

Xem các cuộc thảo luận, số liệu thống kê và hồ sơ tác giả của ấn phẩm này tại: <https://www.researchgate.net/publication/356151677>

# Nhận dạng thực vật bằng mạng lưới thần kinh tích chập và học chuyển giao

**Bài báo** *TRON* Tập chí Khoa học Kỹ thuật Đại học Pamukkale · Tháng 1 năm 2021

DOI: 10.5505/pajes.2020.84042

TRÍCH DẪN

4

ĐỌC

63

## 2 tác giả:



**Tolga Karahan**  
Hypatos GmbH

1 SỰ XUẤT BẢN 4 TRÍCH DẪN

XEM HỒ SƠ



**Vasif Nabiye**  
Đại học Kỹ thuật Karadeniz

159 CÔNG BỐ 1.214 TRÍCH DẪN

XEM HỒ SƠ

Một số tác giả của ấn phẩm này cũng đang thực hiện các dự án liên quan sau:



Thiết kế, triển khai và đánh giá Hệ thống dạy kèm thông minh thích ứng dựa trên trò chơi nhằm cải thiện kỹ năng giải quyết vấn đề của học sinh [Xem Kế hoạch](#)



Ortaokul Öğrencileri trong chương trình Öğretimine Yönelik Oyun Tabanlı Öğretim Uyarlanabilir Bir Ortamın Geliştirilmesi ve Öğrencilerin Bilgi Lâemsel Düşünme Becerisine Etkisinin İncelenmesi (Tübitak 1001 - 118R034 dự án nolu) [Xem Kế hoạch](#)

## Nhận dạng thực vật bằng mạng lưới thần kinh tích chập và chuyển giao học hỏi

## Bạn có thể trả tiền cho việc chuyển nhượng và chuyển nhượng của mình với bạn

*Tolgahan KARAHAN*<sup>1</sup> , *Vasif NABIYEV*<sup>2\*</sup> 

1,2Khoa Kỹ thuật Máy tính, Khoa Kỹ thuật, Đại học Kỹ thuật Karadeniz, Trabzon, Thổ Nhĩ Kỳ.  
[karahantlg@gmail.com](mailto:karahantlg@gmail.com) , [vasif@ktu.edu.tr](mailto:vasif@ktu.edu.tr)

Đã nhận/Geliş Tarihi:30.04.2020  
Chấp nhận/Kabul Tarihi:20.11.2020

Ön tập/Düzelme Tarihi:14.11.2020

doi: 10.5505/pajes.2020.84042 Bài  
báo nghiên cứu/Araştırma Makalesi

trừ tượng

Thiên nhiên phong phú với số lượng lớn các loài thực vật và hoa và vì sự đa dạng to lớn của chúng; việc xác định các loài này đòi hỏi phải có chuyên môn trong lĩnh vực này. Việc phát triển hệ thống nhận dạng thực vật tự động có thể giúp quá trình này dễ dàng hơn. Trong công việc này, Mạng lưới thần kinh chuyển đổi sâu và Học chuyển giao đã được sử dụng để phát triển một hệ thống nhận dạng như vậy. Hình ảnh trong cơ sở dữ liệu đã được thu thập từ các cơ sở dữ liệu khác và web và tổng cộng nó bao gồm 5.345 hình ảnh hoa và thực vật thuộc 76 loài. 65 loài trong số đó là các loài hoa khác nhau và 11 loài trong số đó là các loài thực vật khác. Kỹ thuật tăng dữ liệu được áp dụng nhằm tăng số lượng ảnh trong cơ sở dữ liệu và nâng cao khả năng khái quát hóa của mô hình. Để tăng cường dữ liệu, xoay ngẫu nhiên ở bốn góc, thay đổi độ sáng ngẫu nhiên trong phạm vi [-0,2, 0,2] và lật ngang đã được áp dụng. Ngoài ra, các kỹ thuật tiền xử lý như cắt xén trung tâm và chuẩn hóa cũng đã được áp dụng cho hình ảnh trước khi đưa chúng vào mô hình. Trong nhận dạng thực vật tự động, độ chính xác đạt được là 0,9971 trên tập huấn luyện và độ chính xác 0,9897 đạt được trên tập kiểm tra.

**Từ khóa:** Trí tuệ nhân tạo, Học sâu, Mạng lưới thần kinh tích chập, Nhận dạng, Phân loại.

## 1. Giới thiệu

Việc phân loại các loài thực vật và hoa đã được nghiên cứu khác nhau. Nhiều phương pháp và mô hình khác nhau đã được phát triển để phân loại các loài thực vật hoặc hoa. Trong những năm gần đây, Mạng thần kinh chuyển đổi đã trở nên nổi bật trong số các phương pháp tiếp cận khác và chúng có một số lợi thế quan trọng so với các phương pháp truyền thống. Chúng có ít tham số hơn nhiều so với mạng nơ-ron nhân tạo truyền thống, vì vậy chúng hiệu quả hơn về mặt tính toán. Chúng ít gặp phải vấn đề trang bị quá mức do việc chia sẻ tham số giữa các vùng khác nhau của hình ảnh [1]. Ngoài ra, họ có khả năng tìm hiểu thêm các tính năng trừu tượng. Vì những lý do này, các mô hình phân loại hình ảnh hiện đại dựa trên Mạng thần kinh chuyển đổi.

CNN có ba loại lớp chính [1]. Các lớp này là các lớp tích chập, các lớp gộp và các lớp được kết nối đầy đủ. Các lớp tích chập áp dụng thao tác tích chập cho đầu vào của chúng thông qua các hạt nhân và giá trị của các hạt nhân này được học trong quá trình đào tạo. Các lớp gộp áp dụng phương pháp lấy mẫu xuống cho đầu vào của chúng để giảm kích thước không gian và các lớp được kết nối đầy đủ sử dụng các tính năng đã học được qua các lớp tích chập trước đó và thực hiện phân loại.

