

Hệ thống nhận dạng thực vật dựa trên  
Mạng lưới thần kinh chuyển đổi cho LifeClef  
Nhiệm vụ phân loại thực vật năm 2016

Sue Han Lee, Yang Loong Chang, Chee Seng Chan, và Paolo Remagnino†

Trung tâm xử lý hình ảnh và tín hiệu, Fac. Comp. Khoa học. & Thông  
tin. Tech., Đại học Malaya,  
Malaysia #Khoa Khoa học Máy tính, Đại học Kingston, Vương quốc  
Anh {leesuehan,yangloong}@siswa.um.edu.my  
cs.chan@um.edu.my  
p.remagnino@kingston.ac.uk

Trừu tượng. Trong bài viết này, chúng tôi mô tả kiến trúc của hệ thống phân loại nhà máy cho thử thách LifeClef 2016 [14]. Mục tiêu của nhiệm vụ là xác định 1000 loài hình ảnh thực vật tương ứng với 7 cơ quan thực vật khác nhau, cũng như tự động phát hiện các loài xâm lấn thuộc các lớp chưa biết. Để giải quyết thách thức [10], chúng tôi đã đề xuất một hệ thống phân loại thực vật sử dụng mạng nơ ron tích chập (CNN).

Từ khóa: Phân loại thực vật, deep learning, mạng nơ ron tích chập

## 1. Giới thiệu

Phân loại thực vật đã nhận được sự quan tâm đặc biệt trong lĩnh vực thị giác máy tính do ý nghĩa quan trọng của nó trong tự động hóa nông nghiệp cũng như bảo tồn môi trường. Ví dụ, kiến thức thực vật học về thực vật là cần thiết để cải thiện sự phát triển nông nghiệp. Các nhà nghiên cứu về thị giác máy tính [19, 17, 16, 2, 6, 12, 20] đã sử dụng các biến thể về đặc điểm của lá làm công cụ so sánh để phân loại thực vật. Nguyên nhân là do ký tự lá đã được sử dụng rộng rãi trong phân loại văn bản truyền thống; chúng là chìa khóa để nhận dạng thực vật kể từ những ngày đầu của khoa học thực vật [5, 4]. Mặc dù các đặc điểm cấu trúc của lá đóng vai trò quan trọng trong nhiệm vụ nhận dạng thực vật, nhưng đối với một số loài thực vật nhất định, chẳng hạn như cây rụng lá hoặc cây bán thường xanh, lá không có sẵn trong các thời kỳ khác nhau trong năm. Hơn nữa, một số loài khó có thể phân biệt được nếu chỉ sử dụng cơ quan lá vì lá trong tự nhiên có thể có hình dạng và màu sắc rất giống nhau [20, 9]. Vì vậy, các nhà thực vật học thường mở rộng sự quan sát của họ đến nhiều cơ quan như thân, hoa, cành hoặc quả. Kể từ năm 2013, thử thách LifeClef [11] đã cung cấp bộ dữ liệu thực vật đa cơ quan đầu tiên không chỉ bao gồm các hình ảnh dựa trên lá mà còn các cơ quan khác nhau của từng cây cụ thể. Những hình ảnh thực vật như vậy được thu thập trong một môi trường không bị hạn chế, vào những khoảng thời gian khác nhau trong năm, bởi những người dùng khác nhau. Mục tiêu của

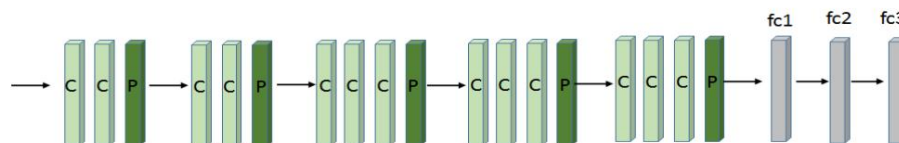
Nhiệm vụ nhận dạng thực vật [10] trong thử thách LifeClef 2016 [14] là xác định 1000 loài hình ảnh thực vật tương ứng với 7 cơ quan khác nhau, cũng như tự động phát hiện các loài xâm lấn thuộc các lớp chưa xác định.

