư đã cho thấy sự hấp dẫn độ chính xác và hành vi hội tụ tốt trên một số tác vụ nhận dạng hình ảnh quy mô lớn, chẳng hạn như ImageNet [23] và MS COCO [25] cuộc thi.

2.2.1. Khối xây dựng còn lại. Đơn vị cấu trúc còn lại sử dụng các kết nối phím tắt với sự trợ giúp của ánh xạ nhận dạng. Kết nối lối tắt là những kết nối bỏ qua một hoặc nhiều lớp. Ánh xạ cơ bản ban đầu có thể được thực hiện bằng cách chuyển tiếp mạng lưới thần kinh bằng các kết nối phím tắt. Các khối xây dựng được minh họa trong Hình 2 được định nghĩa là

$$\begin{aligned} &= ( \cdot , \{ \cdot \} ) + \cdot , \\ &= 2 ( 1 ) , \end{aligned} \tag{1}$$
$$( \cdot ) = \text{tối đa } ( 0 , ) ,$$

ở đâu và là các vectơ đầu vào và đầu ra của xếp chồng lên nhau các lớp tương ứng. Hàm  $( \cdot , \{ \cdot \} )$  đại diện cho ánh xạ dư cần phải học. Chức năng  $( \cdot )$  biểu thị ReLU [26] và các độ lệch được bỏ qua để đơn giản hóa ký hiệu. Kích thước của và phải bằng thực hiện phép cộng theo từng phần tử. Nếu đây không phải là trường hợp, một Phép chiếu tuyến tính được áp dụng để khớp với các kích thước của và :

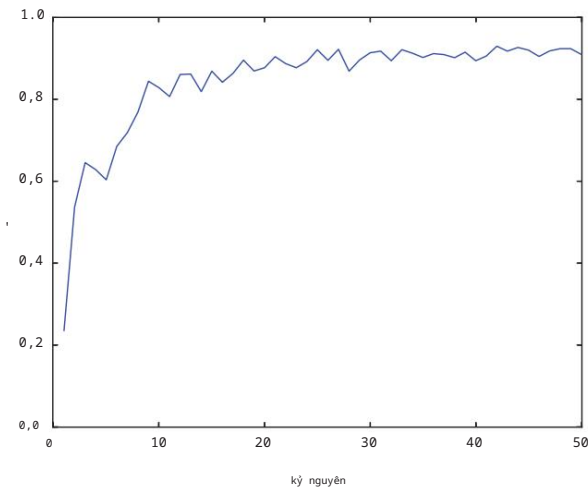
$$= ( \cdot , \{ \cdot \} ) + \cdot . \tag{2}$$

Khối xây dựng đường cơ sở được hiển thị trong Hình 2 (a). Một kết nối phím tắt được thêm vào mỗi cặp bộ lọc  $3 \times 3$ . Liên quan đến thời gian huấn luyện trên lưới sâu hơn, một điểm nghẽn khối xây dựng được thiết kế như trong Hình 2(b). Ba lớp là các phép chập  $1 \times 1$ ,  $3 \times 3$  và  $1 \times 1$ , trong đó  $1 \times 1$  lớp có nhiệm vụ giảm bớt và khôi phục lại kích thước, để lại lớp  $3 \times 3$  một nút cổ chai với kích thước nhỏ hơn kích thước đầu vào/đầu ra [23]. Sử dụng khối xây dựng cổ chai ít tham số hơn để thu được nhiều lớp trừu tượng hơn.