## Öz

Doğa büyük miktarda bitki ve çiçek türü zenginliğine sahiptir ve bu çeşitlilik nedeniyle bu türlerin tanımlanması botanik alanında uzmanlık gerektirmektedir. *Bạn có thể làm điều đó bằng cách sử dụng hệ thống hỗ trợ của mình và có thể đảm bảo rằng bạn có đủ khả năng. Bạn có thể làm điều đó bằng cách sử dụng hệ thống của bạn để giúp bạn trả tiền cho việc chuyển nhượng và chuyển giao công việc của mình.* Veritabanındaki đã xác định được kế hoạch xác thực và kế hoạch webden và toplamda 76 tûre ait 5.345 và bitkiden oluşmaktadır. *Türlerin 65 tanesi çeşitli çiçek türleridir ve 11 tanesi ise diğer bitki çeşitlerindendir. Veritabanındaki görüntü sayısını ve modelin genelleme kapasitesini arttırmak için çeşitli veri çoğaltma yöntemleri uygulanmıştır. Veri çoğaltmak için, 4 açıdan rastgele döndürme, [-0.2, 0.2] aralığında rastgele parlaklık değişimi ve yatay yansıtma işlemleri uygulanmıştır. Eğer bạn muốn, bạn có thể làm theo mô hình của mình, bạn có thể làm điều đó và bình thường hóa nó. Geliştirilen mô hình eğitim verileri için 0.9971, test verileri için ise 0.9897 isabet oranı elde etmiştir.*

**Anahtar kelimeler:**Yapay zekâ, Derin öğrenme, Konvolüsyonel yapay sinir ağıları, Tanımlama, Sınıflandırma.

Sự phát triển về phần cứng và việc sử dụng GPU để triển khai song song giúp có thể đào tạo các mạng sâu trên cơ sở dữ liệu quy mô lớn. ImageNet là một cơ sở dữ liệu lớn và nó thường được sử dụng để đánh giá hiệu suất của các thuật toán nhận dạng hình ảnh [2]. Các mô hình Mạng thần kinh chuyển đổi sâu đã thể hiện hiệu suất ấn tượng trên điểm chuẩn ImageNet [3]. Kết quả là chúng được sử dụng trong rất nhiều nhiệm vụ Thị giác Máy tính.

Sử dụng Mạng thần kinh chuyển đổi với kỹ thuật có tên là Transfer Learning làm tăng đáng kể kết quả phân loại. Vì vậy, kỹ thuật Transfer Learning cũng được sử dụng trong nghiên cứu này. Học chuyển giao có thể được thực hiện bằng cách sử dụng các tham số đã học từ một miền khác nhưng có liên quan. Bằng cách này, chúng tôi có thể điều chỉnh các tham số đã học cho phù hợp với nhiệm vụ của mình và làm cho mô hình của chúng tôi đạt được khả năng khái quát hóa tốt hơn [4]. Về vấn đề này, học chuyển giao có thể được sử dụng bằng cách điều chỉnh các tham số quy mô lớn được đào tạo trên ImageNet cho phù hợp với nhiệm vụ hiện tại.

Một kỹ thuật hữu ích khác để cải thiện khả năng khái quát hóa của các mô hình Thị giác Máy tính là tăng cường dữ liệu. Trong tăng cường dữ liệu, các mẫu có thể được tăng lên bằng cách tạo dữ liệu tổng hợp. Các mẫu mới có thể được tạo bằng cách xoay, lật, cắt hoặc làm biến dạng các mẫu ban đầu. Nó có thể cải thiện khả năng khái quát hóa của mô hình, bởi vì các mẫu

\*Đồng tác giả/Yazışılan Yazar

được giới thiệu vào mô hình trong các tình huống khác nhau [5]. Vì vậy, một mô hình mạnh mẽ hơn đã đạt được. Các kỹ thuật tăng cường dữ liệu như xoay ngẫu nhiên theo bốn góc, thay đổi độ sáng ngẫu nhiên trong phạm vi  $[-0,2, 0,2]$  và lật ngang đã được sử dụng trong nghiên cứu này.

Để phát triển hệ thống nhận dạng cây và hoa, một bộ dữ liệu đã được tạo. Cơ sở dữ liệu hình ảnh đã được chia thành các tập huấn luyện và kiểm tra bằng cách lấy mẫu phân tầng [6]. Tập huấn luyện cuối cùng bao gồm 29932 hình ảnh và tập kiểm tra cuối cùng bao gồm 7483 hình ảnh. Nhờ lấy mẫu phân tầng, mỗi lớp đã được đưa vào cả tập huấn luyện và tập kiểm tra.

Việc cắt xén và chuẩn hóa trung tâm đã được áp dụng cho hình ảnh trước khi chúng được cung cấp làm đầu vào cho mô hình. Sau 15 kỳ nguyên, mô hình đề xuất đã đạt độ chính xác 0,9971 trên tập huấn luyện và 0,9897 độ chính xác trên tập kiểm tra.

Mô hình được đào tạo cũng đã được sử dụng để phát triển ứng dụng nhận dạng cây/hoa di động cho nền tảng Android. Bằng cách sử dụng ứng dụng di động phát triển, mô hình đề xuất đã được thử nghiệm trên thực vật hoang dã.

## **2 Công việc liên quan**

O'Shea và cộng sự. đã mô tả Mạng thần kinh chuyển đổi và những ưu điểm của chúng so với mạng thần kinh truyền thống. Các tác giả đã giải thích về các khối xây dựng của CNN, kiến trúc tổng thể và hoạt động của từng lớp [1].

Alex Krizhevsky và cộng sự. đã đề xuất mô hình CNN của họ để phân loại ImageNet. Họ đã trình bày sự thành công của các mô hình CNN và việc sử dụng GPU để tăng tốc Deep Learning [7].

Asifullah Khan và cộng sự. đã trình bày một cuộc khảo sát về kiến trúc CNN gần đây. Họ đã mô tả lịch sử của CNN, các thành phần của chúng và sự phát triển kiến trúc của các mô hình CNN. Họ cũng đề cập đến các ứng dụng của CNN và các kỹ thuật gần đây được sử dụng cùng với chúng [8].

C Tân và cộng sự. đã trình bày một cuộc khảo sát về Deep Transfer Learning. Họ đã mô tả những vấn đề có thể khắc phục được bằng cách học chuyển tiếp và những lợi ích mà nó mang lại. Họ đã phân loại các phương pháp tiếp cận chuyển giao học tập và trình bày loại này [9].

L.Perez và J.Wang đã so sánh các kỹ thuật tăng cường dữ liệu khác nhau và ảnh hưởng của chúng đến việc phân loại hình ảnh. Họ đã áp dụng các phương pháp tăng cường dữ liệu truyền thống và gần đây trên một tập hợp con của bộ dữ liệu ImageNet và trình bày kết quả của mình. Họ cũng đề xuất phương pháp riêng của mình gọi là tăng cường thần kinh. Trong phương pháp này, mạng lưới thần kinh đã được sử dụng để học các phép tăng cường nhằm cải thiện tốt nhất các bộ phân loại [10].