Lấy cảm hứng từ sự đột phá của deep learning trong phân loại hình ảnh, nhiều nhà nghiên cứu đã bắt đầu sử dụng các mô hình deep learning như CNN để tìm hiểu cách biểu diễn hình ảnh thực vật mạnh mẽ [1, 21, 3, 8, 24]. Trong công việc này, chúng tôi sử dụng mô hình CNN để xây dựng hệ thống phân loại thực vật. Chúng tôi tái sử dụng mạng VGG tiên tiến nhất hiện nay [22] để kết hợp các đặc điểm loài và cơ quan, đồng thời giải quyết vấn đề phân loại thực vật đa cơ quan.

Phần còn lại của bài báo được tổ chức như sau. Trong Phần 2, chúng tôi trình bày phương pháp luận của kiến trúc được đề xuất. Phần 3 minh họa chương trình đào tạo của nó. Phần 4 trình bày các thử nghiệm và kết quả cho cả bộ xác nhận và bộ thử nghiệm. Cuối cùng, Phần 5 trình bày kết luận và công việc tiếp theo.

## 2 Mô tả phương pháp

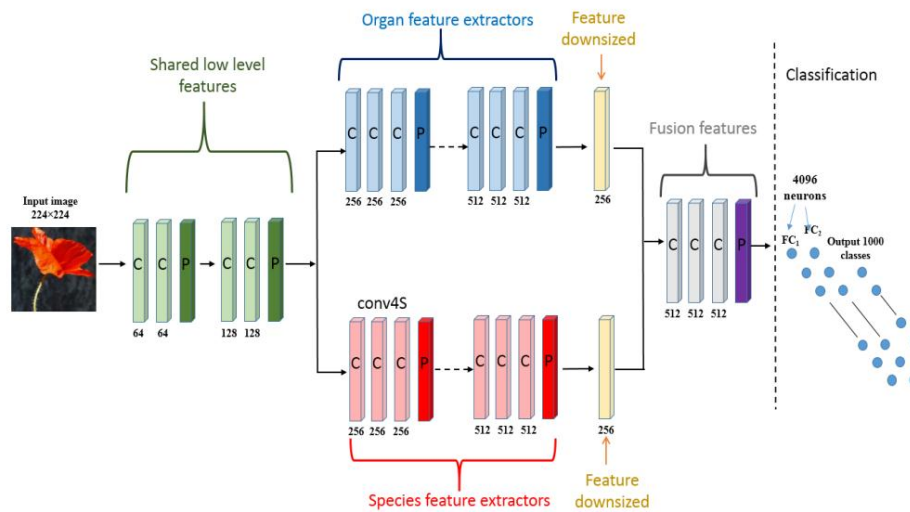
Mô hình CNN ban đầu được thiết kế để xử lý nhiều mảng dữ liệu như hình ảnh màu (RGB), tín hiệu hoặc chuỗi cũng như video. Do khả năng sẵn có của các bộ dữ liệu hình ảnh quy mô lớn, chẳng hạn như ImageNet [7], và tiếp theo là sự tiến bộ của công nghệ, chẳng hạn như Bộ xử lý đồ họa (GPU), CNN hiện là một mặt hàng trong lĩnh vực thị giác máy tính. Mạng VGG [22] hiện cung cấp kết quả tiên tiến nhất để phân loại hình ảnh.



Hình 1: Kiến trúc 16 lớp VGG-net

Hình 1 mô tả cấu hình của mạng VGG 16 lớp được đề xuất. Trong công việc này, chúng tôi khởi tạo kiến trúc mô hình của mình dựa trên mạng VGG và sửa đổi lớp chập cấp cao hơn của nó để tìm hiểu các đặc điểm cơ quan và loài kết hợp. Hình 2 cho thấy kiến trúc được đề xuất của chúng tôi. Chúng tôi không tạo thủ công bất kỳ bộ mô tả tính năng nào cho các tính năng tổng hợp mà chỉ giới thiệu các lớp tích chập để tự tìm hiểu các bộ lọc. Chúng ta có thể xem các bộ lọc này như các bộ mô tả tính năng đã học mã hóa các cấu trúc hợp nhất đặc biệt.