Kiến trúc mạng tổng thể của ResNet 26 lớp của chúng tôi, tức là ResNet26, mô hình được mô tả trong Hình 3. Như Hình 3 cho thấy, mô hình chủ yếu được thiết kế bằng cách sử dụng nút cổ chai khu nhà. Hình ảnh đầu vào được đưa vào lớp tích chập  $7 \times 7$  và lớp tổng hợp tối đa  $3 \times 3$ , theo sau là 8 khối xây dựng cổ chai. Khi kích thước tăng lên,  $1 \times 1$  tích chập được sử dụng trong nút cổ chai để khớp với kích thước. Tích chập  $1 \times 1$  làm phong phú thêm mức độ trừu tượng và làm giảm độ phức tạp về thời gian. Mạng kết thúc với một mạng toàn cầu tổng hợp trung bình, một lớp được kết nối đầy đủ và một lớp softmax. Chúng tôi áp dụng chuẩn hóa hàng loạt (BN) [27] ngay sau mỗi lớp tích chập và trước lớp kích hoạt ReLU [26]. Lấy mẫu xuống được thực hiện bởi lớp chập đầu tiên, tối đa lớp tổng hợp và các khối xây dựng nút cổ chai 3, 5 và 7.

3. Thí nghiệm và kết quả

3.1. Thực hiện và tiền xử lý. Việc triển khai mô hình dựa trên khung học sâu nguồn mở



Hình 4: Sự phát triển của độ chính xác phân loại trong bộ kiểm tra.

máy ảnh [28]. Tất cả các thử nghiệm được thực hiện trên máy chủ Ubuntu 16.04 Linux với CPU i7-3770 3,40 GHz (bộ nhớ 16 GB) và GPU GTX 1070 (bộ nhớ 8 GB). 100 mẫu của mỗi lớp được chia thành 80 mẫu huấn luyện và 20 mẫu thử nghiệm. So với các phương pháp phân loại thông thường, quá trình tiền xử lý dữ liệu trên các phương pháp học sâu đơn giản hơn nhiều. Trong bài báo này, đầu vào của mạng là hình ảnh màu RGB. Tất cả hình ảnh chỉ cần được thay đổi kích thước thành  $224 \times 224$  pixel và sau đó giá trị mỗi pixel được chia cho 255 để chuẩn hóa.

3.2. Thuật toán đào tạo. Trong giai đoạn lan truyền ngược, tham số mô hình được huấn luyện bằng thuật toán giảm dần độ dốc ngẫu nhiên (SGD), với hàm mất entropy chéo phân loại làm đối tượng tối ưu hóa. SGD có thể được biểu thị như sau:

$$\begin{aligned} \theta &= \theta + \eta \cdot \nabla_{\theta} L(\theta) \\ \Delta &= \eta \cdot \nabla_{\theta} L(\theta) \end{aligned} \tag{3}$$

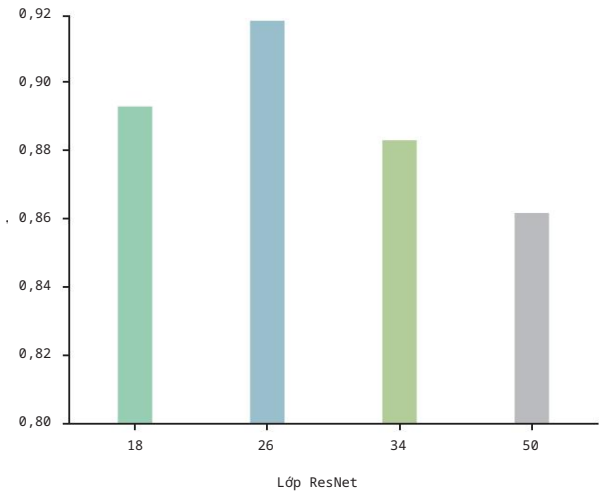
trong đó  $\eta$  là độ nhảy,  $\nabla_{\theta} L(\theta)$  là độ lệch nhân,  $\theta$  biểu thị rằng mỗi phần tử được nhân lên,  $\Delta$  là lấy mẫu lên, xuống là lấy mẫu xuống,  $\Delta$  biểu thị cập nhật trọng số của lớp và là tốc độ học tập. Hàm mất entropy chéo được định nghĩa là

$$H = - \sum_{i=1}^n p_i \log(p_i) \tag{4}$$

phần tử thứ  $i$  ở đầu trong vectơ điểm phân loại

Sau một số thử nghiệm huấn luyện sơ bộ, tốc độ học cơ bản được đặt thành 0,001, tốc độ này giảm dần ở mỗi kỷ nguyên. Tốc độ phân rã là  $10^{-6}$  và động lượng là 0,9. Hình 4 thể hiện quá trình huấn luyện của mô hình ResNet26. Độ chính xác của thử nghiệm được cải thiện nhanh chóng kể từ kỷ nguyên đầu tiên và ổn định sau 40 kỷ nguyên.