Dyrmann và cộng sự. đã phát triển mô hình CNN để phân loại 22 loài cỏ dại và cây trồng khác nhau ở giai đoạn phát triển ban đầu nhằm giúp quản lý cỏ dại cụ thể tại địa điểm. Bộ dữ liệu hình ảnh của họ bao gồm 10.413 hình ảnh được tổng hợp từ sáu bộ dữ liệu khác nhau. Họ cũng sử dụng các kỹ thuật tiền xử lý và tăng cường dữ liệu để cải thiện khả năng phân loại của mạng. Họ đã loại bỏ tất cả các pixel không có màu xanh lục khỏi hình ảnh thông qua phân đoạn và bằng cách phản chiếu và xoay các hình ảnh gốc mà họ đã áp dụng tính năng tăng cường dữ liệu. Họ đạt được độ chính xác 86,20% trên tập thử nghiệm [11].

Guillermo L. Grinblat, và cộng sự. đã thay thế các giai đoạn phân loại và trích xuất tính năng của quy trình máy học trước đây bằng mô hình CNN để phân loại đậu trắng, đậu đỏ và đậu nành. Họ đã sử dụng các mô hình hình thái tĩnh mạch. Trong đường dẫn đề xuất của họ, các mẫu hình thái tĩnh mạch được trích xuất thông qua

UHMT và trích xuất bản vá trung tâm được áp dụng trên đầu ra ảnh nhị phân của phương pháp UHMT. Các giai đoạn này đảm bảo rằng thông tin về màu sắc và hình dạng sẽ bị xóa và chỉ các mẫu hình thái tĩnh mạch được sử dụng để phân loại. Đầu ra của các giai đoạn này được sử dụng làm đầu vào cho mô hình CNN và chúng có hai thiết lập được đặt tên là S1 và S2. Họ đã đào tạo một số mô hình CNN từ các mô hình có 2 lớp đến 6 lớp trên những hình ảnh về mô hình hình thái tĩnh mạch này. Họ đạt được kết quả tốt nhất ở độ sâu 5 lớp cho cả hai thiết lập. Đối với thiết lập S1, chúng đạt độ chính xác trung bình 92,6% và đối với thiết lập S2, chúng đạt độ chính xác trung bình 96,9% [12].

Kaya, A., và cộng sự. đã phân tích các phương pháp học chuyển giao và các kịch bản học chuyển giao khác nhau về nhận dạng hoa tự động. Họ đã thiết kế và thực hiện năm mô hình phân loại. Họ đã sử dụng bốn bộ dữ liệu khác nhau để đánh giá các mô hình và đã cho thấy tác động tích cực của việc sử dụng học chuyển giao trên các mô hình học sâu để cải thiện hiệu suất [13].

Sue Han Lee và cộng sự đã sử dụng Mạng thần kinh chuyển đổi để tìm các đặc điểm thể hiện tốt nhất các mẫu lá. Họ đã sử dụng mô hình CNN để tìm hiểu cách biểu diễn hình ảnh của những chiếc lá, sau đó họ phân tích các đặc điểm để tìm ra những đặc điểm quan trọng nhất bằng kỹ thuật trực quan hóa đặc điểm. Họ cũng định lượng các tính năng và tìm các tính năng cần thiết đại diện cho dữ liệu lá bằng cách đào tạo mô hình dựa trên CNN trên dữ liệu lá thô và sau đó áp dụng Mạng giải mã cho các đầu ra của mô hình dựa trên CNN. Họ đã phân tích cách CNN mô tả dữ liệu về lá và phát hiện ra rằng hình dạng của lá không phải là đặc điểm nổi bật mà là thứ tự gân lá. Họ cũng đề xuất các đặc điểm mới mô tả toàn bộ cấu trúc lá và các đặc điểm cục bộ tập trung vào gân lá.

Mario Lasseck đã trình bày các kỹ thuật học sâu cho LifeClef 2017. Trong tác phẩm này, Lasseck đã sử dụng Mạng chuyển đổi sâu và chuyển giao việc học thông qua tinh chỉnh để phân loại 10.000 loài và huấn luyện một số mô hình trên các bộ dữ liệu khác nhau. Ông cũng sử dụng việc đóng gói kết quả từ các mô hình và kỹ thuật tăng cường dữ liệu như cắt các mảng hình vuông từ mỗi hình ảnh ở các vị trí ngẫu nhiên, lật ngang, xoay, biến đổi ngẫu nhiên của độ bão hòa và biến đổi ngẫu nhiên của độ sáng được áp dụng. Hệ thống của ông đã đạt được thứ hạng tương hỗ trung bình (MRR) là 92% và độ chính xác thuộc top 5 là 96% trên bộ thử nghiệm chính thức của PlantCLEF [15].

Ghazi MM và cộng sự. đã sử dụng phương pháp học chuyển giao bằng cách tinh chỉnh các mô hình được đào tạo trước của GoogleLeNet, AlexNet và VGGNet trên tập dữ liệu LifeCLEF2015. Để giảm khả năng trang bị quá mức, họ cũng áp dụng các kỹ thuật tăng cường dữ liệu như xoay, dịch, phản chiếu và chia tỷ lệ. Họ đã hợp nhất các bộ phân loại khác nhau này để cải thiện hiệu suất tổng thể. Hệ thống kết hợp của họ đạt được độ chính xác 80% trên bộ xác thực [16].

Milan Sulc và cộng sự. đã đề xuất một hệ thống nhận dạng tự động gồm 10.000 loài thực vật cho ExpertLifeCLEF 2018. Hệ thống được đề xuất của họ dựa trên kiến trúc Mạng thần kinh chuyển đổi Inception-ResNet-v2 và Inception-v4 cũng như tập hợp của chúng. Họ đã sử dụng các kỹ thuật tiền xử lý như cắt ngẫu nhiên, lật trái phải ngẫu nhiên, biến dạng độ sáng và độ bão hòa. Thay vì sử dụng các giá trị sau kỳ nguyên huấn luyện cuối cùng, họ sử dụng giá trị trung bình đang chạy của các tham số mạng được huấn luyện và nhận thấy rằng điều đó làm tăng độ chính xác của mô hình của họ. Bởi vì tần số lớp trong dữ liệu huấn luyện tuân theo phân phối đuôi dài và do đó chúng có xác suất trước khác với tập kiểm tra nên họ đã sử dụng thuật toán EM cho

ước tính các ưu tiên của tập kiểm tra bằng cách tối đa hóa khả năng quan sát của tập kiểm tra. Kết quả là họ đạt được độ chính xác 88,4% trên toàn bộ bộ thử nghiệm [17].

### 3 Vật liệu và phương pháp

#### 3.1 Nguồn dữ liệu

Các hình ảnh trong bộ cơ sở dữ liệu được tổng hợp từ các tìm kiếm trên web và Bộ dữ liệu Oxford Flowers [18]. Tổng cộng, nó bao gồm 5345 hình ảnh hoa và cây. Ngoài ra, việc tăng cường dữ liệu đã được áp dụng cho tập dữ liệu đã tạo và nhờ việc tăng cường dữ liệu, kích thước của nó đã được tăng lên 37415 hình ảnh [19].