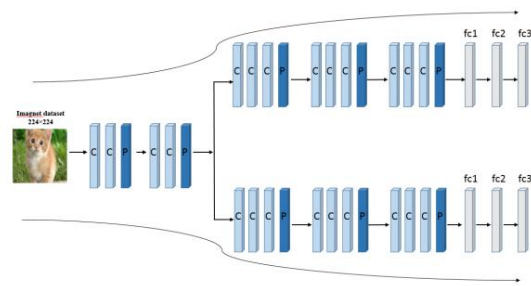
Kiến trúc của chúng tôi chủ yếu bao gồm bốn thành phần: (i) các lớp dùng chung, (ii) các lớp cơ quan, (iii) các lớp loài và (iv) các lớp hợp nhất xử lý sự kết hợp của cả các đặc điểm loài và cơ quan. Chúng tôi giới thiệu các lớp dùng chung cho cả loài và thành phần cơ quan. Những lý do có ba phần. Đầu tiên, [25, 23] chứng minh rằng các lớp trước đó trong mạng sâu phản ứng với các tính năng cấp thấp như góc và cạnh. Vì cả loài cấp cao hơn hoặc các thành phần cơ quan



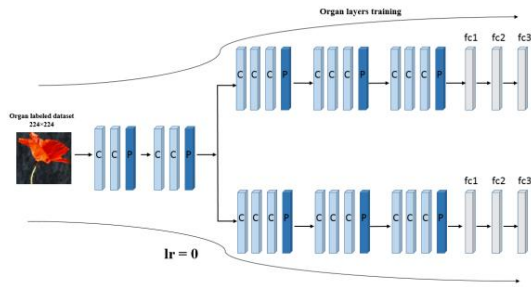
Hình 2: Kiến trúc tổng hợp cấp cao của loài-cơ quan

yêu cầu các tính năng cấp thấp để xây dựng các tính năng cấp cao hơn, chúng tôi giới thiệu lớp tiền nhượng quyền chung cho cả hai thành phần. Thứ hai, theo [23], các lớp chia sẻ có thể giảm cả hoạt động đầu phẩy động và dung lượng bộ nhớ khi thực thi mạng, điều này có tầm quan trọng đối với ứng dụng trong thế giới thực. Cuối cùng, việc sử dụng các lớp chia sẻ giúp giảm số lượng tham số huấn luyện, điều này có lợi cho hiệu suất kiến trúc.

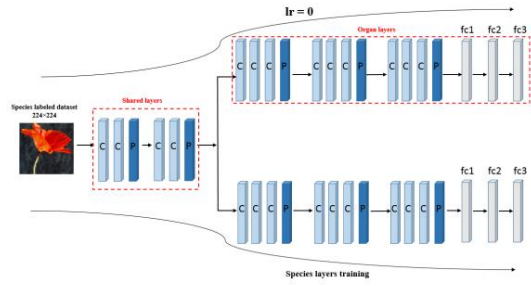
Để kết hợp cả đặc điểm cơ quan và loài, trước tiên chúng tôi huấn luyện các lớp cơ quan CNN dựa trên các lớp cơ quan. Sau đó, chúng tôi giữ nguyên các lớp chia sẻ và các lớp cơ quan, để có thể tái sử dụng nó nhằm huấn luyện các lớp loài. Sau khi huấn luyện các lớp loài, chúng tôi xếp tầng cả hai lớp để tìm hiểu các tính năng kết hợp. Trước khi xếp tầng cả hai đối tượng, một lớp tích chập thu nhỏ đối tượng sẽ được thêm vào trong mỗi lớp để giảm kích thước bản đồ đối tượng và tạo ra các đối tượng dựa trên nhỏ gọn. Bước này rất cần thiết để giảm số lượng tham số huấn luyện, bù đắp cho vấn đề trang bị quá mức. Cuối cùng nhưng không kém phần quan trọng, chúng tôi đào tạo ba lớp được kết nối đầy đủ làm bộ phân loại để phân loại hình ảnh đầu vào thành các lớp loài tương ứng. Để những bốn thành phần vào một quy trình và được đào tạo chung từ đầu đến cuối, chúng tôi sử dụng quy trình đào tạo nhiều bước như được nêu trong Phần. 3.