Trí tuệ tính toán và khoa học thần kinh



Hình 5: Kiểm tra độ chính xác của mô hình ResNet18, ResNet34, ResNet50 [23] và ResNet26. ResNet26 được đề xuất vượt trội hơn ResNet tham chiếu tốt nhất 2,51%.

3.3. Phân tích kết quả. Để tìm ra mạng dư sâu tốt nhất, một loạt thử nghiệm đã được tiến hành trên bộ dữ liệu BJJU100. Hình 5 cho thấy sự so sánh độ chính xác của thử nghiệm giữa mô hình ResNet26 được đề xuất và mô hình ResNet ban đầu gồm 18, 34 và 50 lớp [23] được thiết kế cho ImageNet. ResNet18, ResNet34 và ResNet50 mang lại độ chính xác thử nghiệm lần lượt là 89,27%, 88,28% và 86,15%. ResNet26 được đề xuất mang lại độ chính xác 91,78%, giúp tăng hiệu suất tổng thể lên tới 2,51%.

ResNet26 là sự cân bằng tốt nhất giữa dung lượng mô hình và độ khó tối ưu hóa. Đối với kích thước của BJJU100, ResNet26 chứa đủ tham số có thể huấn luyện để tìm hiểu các lông nguyên bản phân biệt, giúp ngăn ngừa việc trang bị thiếu. So với mô hình lớn hơn, ResNet26 mang lại sự hội tụ nhanh và mạnh mẽ trong quá trình tối ưu hóa SGD, giúp ngăn ngừa việc trang bị quá mức hoặc rơi vào mức tối ưu cục bộ.

4. ResNet26 trên Bộ dữ liệu Flavia

Để cho thấy tính hiệu quả của mô hình ResNet26 được đề xuất, một loạt thử nghiệm đã được thực hiện trên bộ dữ liệu lá Flavia [29] có sẵn công khai. Nó bao gồm 1907 hình ảnh có độ phân giải  $1600 \times 1200$  pixel, với 32 danh mục. Một số mẫu được hiển thị trong Hình 6. Chúng tôi chọn ngẫu nhiên 80% tập dữ liệu để huấn luyện và 20% để thử nghiệm.

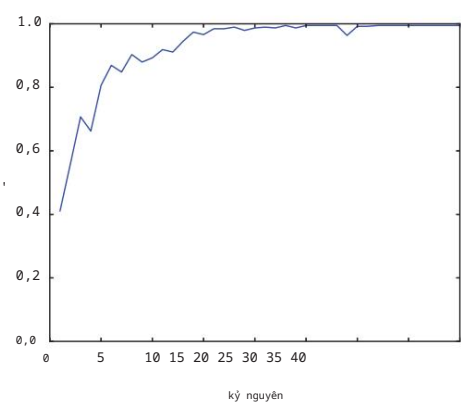
Tất cả các hình ảnh được nhân đôi và thay đổi kích thước thành  $224 \times 224$  pixel. Giá trị mỗi pixel được chia cho giá trị tối đa và trừ đi giá trị trung bình của dữ liệu.

Thuật toán huấn luyện hoàn toàn giống với thuật toán áp dụng cho tập dữ liệu BJJU100. Hình 7 thể hiện quá trình huấn luyện của mô hình ResNet26. Độ chính xác của thử nghiệm được cải thiện nhanh chóng kể từ kỷ nguyên đầu tiên và ổn định sau 30 kỷ nguyên.

Độ chính xác kiểm tra của từng mô hình được ước tính bằng xác thực chéo 10 lần, như được hiển thị trong Hình 8. ResNet18, ResNet34 và ResNet50 đạt được độ chính xác kiểm tra lần lượt là 99,44%, 98,95% và 98,60%. ResNet26 được đề xuất đạt độ chính xác 99,65% giúp tăng hiệu quả tổng thể



Hình 6: Hình ảnh ví dụ về tập dữ liệu Flavia.