Có 76 loài thực vật khác nhau trong cơ sở dữ liệu hình ảnh. 64 loài hoa trong số đó và 12 loài thực vật khác. Các loài thuộc hệ thực vật Thổ Nhĩ Kỳ hoặc được mang từ nước ngoài vào Thổ Nhĩ Kỳ được chọn làm loài mục tiêu.

Nguồn dữ liệu thứ hai lưu trữ thông tin về loài trong cơ sở dữ liệu hình ảnh đã được tạo [19]. Mỗi loài có hồ sơ tương ứng trong nguồn dữ liệu này. Thông tin về các loài được tìm thấy trên các nguồn trực tuyến và được lưu trữ dưới dạng tệp JSON [20],[21]. Thông tin về loài bao gồm tên, họ, chi và nguồn gốc của chúng.

#### 3.2 Tăng cường dữ liệu

Sau khi tạo cơ sở dữ liệu hình ảnh, các kỹ thuật tăng cường dữ liệu đã được áp dụng cho hình ảnh. Kỹ thuật tăng cường dữ liệu rất hữu ích để tăng khả năng khái quát hóa của mô hình. Để tăng dữ liệu bốn lần xoay ngẫu nhiên, thay đổi độ sáng ngẫu nhiên và lật ngang đã được áp dụng.

Để xoay ngẫu nhiên, bốn góc đã được lấy mẫu thống nhất trong khoảng -90 và 90 độ mà không cần thay thế. Sau đó, mỗi hình ảnh đã được xoay bốn lần theo từng góc độ này. Sự phản chiếu đã được sử dụng như một phương pháp nội suy các giá trị pixel sau khi xoay.

Sau khi xoay ngẫu nhiên, thay đổi độ sáng ngẫu nhiên đã được áp dụng trong khoảng -0,2 và 0,2. Giá trị pixel của hình ảnh đầu ra thay đổi ngẫu nhiên trong phạm vi 0,2 lần dưới và trên giá trị pixel ban đầu.

Cuối cùng, thao tác lật ngang đã được áp dụng cho hình ảnh. Sau khi áp dụng các phép biến đổi này, tập dữ liệu đã tăng gấp 6 lần. Các phép biến đổi được cho trong Hình 1.



Hình 1. Ảnh gốc và ảnh chuyển đổi. (Xoay ngẫu nhiên, thay đổi độ sáng ngẫu nhiên và lật ngang có được áp dụng tương ứng cho ảnh gốc).

#### 3.3 Tiền xử lý

Các hoạt động tiền xử lý đã được áp dụng cho hình ảnh trước khi cung cấp chúng cho mô hình CNN làm đầu vào.

Để trích xuất các tính năng hữu ích nhất, hình ảnh được cắt xung quanh trung tâm. Sau khi cắt trung tâm, hình ảnh đã được thay đổi kích thước thành kích thước 224 x 224.

Cuối cùng, quá trình chuẩn hóa đã được áp dụng cho các hình ảnh trong tập dữ liệu và các giá trị pixel được chia tỷ lệ trong phạm vi -1 và 1. Trong thực tế, quá trình chuẩn hóa có thể tăng tốc độ học.

#### 3.4 Mô hình đề xuất

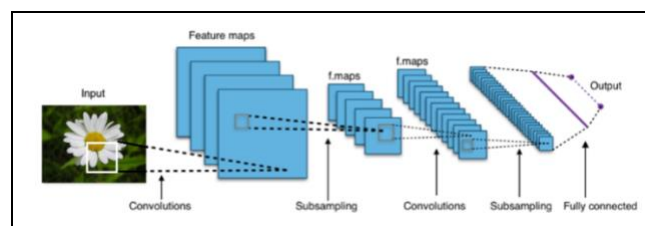
Phương pháp tiếp cận được đề xuất dựa trên Học chuyển giao. Transfer Learning có thể được sử dụng cho mạng thần kinh theo hai cách: bằng cách sử dụng mạng được đào tạo trước làm trình trích xuất tính năng hoặc bằng cách loại bỏ lớp cuối cùng của mạng được đào tạo trước và thêm bộ phân loại cho tác vụ được chỉ định. Cách thứ hai có thể coi là hai mạng được xếp chồng lên nhau. Tất cả hoặc một số tham số của mạng được huấn luyện trước có thể được sử dụng hoặc có thể được huấn luyện bằng các tham số của bộ phân loại được thêm vào. Cách tiếp cận này được gọi là tinh chỉnh [22].

Trong công việc này, phương pháp tinh chỉnh đã được sử dụng để chuyển việc học từ miền ImageNet sang nhiệm vụ của chúng tôi. Là mạng được đào tạo trước MobileNetV2 đã được sử dụng, mạng này đã được đào tạo trên cơ sở dữ liệu ImageNet.

MobileNet là một kiến trúc CNN nhẹ, hiệu quả nhằm vào các nền tảng di động và nhúng có tài nguyên tính toán hạn chế [23]. Trong kiến trúc này, các lớp tích chập được thay thế bằng cái gọi là các lớp tích chập có thể phân tách theo chiều sâu để thực hiện thao tác này một cách hiệu quả hơn.

Các lớp tích chập là xương sống của các mô hình CNN. Lớp tích chập đầu tiên lấy hình ảnh làm đầu vào và thực hiện thao tác tích chập trên hình ảnh bằng cách áp dụng bộ lọc tích chập. Số lượng bộ lọc tích chập là một siêu tham số và chúng cũng xác định độ sâu của đầu ra. Các lớp tích chập không phá vỡ cấu trúc không gian của hình ảnh đầu vào và mỗi bộ lọc được áp dụng sẽ tạo ra đầu ra hai chiều. Bằng cách xếp chồng các đầu ra hai chiều này lại với nhau, sẽ thu được bản đồ đặc trưng. Trong các lớp chập tiếp theo, đầu vào và đầu ra chỉ là bản đồ đặc trưng và độ sâu của các bản đồ đặc trưng này phụ thuộc vào số lượng bộ lọc trong mỗi lớp chập.

Hình 2 cho thấy các hoạt động của Mạng thần kinh chuyển đổi [24].



Hình 2. Hoạt động của lớp tích chập tiêu chuẩn.