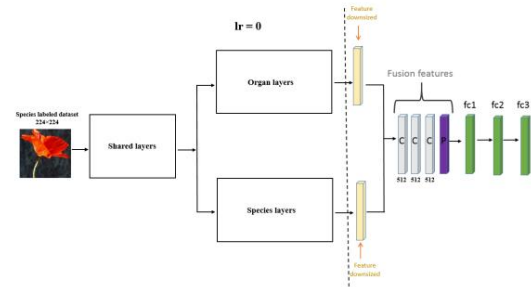
(a) CNN hai đường dẫn huấn luyện trước



(b) Khởi tạo các lớp nội tạng



(c) Khởi tạo các lớp loài



(d) Khởi tạo các lớp kết hợp

Hình 3: Sơ đồ đào tạo nhiều bước cho sự kết hợp cấp độ cao giữa các loài nội tạng

### 3 Đào tạo

Thuật toán bên dưới hiển thị quy trình đào tạo của kiến trúc tổng hợp cấp cao được đề xuất:

---

Thuật toán của kiến trúc tổng hợp cấp cao được đề xuất của chúng tôi:

---

- Quy trình đào tạo kiến trúc tổng hợp cấp cao
- Bước 1: Đào tạo trước CNN hai đường dẫn
  - Bước 2: Khởi tạo các lớp Organ
  - Bước 3: Khởi tạo lớp loài
  - Bước 4: Khởi tạo các lớp kết hợp
  - Bước 5: Tinh chỉnh sự kết hợp cấp cao giữa các loài cơ quan

Bước 1: Huấn luyện trước CNN hai đường dẫn Ban đầu, chúng tôi thiết kế CNN hai đường dẫn nhằm mục đích huấn luyện hai thành phần khác nhau (loài và cơ quan) như trong Hình 3a. Mỗi đường dẫn của cấu hình CNN tương tự như kiến trúc 16 lớp VGG-net [22], ngoại trừ việc mỗi đường dẫn đó chia sẻ các lớp trước đó.

Bước 2: Khởi tạo các lớp cơ quan Sau khi chúng tôi đã huấn luyện trước CNN hai đường dẫn với bộ dữ liệu Imagenet, chúng tôi tái sử dụng một trong các đường dẫn để huấn luyện lớp cơ quan như trong Hình 3b. Chúng tôi thực hiện tinh chỉnh với bảy nhãn cơ quan: nhánh, toàn bộ, hoa, quả, lá, quét lá và thân. Các nhãn cơ quan này được ký hiệu an cùng với các loài trong Bộ dữ liệu PlantClef. Chúng tôi tinh chỉnh mạng VGG-16 bằng cách thay thế lớp được kết nối đầy đủ cuối cùng bằng tổng cộng bảy nơtron tương ứng với bảy lớp.

Bước 3: Khởi tạo các lớp loài Sau khi có được các lớp cơ quan, chúng tôi huấn luyện các lớp loài dựa trên tập dữ liệu được gắn nhãn loài như trong Hình 3c. Như đã đề cập trong Sec. 2, chúng tôi cho phép cả loài và lớp cơ quan chia sẻ các lớp tiến hành chung. Do đó, để có thể chia sẻ các bộ lọc, chúng tôi giữ nhất quán trọng số của hai lớp chập đầu tiên bằng cách đặt tốc độ học về 0 trong quá trình huấn luyện các lớp loài. Ngoài ra, chúng tôi đặt tốc độ học tập của các lớp cơ quan thành 0 để tránh việc bộ lọc của chúng bị thay đổi trong quá trình huấn luyện các lớp loài.

Bước 4: Khởi tạo các lớp kết hợp Sau khi tinh chỉnh cả thành phần cơ quan và loài trên CNN hai đường, trước tiên, chúng tôi di chuyển các lớp tích chập của nó sang một kiến trúc mới như trong Hình 3d. Sau đó, chúng tôi thêm một lớp chập vào mỗi thành phần để giảm kích thước bản đồ đặc trưng và tiếp theo là hai lớp tích chập khác để tìm hiểu các tính năng kết hợp. Cuối cùng, chúng tôi chỉ định ba lớp được kết nối đầy đủ để phân loại loài. Khi đào tạo, chúng tôi đặt tốc độ học tập của các lớp được xác định trước thành 0 và chỉ huấn luyện các lớp tích chập mới được chỉ định với tập dữ liệu được gắn nhãn loài. Thông thường, chúng ta có thể coi bước thứ ba này là giai đoạn tìm hiểu tính năng tổng hợp.