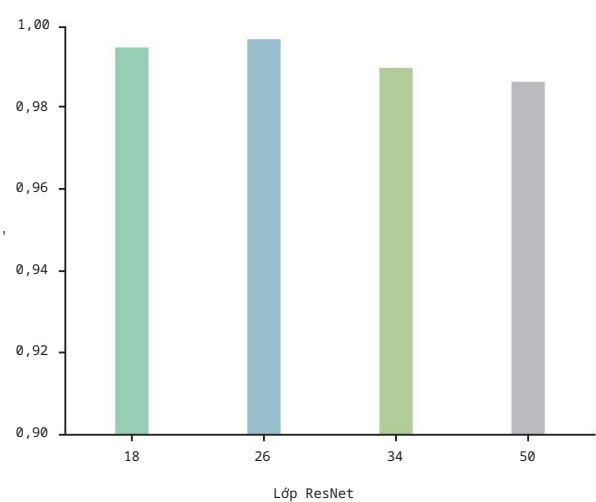
Hình 7: Sự phát triển của độ chính xác phân loại trong bộ kiểm tra.

lên tới 0,21%. Bảng 1 tóm tắt kết quả của chúng tôi và các kết quả được công bố trước đây trên tập dữ liệu lá Flavia [29]. Mô hình ResNet26 đạt được mức cải thiện 0,28% so với phương pháp hoạt động tốt nhất.

5. Kết luận

Thiết bị di động đầu tiên thu được bộ dữ liệu BJFU100 chứa 10.000 hình ảnh của 100 loài thực vật, cung cấp nền tảng dữ liệu cho nghiên cứu nhận dạng thực vật tiếp theo. Chúng tôi tiếp tục mở rộng bộ dữ liệu BJFU100 bằng cách bao phủ rộng hơn về loài và mùa. Tập dữ liệu này mở cho cộng đồng học thuật, có sẵn tại <http://pan.baidu.com/s/1jILsypS>. Công việc này cũng nghiên cứu cách tiếp cận học sâu để tự động khám phá các biểu diễn cần thiết cho việc phân loại, cho phép sử dụng một đầu cuối thống nhất. Đường ống đầu cuối để nhận biết thực vật trong môi trường tự nhiên. Mô hình ResNet26 đề xuất cho kết quả chính xác 91,78% trong bộ thử nghiệm, chứng tỏ rằng học sâu là công nghệ đầy hứa hẹn để phân loại thực vật quy mô lớn trong môi trường tự nhiên.

Trong tương lai, cơ sở dữ liệu BJFU100 sẽ được mở rộng cho nhiều loài thực vật hơn ở các giai đoạn khác nhau của vòng đời và chú thích chi tiết hơn. Mô hình học sâu sẽ



Hình 8: Kiểm tra độ chính xác của mô hình ResNet18, ResNet34, ResNet50 [23] và ResNet26 trên tập dữ liệu Flavia. ResNet26 được đề xuất tốt hơn ResNet tham chiếu tốt nhất 0,21%.

Bảng 1: So sánh tỷ lệ nhận dạng trên tập dữ liệu Flavia.

Phương pháp	Tỷ lệ nhận dạng
PBPNN [21]	93,82%
SVM [22]	96,00%
DBN (có "bỏ học") [9]	99,37%
Công việc của chúng ta	99,65%

mở rộng từ nhiệm vụ phân loại đến dự đoán năng suất, phát hiện côn trùng, phân chia bệnh, v.v.

Xung đột lợi ích

Các tác giả tuyên bố rằng họ không có xung đột lợi ích.

Tác giả đóng góp

Yu Sun và Yuan Liu đã đóng góp như nhau cho công việc này.

Sự nhìn nhận

Công trình này được hỗ trợ bởi Quỹ nghiên cứu cơ bản cho các trường đại học miền Trung: YX2014-17 và TD2014-01.