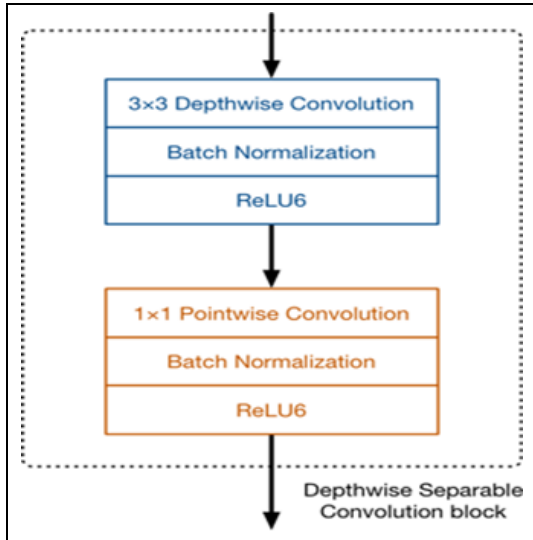
Tích chập có thể phân tách theo chiều sâu bao gồm hai thao tác: tích chập theo chiều sâu và tích chập theo điểm. Tích chập theo chiều sâu áp dụng một bộ lọc duy nhất cho mỗi kênh đầu vào, trái ngược với các lớp tích chập tiêu chuẩn trong đó bộ lọc được áp dụng cho tất cả các kênh cùng một lúc. Sau đó, tích chập theo chiều áp dụng tích chập 1x1 cho đầu ra của phép tích chập theo chiều sâu và kết hợp chúng [23]. Hoạt động của khối tích chập có thể phân tách theo chiều sâu được đưa ra trong Hình 3 [25].

Đầu ra của lớp chập tiêu chuẩn giả sử sai bước một và phần đệm được tính như sau [23]:

$$O_{x,y} = \sum_{i,j} I_{x+i,y+j} * W_{i,j} + b \quad (1)$$

Hoạt động tích chập tiêu chuẩn yêu cầu  $\times \times \times$  thông số ở đầu là chiều rộng và chiều cao không gian của bình phương

hạt nhân,  $C_{in}$  là số kênh đầu vào và  $C_{out}$  là số lượng kênh đầu ra. Tổng chi phí tính toán là  $\times \times \times \times \times$ , Trong đó  $F$  là chiều rộng và chiều cao không gian của bản đồ đặc trưng hình vuông.



Hình 3. Hoạt động của một phép chập có thể phân tách theo chiều sâu.

Trong khối tích chập có thể phân tách theo chiều sâu, tích chập theo chiều sâu được sử dụng để áp dụng một bộ lọc duy nhất cho mỗi kênh đầu vào và tích chập 1x1 được gọi là tích chập theo điểm sau đó được sử dụng để tạo ra sự kết hợp tuyến tính của đầu ra của lớp tích chập theo chiều sâu. Ngoài ra, chuẩn hóa hàng loạt và phi tuyến ReLU cũng được sử dụng cho cả hai lớp. Tích chập theo chiều sâu được tính như sau [23]:

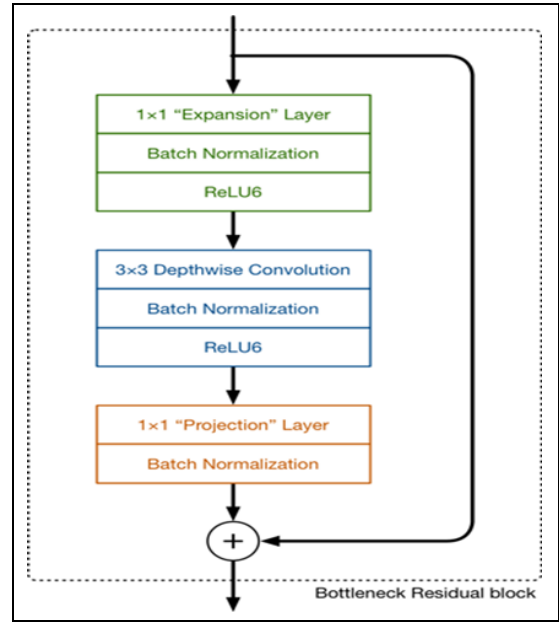
$$Z_{ij} = \sum_k X_{ik} * Y_{kj} + b_j \quad (2)$$

Tích chập theo chiều sâu có  $\times \times \times \times$  chi phí tính toán. Tích chập theo điểm cũng có  $\times \times \times$ , vì vậy tổng chi phí của khối phân tách theo chiều sâu là  $\times \times \times \times + \times \times \times$ . So sánh với hoạt động tiêu chuẩn của lớp tích chập, nó dẫn đến giảm tính toán:

$$\frac{K \cdot K \cdot C_i \cdot F \cdot F + C_i \cdot C_o \cdot F \cdot F}{K \cdot K \cdot C_i \cdot C_o \cdot F \cdot F} = \frac{1}{C_i} + \frac{1}{K^2} \quad (3)$$

MobileNetV2 có một số kiến trúc thay đổi qua MobileNetV1. Bên cạnh đó, nó vẫn sử dụng các phép tích chập theo chiều sâu và có thể phân tách theo chiều sâu, khối xây dựng chính của nó đã được thay đổi. Ngoài hai lớp trong khối có thể phân tách theo chiều sâu trước đó, một lớp chập 1x1 mới đã được giới thiệu trước lớp chập chiều sâu [23]. Do đó, lớp tích chập 1x1 đầu tiên làm tăng số lượng kênh của bản đồ đối tượng trước khi được đưa vào lớp tích chập theo chiều sâu và lớp tích chập 1x1 thứ hai làm giảm số lượng kênh của bản đồ tính năng đầu ra của lớp tích chập theo chiều sâu. Ngoài ra, một kết nối còn lại hoạt động giống như một đường tắt đã được đưa vào mạng.

Các kết nối ngắn giúp giảm bớt vấn đề biến mất độ dốc của các mạng sâu và cải thiện khả năng truyền độ dốc trên mạng [26]. Khối xây dựng chính của MobileNetV2, được gọi là khối dư thắt cổ chai, được đưa ra trong Hình 4 [25].



Hình 4. Khối xây dựng chính của MobileNetV2.

So với các khối có thể phân tách theo chiều sâu của MobileNetV1, các khối dư do thắt cổ chai có thêm một lớp tích chập 1x1; tuy nhiên kiến trúc mới cung cấp khả năng sử dụng kích thước đầu vào và đầu ra nhỏ hơn và hiệu quả hơn.

Mỗi lớp được lặp lại  $n$  lần trong kiến trúc.  $e$  biểu thị hệ số mở rộng,  $c$  biểu thị số kênh đầu ra và  $s$  biểu thị bước tiến được sử dụng.

Để tinh chỉnh, lớp cuối cùng của kiến trúc MobileNetV2 đã bị loại bỏ và thay vào đó là mạng thần kinh một lớp bộ phân loại bao gồm 76 nơ-ron đã được sử dụng.