Bước 5: Tinh chỉnh sự kết hợp cấp cao giữa các loài cơ quan Cuối cùng, chúng tôi tinh chỉnh toàn bộ kiến trúc từ đầu đến cuối bằng cách sử dụng cùng một tốc độ học tập.

## 4 thí nghiệm và kết quả

Chúng tôi đánh giá các kiến trúc trên bộ dữ liệu PlantClef2015 [15]. Các mô hình được huấn luyện bằng phần mềm Caffe [13]. Đối với cài đặt tham số trong đào tạo, chúng tôi sử dụng chính sách học cố định. Chúng tôi đặt tốc độ học thành 0,01, sau đó giảm hệ số 10 khi độ chính xác của bộ xác thực ngừng cải thiện. Động lượng được đặt thành 0,9 và trọng lượng giảm dần thành 0,0001. Các mạng được huấn luyện với phương pháp lan truyền ngược, sử dụng phương pháp giảm độ dốc ngẫu nhiên [18]. Chúng tôi chạy thử nghiệm bằng nhiều GPU trên hai card đồ họa NVIDIA TitanX.

### 4.1 Tăng cường dữ liệu

Các hình ảnh được sử dụng để đào tạo chụp thực vật sống, trong đó mọi đối tượng trong ảnh có thể được chụp ở các kích cỡ khác nhau. Do đó, đào tạo đa quy mô được đề xuất. Chúng tôi thay đổi tỷ lệ đẳng hướng các hình ảnh huấn luyện thành ba kích thước khác nhau: 256, 385 và 512, sau đó cắt ngẫu nhiên 224 \* 224 pixel từ các hình ảnh đã được thay đổi tỷ lệ để đưa vào mạng phục vụ huấn luyện. Do đó, phần cắt từ các hình ảnh có tỷ lệ lớn hơn sẽ tương ứng với phần nhỏ của hình ảnh hoặc phần phụ đặc biệt của cơ quan có thể là một đặc điểm quan trọng để nhận dạng. Bên cạnh đó, chúng tôi cũng tăng kích thước dữ liệu bằng cách phản chiếu hình ảnh đầu vào trong quá trình đào tạo. Cuối cùng, chúng ta có một bộ ảnh huấn luyện mới chứa 272892 hình ảnh và một bộ xác thực gồm 66711 hình ảnh.

### 4.2 Kết quả thực nghiệm trên bộ xác nhận

Để đánh giá bộ xác thực của chúng tôi, chúng tôi trực tiếp sử dụng đầu ra softmax từ mô hình, tức là chúng tôi gán cho mỗi hình ảnh thử nghiệm nhãn có đầu ra softmax tối đa từ các bộ phân loại và đo số lượng nhãn được gán chính xác trên tất cả các hình ảnh thử nghiệm.

Đóng góp của việc đào tạo trước Imagenet. Trong phần này, chúng tôi đánh giá sự đóng góp của việc huấn luyện trước bộ dữ liệu Imagenet [7] cho nhiệm vụ phân loại thực vật. Chúng tôi thực hiện các thử nghiệm học chuyển giao trên bộ dữ liệu PlantClef2015. Chúng tôi tái sử dụng lưới VGG-16 bằng cách thực hiện tinh chỉnh trên lớp được kết nối đầy đủ trên cùng. Chúng tôi so sánh nó với mô hình được đào tạo trực tiếp từ đầu bằng cách sử dụng bộ dữ liệu PlantClef2015 mà không cần đào tạo trước. Kết quả phân tích của chúng tôi cho thấy mạng VGG-16 được tiền xử lý bằng dữ liệu Imagenet cải thiện 9,5%, điểm số ban đầu là 61,7%. Do đó, việc huấn luyện trước từ tập dữ liệu đa dạng lớn hơn như Imagenet [7] rõ ràng là có lợi vì nó giúp cải thiện độ chính xác tổng quát của mô hình.