Người giới thiệu

[1] A. Joly, H. Goeau, P. Bonnet và cộng sự, "Nhận dạng thực vật tương tác dựa trên dữ liệu hình ảnh xã hội," Tin học sinh thái, tập. 23, trang 22-34, 2014.

[2] H. Goeau, P. Bonnet và A. Joly, "Nhiệm vụ xác định thực vật LifeCLEF 2015," trong Kỷ yếu của Hội nghị và Phòng thí nghiệm của Diễn đàn Đánh giá (CLEF '15), 2015.

[3] H. Goeau, P. Bonnet và A. Joly, "Nhận dạng thực vật trong thế giới mở (lifecyclef 2016)," trong Kỷ yếu của ghi chú làm việc của CLEF, tập. 2016, 2016.



6

Trí tuệ tính toán và khoa học thần kinh

[4] O. Soderkvist, " Phân loại lá cây Thụy Điển bằng thị giác máy tính, 2001.

[5] H. Fu, Z. Chi, J. Chang và C. Fu, "Khái thác các đặc tính gân lá dựa trên mạng lưới thần kinh nhân tạo-Nghiên cứu về nhận dạng thực vật sống I," *Bản tin Thực vật học Trung Quốc*, tập. 21, trang 429-436, 2003.

[6] Y. Li, Q. Zhu, Y. Cao và C. Wang, "Phương pháp trích xuất gân lá dựa trên kỹ thuật rần," trong *Kỷ yếu của Hội nghị quốc tế về Mạng lưới thần kinh và não bộ (ICNN&B '05)*, trang 885-888, 2005.

[7] P. He và L. Huang, "Trích xuất đặc điểm và nhận dạng lá cây," *Tạp chí Nghiên cứu Cơ giới hóa Nông nghiệp*, tập. 6, tr. 52, 2008.

[8] G. Cerutti, L. Tougne, J. Mille, A. Vacavant và D. Coquin, "Hiểu về lá trong hình ảnh tự nhiên - một cách tiếp cận dựa trên mô hình để xác định loài cây," *Thị giác máy tính và Hiểu hình ảnh*, tập. 117, không. 10, trang 1482-1501, 2013.

[9] N. Liu và J.-M. Kan, "Nhận dạng lá cây dựa trên phương pháp kết hợp đa tính năng và mạng lưới niềm tin sâu sắc," *Tạp chí của Đại học Lâm nghiệp Bắc Kinh*, tập. 38, không. 3, trang 110-119, 2016.

[10] M.-E. Nilsback và A. Zisserman, "Tìm hiểu sâu hơn về sự phân chia hoa," *Máy tính Hình ảnh và Tầm nhìn*, tập. 28, không. 6, trang 1049-1062, 2010.

[11] C. Zhang, J. Liu, C. Liang, Q. Huang và Q. Tian, "Phân loại hình ảnh bằng cách sử dụng phép biến đổi giống Harr của các đặc điểm cục bộ với phần dư mã hóa," *Xử lý tín hiệu*, tập. 93, không. 8, trang 2111-2118, 2013.

[12] YJ Wang, YW Zhang, DL Wang, X. Yin và WJ Zeng, "Thuật toán nhận dạng hình ảnh hoa hồng ăn được dựa trên mạng lưới thần kinh," *Tạp chí của Đại học Nông nghiệp Trung Quốc*, tập. 19, không. 4, trang 180-186, 2014.

[13] X. Li, L. Li, Z. Gao, J. Zhou và S. Min, "Nhận dạng hình ảnh của quả hoa trà dựa trên sở thích tích hợp đa tính năng aiNET," *Giao dịch của Hiệp hội Kỹ thuật Nông nghiệp Trung Quốc*, tập. 28, không. 14, trang 133-137, 2012.

[14] N. Kumar, PN Belhumeur, A. Biswas và cộng sự, "Leafsnap: một hệ thống thị giác máy tính để nhận dạng các loài thực vật tự động," trong *Kỷ yếu của Tầm nhìn Máy tính-ECCV 2012*, trang 502-516, 2012.