Trong giai đoạn huấn luyện, các tham số của lớp đầu tiên của MobileNet được giữ không đổi và chỉ 20 lớp cuối cùng của nó và các tham số của bộ phân loại bổ sung được huấn luyện.

Hình ảnh đã được cung cấp làm đầu vào cho mạng theo đợt. Mỗi lô bao gồm 32 hình ảnh.

Hàm entropy chéo phân loại đã được sử dụng làm hàm mất mát để đánh giá quá trình huấn luyện và phương pháp tối ưu hóa Adamax [27] đã được sử dụng làm trình tối ưu hóa giảm độ dốc. Tỷ lệ học tập ban đầu đã được chọn là 0,02.

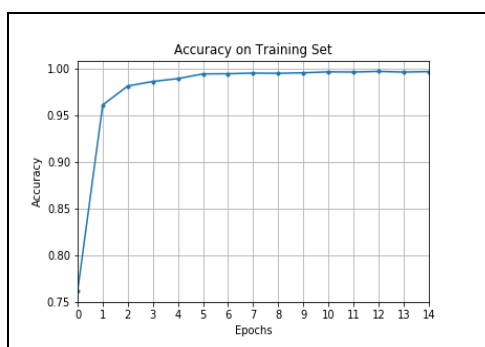
## 4 Kết quả thực nghiệm

Mô hình này được phát triển bằng ngôn ngữ lập trình Python và được đào tạo trên Google Colab để sử dụng GPU. Keras, OpenCV và nhiều thư viện khác được sử dụng. Chúng tôi cũng đã phát triển một ứng dụng nhận dạng thực vật di động dựa trên mô hình đề xuất cho nền tảng Android [19].

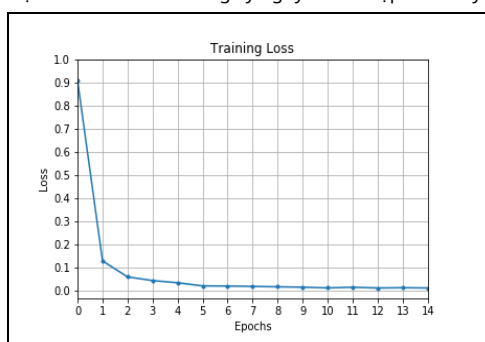
Trước giai đoạn huấn luyện và kiểm tra, các mẫu trong cơ sở dữ liệu đã được xáo trộn và sau đó được chia thành các tập huấn luyện và kiểm tra bằng cách lấy mẫu phân tầng. Tập huấn luyện bao gồm 29932 hình ảnh và tập kiểm tra bao gồm 7483 hình ảnh. Tỷ lệ hình ảnh trong bộ thử nghiệm so với tất cả hình ảnh trong cơ sở dữ liệu là 1/5.

Trong quá trình huấn luyện, tỷ lệ chính xác đạt được là 0,9971 trên tập huấn luyện sau 15 kỷ nguyên. Trong mỗi tập kiểm tra kỷ nguyên cũng được đánh giá. Đồ thị về độ chính xác và độ mất mát của tập huấn luyện được đưa ra tương ứng trong Hình 5 và Hình 6.



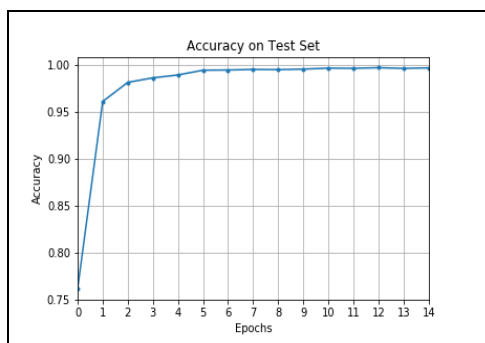


Hình 5. Độ chính xác đối với từng kỷ nguyên trên tập huấn luyện.

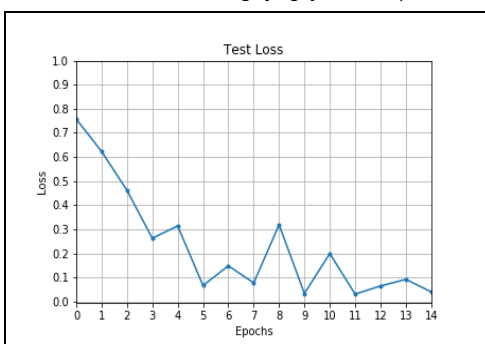


Hình 6. Mất mát theo từng kỷ nguyên trên tập huấn luyện.

Mô hình đã được thử nghiệm trên tập kiểm tra sau mỗi kỷ nguyên và kết quả là đạt được tỷ lệ chính xác 0,9897 trên tập kiểm tra. Đồ thị độ chính xác và độ mất mát của bộ thử nghiệm được cho trong Hình 7 và Hình 8.

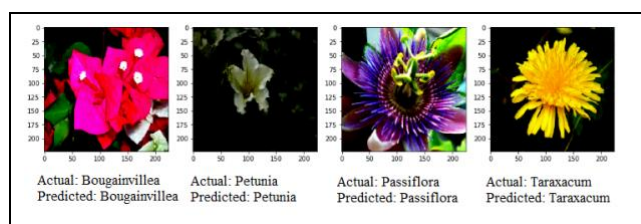


Hình 7. Độ chính xác đối với từng kỷ nguyên trên tập kiểm tra.



Hình 8. Mất mát theo từng kỷ nguyên trên tập kiểm tra.

Sau khi đào tạo, một số ví dụ đã được rút ngẫu nhiên từ bộ kiểm tra và được sử dụng để dự đoán nhằm đánh giá mô hình một cách trực quan. Một số mẫu từ bộ thử nghiệm được cho trong Hình 9. Chúng đã được dự đoán chính xác.



Hình 9. Một số bài kiểm tra trực quan bổ sung trên bộ bài kiểm tra.

Ngoài ra, các số liệu về độ chính xác, thu hồi và điểm f1 đã được đánh giá cho từng lớp. Các thuật ngữ dương và âm đề cập đến các dự đoán của bộ phân loại và các thuật ngữ đúng và sai đề cập đến việc lớp được dự đoán có đúng hay không [28]. Tính toán của họ như sau:

$$= \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

$$= \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

So sánh với các nghiên cứu khác được đưa ra trong Bảng 1. Các số liệu về độ chính xác, thu hồi và điểm F1 trong nội bộ lớp được cung cấp trong phụ lục A.