Đóng góp của việc tăng cường dữ liệu Bảng 1 trình bày kết quả của việc tăng cường dữ liệu. Chúng ta có thể quan sát thấy rằng rỗng VGG-16 tăng 14,8% trong khi

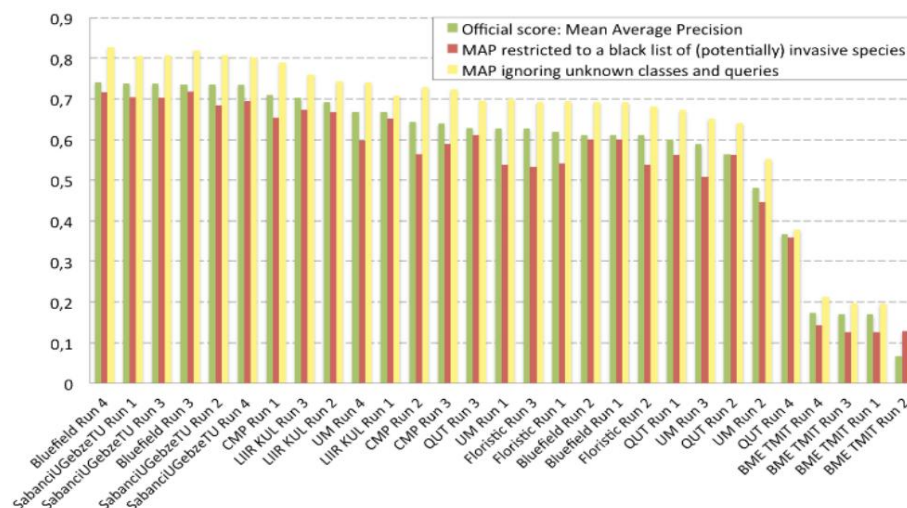
đề xuất nhiệt hạch cấp cao tăng 14,5%. Do đó, có thể suy ra rằng việc tăng cường dữ liệu rất quan trọng đối với nhiệm vụ phân loại thực vật, đặc biệt là khi đào tạo một CNN lớn với lượng dữ liệu hạn chế. Thật vậy, điều này cho phép các mô hình tiếp cận lượng dữ liệu lớn hơn với độ đa dạng cao hơn. Tuy nhiên, mức cao được đề xuất phương pháp tổng hợp cấp độ có độ chính xác phân loại thấp hơn so với phương pháp tinh chỉnh Lưới VGG-16. Điều này có thể là do mạng VGG chỉ sử dụng các đặc điểm của loài, hoạt động tốt hơn sự kết hợp của các tính năng, đủ mạnh để đại diện cho nhà máy hình ảnh.

Bảng 1: So sánh hiệu suất sử dụng tập dữ liệu tăng cường và không tăng cường

Phương pháp	Không tăng cường (%)	Tăng cường (%)
Lớp trên cùng VGG-16 được tinh chỉnh	56,4	71,2
Phản ứng tổng hợp cấp cao (đề xuất)	54,4	68,9

#### 4.3 Kết quả thực nghiệm trên tập thử nghiệm

Chúng tôi đã gửi bốn lần chạy bằng cách sử dụng LifeClef 2016 - bộ dữ liệu thực vật đa cơ quan. Đặc điểm của từng lần chạy được nêu như sau:



Hình 4: Kết quả nhiệm vụ phân loại thực vật đa cơ quan LifeClef2016

- UM RUN 1: pretrain VGG-16 với Imagenet 2012, sau đó tinh chỉnh phần trên hầu hết các lớp với tập dữ liệu đào tạo PlantClef2015 tăng cường.
- UM Run 2: huấn luyện lưới VGG-16 từ đầu với PlantClef2015 tăng cường tập dữ liệu huấn luyện.

- UM Run 3: huấn luyện kiến trúc tổng hợp cấp cao được đề xuất với tập dữ liệu huấn luyện PlantClef2015 tăng cường -
- UM Run 4: tinh chỉnh kiến trúc UM Run 1 với bộ xác thực.