[15] <https://www.microsoft.com/en-us/research/project/flowerre> đồng/.

[16] Y. Bengio, A. Courville và P. Vincent, "Học đại diện: đánh giá và quan điểm mới," *Giao dịch của IEEE về Phân tích Mẫu và Trí thông minh Máy*, tập. 35, không. 8, trang 1798-1828, 2013.

[17] Y. LeCun, Y. Bengio và G. Hinton, "Học sâu," *Tự nhiên*, tập. 521, không. 7553, trang 436-444, 2015.

[18] A. Krizhevsky, I. Sutskever và GE Hinton, "Phân loại Imagenet với mạng lưới thần kinh tích chập sâu," trong *Những tiến bộ trong hệ thống xử lý thông tin thần kinh*, trang 1097-1105, 2012.

[19] <http://www.image-net.org/challenges/LSVRC/2012/>.

[20] B. Huval, T. Wang, S. Tandon và cộng sự, "Đánh giá thực nghiệm về học sâu khi lái xe trên đường cao tốc," <https://arxiv.org/abs/1504.01716>.

[21] A. Kulkarni, H. Rai, K. Jahagirdar và P. Upparamani, "Một kỹ thuật nhận dạng lá để phân loại thực vật bằng cách sử dụng các khoảng khắc RBPNN và Zernike," *Tạp chí Quốc tế về Nghiên cứu Nâng cao về Kỹ thuật Máy tính và Truyền thông*, tập. 2, trang 984- 988, 2013.

..

[22] C. Sari, CB Akgul và B. Sankur, "Sự kết hợp của các đặc điểm hình dạng thô, bộ mô tả phạm vi và ma trận khoảng cách đa tỷ lệ để nhận dạng lá," trong *Kỷ yếu của Hội nghị chuyên đề quốc tế lần thứ 55 (ELMAR '13)*, trang 23-26, tháng 9 năm 2013.

[23] K. He, X. Zhang, S. Ren và J. Sun, "Học tập dư thừa sâu để nhận dạng hình ảnh," trong *Kỷ yếu của Hội nghị IEEE về Thị giác máy tính và Nhận dạng mẫu (CVPR '16)*, trang 770 - 778, Las Vegas, Nev, Mỹ, tháng 6 năm 2016.

[24] K. He, X. Zhang, S. Ren và J. Sun, "Ảnh xạ nhận dạng trong các mạng dư sâu," trong *Kỷ yếu của Hội nghị Châu Âu về Thị giác Máy tính*, trang 630-645, 2016.

[25] J. Dai, K. He và J. Sun, "Phân đoạn ngữ nghĩa nhận biết cá thể thông qua các tầng mạng đa tác vụ," trong *Kỷ yếu của Hội nghị IEEE về Tầm nhìn Máy tính và Nhận dạng Mẫu (CVPR '16)*, trang. 3150-3158, Las Vegas, Nev, Hoa Kỳ, tháng 6 năm 2016.

[26] V. Nair và GE Hinton, "Các đơn vị tuyến tính được chỉnh lưu cải thiện các máy Boltzmann bị hạn chế," trong *Kỷ yếu của Hội nghị quốc tế lần thứ 27 về Học máy (ICML '10)*, trang. 807-814, tháng 6 năm 2010.

[27] S. Ioffe và C. Szegedy, "Chuẩn hóa hàng loạt: Tăng tốc đào tạo mạng sâu bằng cách giảm sự thay đổi hiệp phương sai nội bộ," [https:// arxiv.org/abs/1502.03167](https://arxiv.org/abs/1502.03167). [28] [https:// keras.io/](https://keras.io/).

[29] SG Wu, FS Bảo, EY Xu, Y.-X. Wang, Y.-F. Chang và Q.- L. Xiang, "Thuật toán nhận dạng lá để phân loại thực vật bằng cách sử dụng mạng thần kinh xác suất," trong *Hội nghị chuyên đề quốc tế của IEEE về xử lý tín hiệu và công nghệ thông tin*, trang 2007. 11-16, Giza, Ai Cập, tháng 12 năm 2007.