Bảng 1. So sánh với các nghiên cứu khác.

tác giả	Số lượng	Tập dữ liệu	Bài kiểm tra
	Các lớp học	Kích cỡ	Sự chính xác
Dyrmann và cộng sự.	22	10.413	86,2%
Lasseck M.	10000	1,7M	96%
Šulc M và cộng sự.	10000	1,75 triệu	88,4%
Ghazi MM và cộng sự.	1000	113.204	80%
Phương án đề xuất	76	37.415	98%

Ngoài ra, một ứng dụng nhận dạng cây trồng di động dựa trên mô hình đề xuất cũng đã được phát triển cho nền tảng Android. Ứng dụng này đã được phát triển bằng ngôn ngữ lập trình Java và mô hình được đào tạo đã được tích hợp vào đó. Nó cho phép người dùng sử dụng hình ảnh từ thư viện của họ hoặc chụp ảnh bằng máy ảnh của họ và ứng dụng sẽ đưa ra dự đoán về loài trong ảnh. Một số thông tin như tên, họ, chi và nguồn gốc được hiển thị cho người dùng dựa trên dự đoán. Ứng dụng này đã được thử nghiệm trên hoa dại và thực vật. Một mẫu được đưa ra trong Hình 10.



Hình 10. Một mẫu từ ứng dụng di động của chúng tôi đã được thử nghiệm trên hoang dã.

## 5. Kết Luận

Trong bài báo này, một mô hình nhận dạng thực vật dựa trên Mạng tích chập và Học chuyển giao đã được đề xuất. Nó đạt được độ chính xác 0,9971 trên tập huấn luyện và tỷ lệ chính xác 0,9897 trên tập kiểm tra.

Tóm lại, các mô hình CNN bằng cách sử dụng phương pháp học chuyển giao và ứng dụng di động để nhận dạng thực vật đã được phát triển. Chúng tôi hy vọng rằng công việc này sẽ hữu ích trong lĩnh vực nhận dạng thực vật và hoa.

## 6 Lời cảm ơn

Các tác giả xin cảm ơn Ban biên tập và những người phản biện ẩn danh đã có những nhận xét, góp ý quý báu nhằm nâng cao chất lượng bản thảo.

## 7 Tuyên bố đóng góp của tác giả

Trong phạm vi nghiên cứu này, Tolgahan KARAHAN, trong việc xem xét tài liệu, thiết kế, thu thập dữ liệu, chỉ định ngăn xếp công nghệ, triển khai, chính tả và kiểm tra bài viết về mặt nội dung; Vasif NABIYEV, trong quá trình hình thành ý tưởng, đánh giá tài liệu, đánh giá các giai đoạn của nghiên cứu, đề xuất nội dung, kiểm tra kết quả và tư vấn đã được đóng góp.

## 8 Sự chấp thuận của ủy ban đạo đức và xung đột của báo cáo lỗi suất

Không cần phải xin phép ủy ban đạo đức đối với bài viết được chuẩn bị. Không có xung đột lợi ích với bất kỳ cá nhân/ tổ chức nào trong bài viết được chuẩn bị.

## 9 tài liệu tham khảo

- [1] O'Shea K và cộng sự. "Giới thiệu về Mạng thần kinh chuyển đổi". <https://arxiv.org/abs/1511.08458> (31.03.2021).
- [2] Hình ảnhNet. "Thử thách nhận dạng hình ảnh quy mô lớn của ImageNet". <http://image-net.org/challenges/LSVRC/> (01.05.2020).
- [3] Zeiler MD, Fergus R. "Trực quan hóa và tìm hiểu các mạng tích chập". *ECCV 2014*. Zürich, Thụy Sĩ, Ngày 6-12 tháng 9 năm 2014.
- [4] Goodfellow I, Bengio Y, Courville A. *Chuyển giao học tập*. Biên tập viên: Dietterich T. Deep Learning, 526-531, Cambridge, Massachusetts, USA, MIT Press, 2016.
- [5] Rút gọn C, Khoshgoftaar TM. "Khảo sát về Tăng cường dữ liệu hình ảnh cho học sâu". *Tạp chí Big Data*, năm 2019. <https://doi.org/10.1186/s40537-019-0197-0>.
- [6] Lohr S. *Lấy mẫu: Thiết kế và phân tích*. Biên tập viên: Blitzstein SK, Faraway JJ, Tanner M, Zidek J. Lấy mẫu phân tầng, 73-115, Boca Raton, FL, Hoa Kỳ, CRC Press, 2012.
- [7] Krizhevsky A, Sutskever I, Hinton GE. "Phân loại Imagenet với mạng lưới thần kinh tích chập sâu". *Hệ thống xử lý thông tin thần kinh*. Hồ Tahoe, Nevada, Mỹ, ngày 3-6 tháng 12 năm 2012.
- [8] Khan A, Sohail A, Zahoora U, Qureshi AS. "Một cuộc khảo sát về các kiến trúc gần đây của mạng lưới thần kinh tích chập sâu". *Đánh giá trí tuệ nhân tạo*, 2020. <https://doi.org/10.1007/s10462-020-09825-6>.
- [9] Tan C, Sun F, Kong T, Zhang W, Yang C, Liu C. "Một cuộc khảo sát về học chuyển giao sâu". *Hội nghị quốc tế về mạng lưới thần kinh nhân tạo*. Rhodes, Hy Lạp, ngày 4-7 tháng 10 năm 2018.
- [10] Wang J, Perez L. "Hiệu quả của việc tăng cường dữ liệu trong phân loại hình ảnh bằng cách sử dụng học sâu". <https://arxiv.org/abs/1712.04621> (31.03.2021).
- [11] Dyrmann M, Karstoft H, Midtby HS. "Phân loại loài thực vật bằng cách sử dụng mạng lưới thần kinh tích chập sâu". *Kỹ thuật hệ thống sinh học*, 151, 72-80, 2016.
- [12] Grinblat GL, Uzal LC, Larese MG. "Học sâu để nhận dạng thực vật bằng cách sử dụng các mẫu hình thái tĩnh mạch". *Máy tính và Điện tử trong Nông nghiệp*, 27, 418-424, 2016.
- [13] Kaya A, Keçeli AS, Çatal Ç, Yalıcı HY, Huseyin T, Tekinerdoğan B. "Phân tích quá trình học chuyển giao cho các mô hình phân loại thực vật dựa trên mạng lưới thần kinh sâu". *Máy tính và Điện tử trong Nông nghiệp*, 158, 20-29, 2019.
- [14] Lee HS, Chan CS, Mayo S, Remagnino P. "Cách deep learning trích xuất và tìm hiểu các đặc điểm của lá để phân loại thực vật". *Nhận dạng mẫu*, 71, 1-13, 2017.
- [15] Lasseck M. "Xác định các loài thực vật dựa trên hình ảnh bằng mạng lưới thần kinh tích chập sâu". *CLEF*, Dublin, Ireland, ngày 11-14 tháng 9 năm 2017.
- [16] Ghazi MM, Yanikoğlu B, Aptoula E. "Nhận dạng thực vật bằng cách sử dụng mạng lưới thần kinh sâu thông qua tối ưu hóa các tham số học chuyển giao". *Điện toán thần kinh*, 235, 228-235, 2017.
- [17] Šulc M, Pícek L, Matas J. "Nhận dạng thực vật bằng mạng khởi động với ước tính trước của lớp thời gian thử nghiệm". *CLEF*, Avignon, Pháp, ngày 10-14 tháng 9 năm 2018.
- [18] Đại học Oxford. "Bộ dữ liệu hoa loại 102". <http://www.robots.ox.ac.uk/~vgg/data/flowers/102/index.html>, (15.06.2019).
- [19] Github. "Tolga-Karahan/Nhận dạng thực vật di động". <https://tolga-karahan.github.io/Mobile-Plant-Identification/> (13.11.2020).
- [20] Atatürk Orman Çiftliği. "Atatürk Orman Çiftliği". <http://www.aoc.gov.tr>, (08.09.2020).
- [21] Yabani Çiçekler. "Yaban Çiçekler". [www.yabanicekler.com](http://www.yabanicekler.com), (08.09.2020).
- [22] Yosinski J, Clune J, Bengio Y, Lipson H. "Các tính năng có thể chuyển nhượng được trong Mạng lưới thần kinh sâu như thế nào?". <https://arxiv.org/abs/1411.1792> (31.03.2021).
- [23] Howard AG và cộng sự. "MobileNets: Mạng thần kinh tích chập hiệu quả cho các ứng dụng thị giác di động". <https://arxiv.org/abs/1704.04861> (31.03.2021).
- [24] Wikipedia Commons. "Tệp: cnn.png điển hình". [https://commons.wikimedia.org/wiki/File:Typical\\_cnn.png](https://commons.wikimedia.org/wiki/File:Typical_cnn.png) (12.11.2020).
- [25] Máy tư duy. "Mạng di động Phiên bản 2". <https://machinethink.net/blog/mobilenet-v2/> (05.05.2020).
- [26] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L. "Mobilenetv2: Phân dư đảo ngược và tắc nghẽn tuyến tính". *CVPR*, Thành phố Salt Lake, Hoa Kỳ, ngày 19-21 tháng 6 năm 2018.
- [27] Kingma DP, Ba JL. "Adam: Một phương pháp tối ưu hóa ngẫu nhiên". *ICLR*, San Diego, Hoa Kỳ, ngày 7-9 tháng 5 năm 2015.
- [28] Kelleher JD, Mac Namee B, D'arcy A. *Nguyên tắc cơ bản của Machine Learning để phân tích dữ liệu dự đoán*. Biên tập viên: Mellon M. Precision, Recall, và F1 Measure, 413-417, Cambridge, Massachusetts, USA, The MIT Press, 2015.