Hình 4 cho thấy kết quả của điểm chính xác trung bình với việc đánh giá độ tin cậy của hệ thống xử lý các danh mục không nhìn thấy được. Chúng tôi nhận thấy rằng Giải chạy 4 là giải chạy tốt nhất trong số các Giải chạy đã gửi. Lý do là vì nó được huấn luyện với tập dữ liệu huấn luyện lớn hơn so với các lần chạy khác. Ngoài ra, Run 1 tốt hơn Run 3, phù hợp với kết quả thu được trong bộ xác thực. Cuối cùng nhưng không kém phần quan trọng, Vòng 2 hiển thị kết quả thấp nhất trong số các lần chạy đã gửi. Điều này một lần nữa cho thấy tầm quan trọng của mô hình đào tạo trước với bộ dữ liệu quy mô lớn. Nhìn chung, tất cả các lượt chạy được gửi đều đạt được kết quả trung bình trong nhiệm vụ phân loại thực vật đa cơ quan LifeClef2016.

## 5 Kết luận và công việc trong tương lai

Bài báo này đề xuất sử dụng mô hình CNN để kết hợp các đặc điểm loài và cơ quan cho nhiệm vụ phân loại thực vật. Chúng tôi đã mô tả phương pháp luận về kiến trúc của mình và phân tích kết quả dựa trên cả tập hợp xác thực và thử nghiệm. Kết quả của kiến trúc nhiệt hạch cấp cao được đề xuất rất hứa hẹn nhưng vẫn còn hạn chế so với mạng VGG. Trong tương lai, chúng tôi sẽ tập trung khám phá mô hình CNN cho nhiệm vụ phân loại thực vật, bao gồm cả ý nghĩa của nó trong nhiệm vụ nhận dạng tập mở.

## Người giới thiệu

1. Champ, J., Lorieul, T., Servajean, M., Joly, A.: Một nghiên cứu so sánh các phương pháp phân loại chi tiết trong bối cảnh thách thức nhận dạng thực vật lifeclef năm 2015. Trong: CLEF 2015. tập. 1391 (2015)
2. Charters, J., Wang, Z., Chi, Z., Tsoi, AC, Feng, DD: Eagle: Một mô tả mới để xác định các loài thực vật sử dụng các đặc điểm mạch máu của lá. Trong: Hội thảo đa phương tiện và hội chợ triển lãm, Hội nghị quốc tế IEEE 2014 về. trang 1-6. IEEE (2014)
3. Choi, S.: Nhận dạng thực vật bằng mạng lưới thần kinh tích chập sâu: Snumedinfo tại nhiệm vụ nhận dạng thực vật lifeclef 2015. Trong: Ghi chú làm việc của hội nghị CLEF 2015 (2015)
4. Clarke, J., Barman, S., Remagnino, P., Bailey, K., Kirkup, D., Mayo, S., Wilkin, P.: Phân tích kiểu gân lá của hình ảnh lá. Trong: Những tiến bộ trong máy tính trực quan, trang 427-436. Mùa xuân (2006)
5. Cope, JS, Corney, D., Clark, JY, Remagnino, P., Wilkin, P.: Nhận dạng loài thực vật bằng hình thái học kỹ thuật số: Đánh giá. Hệ thống chuyên gia có ứng dụng 39(8), 7562-7573 (2012)
6. Cope, JS, Remagnino, P., Barman, S., Wilkin, P.: Việc trích xuất gân lá từ hình ảnh lá bằng các bộ phân loại tính mạch tiến hóa và thuật toán đàn kiến. Trong: Các khái niệm nâng cao cho hệ thống thị giác thông minh. trang 135-144. Mùa xuân (2010)
7. Deng, J., Dong, W., Socher, R., Li, LJ, Li, K., Fei-Fei, L.: Imagenet: Cơ sở dữ liệu hình ảnh phân cấp quy mô lớn. Trong: Hội nghị IEEE về Tầm nhìn Máy tính và Nhận dạng Mẫu, 2009. trang 248-255 (2009)



8. Ge, Z., Mccool, C., Corke, P.: Học tập tính năng cụ thể theo nội dung để phân loại thực vật hạt mìn. Trong: Ghi chú làm việc của hội nghị CLEF 2015 (2015)
9. Go`eau, H., Bonnet, P., Barbe, J., Bakic, V., Joly, A., Molino, JF, Barthelemy, D., Boujemaa, N.: Nhận dạng thực vật đa cơ quan. Trong: Proc. của hội thảo quốc tế ACM lần thứ nhất về phân tích đa phương tiện cho dữ liệu sinh thái. trang 41-44 (2012)
10. Go`eau, H., Bonnet, P., Joly, A.: Nhận dạng thực vật trong thế giới mở (lifeclef 2016). Trong: Ghi chú làm việc của CLEF 2016 (2016)
11. Go`eau, H., Bonnet, P., Joly, A., Bakic, V., Barth'el'emy, D., Boujemaa, N., Molino, JF: Nhiệm vụ xác định thực vật imageclef 2013. Trong: CLEF (2013)
12. Hall, D., McCool, C., Dayoub, F., Sunderhauf, N., Upcroft, B.: Đánh giá các đặc điểm phân loại lá trong điều kiện khó khăn. Trong: Các ứng dụng của Thị giác Máy tính (WACV), Hội nghị Mùa đông IEEE 2015 về. trang 797-804 (2015)
13. Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadar-rama, S., Darrell, T.: Caffe: Kiến trúc tích chập cho tốc độ nhanh nhúng tính năng. Trong: Proc. của Hội nghị quốc tế ACM về đa phương tiện. trang 675-678 (2014)
14. Joly, A., Go`eau, H., Glotin, H., Spampinato, C., Bonnet, P., Vellinga, WP, Champ, J., Planqu'e, R., Palazzo, S., M`uller, H.: Lifeclef 2016: thách thức nhận dạng loài sống đa phương tiện
15. Joly, A., Go`eau, H., Glotin, H., Spampinato, C., Bonnet, P., Vellinga, WP, Planqu'e, R., Rauber, A., Palazzo, S., Fisher, B., và cộng sự: Lifeclef 2015: thách thức nhận dạng loài sống đa phương tiện. Trong: IR thử nghiệm đáp ứng đa ngôn ngữ, đa phương thức và tương tác, trang 462-483. Mùa xuân (2015)
16. Kadir, A., Nugroho, LE, Susanto, A., Santosa, PI: Phân loại lá bằng cách sử dụng các đặc điểm hình dạng, màu sắc và kết cấu. Tạp chí Quốc tế về Xu hướng và Công nghệ Máy tính 1(3), 306-311 (Tháng 7 đến tháng 8 năm 2011)
17. Kalyoncu, C., Toygar, O.: Phân loại lá hình học. Thị giác máy tính và hiểu biết về hình ảnh 133, 102-109 (2015)
18. Krizhevsky, A., Sutskever, I., Hinton, GE: 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. Kumar, N., Belhumeur, PN, Biswas, A., Jacobs, DW, Kress, WJ, Lopez, IC, Soares, JV: Leafsnap: Một hệ thống thị giác máy tính để nhận dạng các loài thực vật tự động. Trong: Computer Vision-ECCV 2012, trang 502-516. Mùa xuân (2012)
20. Lee, SH, Chan, CS, Wilkin, P., Remagnino, P.: Deep-plant: Nhận dạng thực vật bằng mạng lưới thần kinh tích chập. Trong: Hội nghị quốc tế về xử lý hình ảnh của IEEE 2015. trang 452-456 (2015)
21. Reyes, AK, Caicedo, JC, Camargo, JE: Tinh chỉnh các mạng chập sâu để nhận dạng thực vật. Trong: Ghi chú làm việc của hội nghị CLEF 2015 (2015)
22. Simonyan, K., Zisserman, A.: Mạng tích chập rất sâu cho quy mô lớn nhận dạng hình ảnh. CoRR, abs/1409.1556 (2014)
23. Yan, Z., Zhang, H., Piramuthu, R., Jagadeesh, V., DeCoste, D., Di, W., Yu, Y.: Hd-cnn: Mạng lưới thần kinh tích chập sâu phân cấp cho hình ảnh quy mô lớn sự công nhận. Trong: Proc. của Hội nghị quốc tế IEEE về thị giác máy tính. trang 2740-2748 (2015)
24. Yanikoglu, B., Tolga, Y., Tirkaz, C., FuenCaglar, E.: Hệ thống Sabanci-okan tại cuộc thi nhận dạng thực vật lifeclef 2014. Trong: Ghi chú làm việc của hội nghị CLEF 2014 (2014)
25. Zeiler, MD, Fergus, R.: Trực quan hóa và hiểu các mạng tích chập. Trong: Thị giác máy tính-ECCV 2014, trang 818-833. Mùa Xuân (2014)