**Phụ lục A Độ chính xác, Thu hồi và Điểm F1**

Lớp học	Độ chính xác	Nhớ lại	điểm f1
Cây mao lương	1.000	1.000	1.000
Cây sơn tùng Tomentosa	0,968	0,968	0,968
Echinops	1.000	1.000	1.000
Tussilago	1.000	1.000	1.000
Aconitum	1.000	1.000	1.000
dianthus	1.000	1.000	1.000
hoa huệ	1.000	0,986	0,993
Gazania	0,990	1.000	0,995
Vriesea	0,986	1.000	0,993
Dieffenbachia	1.000	1.000	1.000
Cirsium	1.000	1.000	1.000
Eschscholzia	1.000	1.000	1.000
Helleborus	1.000	0,988	0,994
cà độc dược	1.000	0,987	0,993
cây bìm bìm	1.000	0,899	0,947
Cây thông	1.000	1.000	1.000
Euphorbia	1.000	1.000	1.000
hoa oải hương	0,981	0,981	0,981
Nigella	1.000	0,943	0,971
Taraxacum	1.000	1.000	1.000
Iris giả hành	1.000	1.000	1.000
Trollius	1.000	1.000	1.000
thuốc chống sốt rét	0,965	1.000	0,982
Gaillardia	1.000	0,982	0,991
Montera	0,955	1.000	0,977
hoa giấy	1.000	0,994	0,997
Muscari	1.000	1.000	1.000
thuộc dược	1.000	1.000	1.000
Lathyrus	1.000	0,957	0,978
Sansevieria	0,987	1.000	0,993
Zantedeschia	1.000	1.000	1.000
hoa thủy tiên	0,987	1.000	0,994
bệnh thận	1.000	0,961	0,980
Bellis	1.000	1.000	1.000
cây xương rồng	0,989	1.000	0,994
Echinacea	1.000	1.000	1.000
lịch	1.000	0,988	0,994
phlox	0,855	0,985	0,915
vinh quang	1.000	1.000	1.000
Cycas	0,982	0,964	0,973
hoa huệ	1.000	1.000	1.000
Cây huyết dụ	0,961	1.000	0,980
giống anh thảo	0,961	1.000	0,980
tuổi già	0,987	1.000	0,994
hoa tulip	1.000	0,986	0,993
Dracaena Fragrans Massengena	0,974	0,987	0,981
Adenium	0,925	1.000	0,961
canna	1.000	0,989	0,995
Colchi	0,981	1.000	0,991
Passiflora	1.000	1.000	1.000
Papaver	1.000	1.000	1.000
Nymphaea	0,996	1.000	0,998
hoa anh thảo	1.000	0,974	0,987
gentiana	0,923	1.000	0,960
Hedera	0,957	1.000	0,978
Astrantia	0,985	1.000	0,993
Tuyến ức	1.000	0,859	0,924
lưỡng hình	1.000	0,988	0,994
Đỗ quỳ	0,966	0,991	0,979
Cynara	0,960	0,990	0,975
hồng môn	1.000	1.000	1.000
hoa mộc lan	1.000	0,988	0,994
Fritillaria	1.000	0,980	0,990
Oenothera	1.000	1.000	1.000
màu hồng	1.000	0,977	0,989
Platycodon	0,982	0,949	0,966
Viola	0,984	1.000	0,992



### Phụ lục A tiếp theo.

Lớp học	Độ chính xác	Nhớ lại	điểm f1
Pelargonium	0,965	1.000	0,982
cây dã yên thảo	1.000	0,994	0,997
kỹ thuật số	1.000	1.000	1.000
Aechmea	1.000	1.000	1.000
Tagetes	1.000	1.000	1.000
hoa dâm bụt	0,988	1.000	0,994
tối đa	0,961	0,961	0,961
Hoa hướng dương	0,988	1.000	0,994
Cây lay ơn	1.000	0,925	0,961