

PLANT IDENTIFICATION BASED ON A LATE FUSION METHOD WITH PRIORITY WEIGHTS

Nguyen Thi Thanh Nhan*, Nguyen Thu Huong

TNU - University of Information and Communication Technology

ARTICLE INFO	ABSTRACT
Received: 29/01/2021 Revised: 28/02/2021 Published: 28/02/2021	Automatic plant identification based on images is currently very interesting. The challenge with the plant identification problem is the great similarity among species, especially when the number of species is large. In this study, we developed a plant identification method based on the use of late fusion method for identification results on different plant organs. We assigned priority weights by organ/species to the confidence scores of each model. The organ/species with a better identification result was assigned a higher weight. GoogLeNet was used to identify plant based on each organ. Experiments were applied to combine two to six organs according to leaf, flower, fruit, stem, branch, entire. This method is based on combining the product rule using weights assigned to plant organs and species. The experimental results have shown the effectiveness of the proposed method, it outperforms some fusion late methods. The proposed method achieved the highest accuracy when combining 2 organs, 3 organs, 4 organs, 5 organs, and 6 organs with 96.0%, 98.2%, 98.8%, 99%, and 99.2% respectively.
KEYWORDS Late fusion Product rule Plant identification Convolutional neural network Priority weights	

NHẬN DẠNG CÂY DỰA TRÊN PHƯƠNG PHÁP KẾT HỢP MUỘN CÓ GÁN TRỌNG SỐ

Nguyễn Thị Thanh Nhân*, Nguyễn Thu Hương

Trường Đại học Công nghệ Thông tin và Truyền thông – ĐH Thái Nguyên

THÔNG TIN BÀI BÁO	TÓM TẮT
Ngày nhận bài: 29/01/2021 Ngày hoàn thiện: 28/02/2021 Ngày đăng: 28/02/2021	Bài toán nhận dạng cây tự động dựa trên hình ảnh hiện nay đang rất được quan tâm, mục đích hỗ trợ người dùng có khả năng nhận diện được các loài cây để sử dụng hợp lý và bảo tồn sự đa dạng của thực vật. Thách thức với bài toán nhận dạng cây là sự tương tự lớn giữa các loài, đặc biệt khi nhận dạng số lượng loài lớn. Trong nghiên cứu này, chúng tôi phát triển một phương pháp nhận dạng cây dựa trên việc sử dụng phương pháp kết hợp muộn các kết quả nhận dạng cây trên các bộ phận khác nhau của cây. Chúng tôi gán trọng số ưu tiên theo bộ phận/ theo loài cho các độ tin cậy trả về theo mỗi mô hình. Bộ phận/loài nào có kết quả nhận dạng tốt hơn sẽ được gán trọng số cao hơn. Mạng học sâu GoogLeNet được sử dụng để nhận dạng cây dựa trên từng bộ phận. Các thực nghiệm được áp dụng cho kết hợp từ hai cho đến sáu bộ phận của cây theo các bộ phận lá, hoa, quả, thân, cành, toàn bộ cây. Các kết quả thực nghiệm đã chỉ ra hiệu quả của phương pháp đề xuất, vượt trội hơn so với một số phương pháp khác. Phương pháp đề xuất đạt kết quả tốt nhất khi kết hợp hai, ba, bốn, năm và sáu bộ phận đạt được lần lượt là 96,0%, 98,2%, 98,8%, 99,0% và 99,2%.
TỪ KHÓA Kết hợp muộn Luật nhân Nhận dạng cây Mạng nơ ron tích chập Trọng số ưu tiên	

* Corresponding author. Email: nttanhan@ictu.edu.vn

1. Giới thiệu

Thực vật có vai trò rất quan trọng trong cuộc sống của con người, chúng cung cấp thực phẩm, thuốc men, gỗ, sợi, và cân bằng hệ sinh thái... Để giúp cho việc sử dụng, khai thác hiệu quả, hợp lý, bảo tồn sự đa dạng thực vật, bài toán nhận dạng cây tự động bằng máy tính điện tử hiện nay đang rất được quan tâm, từ đó xây dựng được các ứng dụng nhận dạng cây trên các thiết bị di động giúp người dùng dễ dàng nhận dạng được các loài cây họ quan tâm [1]. Thách thức đặt ra đối với bài toán nhận dạng cây là sự tương tự giữa các loài lớn, đặc biệt các loài trong cùng một họ. Các nghiên cứu trước đây thường áp dụng các kỹ thuật học truyền thống trên một bộ phận chủ yếu là ảnh lá trên nền đơn giản, với những cơ sở dữ liệu nhỏ, số lượng loài nhỏ cũng đã đạt được một số kết quả ấn tượng [1]. Tuy nhiên, hiện nay, các nghiên cứu đang hướng đến xử lý những cơ sở dữ liệu thực vật có số lượng loài lớn, các ảnh có nền phức tạp được chụp trực tiếp trên cây ví dụ như cơ sở dữ liệu cây trong cuộc thi LifeCLEF về nhận dạng cây được cập nhật hàng năm [2] – [4]. Việc áp dụng các phương pháp học truyền thống trong các cơ sở dữ liệu lớn như thế này là còn hạn chế, đạt độ chính xác không cao.

Gần đây với sự phát triển của phần cứng máy tính, đặc biệt là đơn vị xử lý đồ họa (GPUs), sự bùng nổ của các phương pháp học sâu đã đạt được các kết quả rất ấn tượng cho các bài toán nhận dạng nói chung và các bài toán nhận dạng cây nói riêng [3]-[4]. Các nghiên cứu thường chỉ tập trung trên một bộ phận của cây. Tuy nhiên, việc sử dụng một bộ phận cây là không đầy đủ thông tin do sự tương tự lớn giữa các loài và sự khác biệt ngay trong một loài. Ví dụ nhiều loài có bộ phận lá giống nhau nhưng bộ phận hoa hoặc quả lại khác nhau. Để giải quyết vấn đề này, chúng tôi đề xuất phương pháp kết hợp muộn nhiều ảnh thuộc các bộ phận khác nhau của cây để cải thiện độ chính xác.

Mỗi bộ phận lớp riêng biệt sẽ được áp dụng cho mỗi bộ phận, sau đó, chúng tôi áp dụng phương pháp kết hợp muộn để kết hợp các kết quả đầu ra cho mỗi bộ phận lớp. Các phương pháp kết hợp muộn thường sử dụng như luật nhân, luật trung bình, luật lớn nhất, luật tổng, luật nhỏ nhất [5], luật IprMNZ, luật bình chọn theo số đông (majority voting rule), luật tổng có trọng số. Các phương pháp này thường được sử dụng do tính đơn giản và hiệu quả. Nghiên cứu [6] đề xuất ra phương pháp kết hợp lai bền vững (viết tắt RHF) thực hiện kết hợp dựa trên luật nhân và phương pháp dựa trên phân lớp đã đạt kết quả rất tốt. Nghiên cứu [7] thực hiện nhận dạng dựa trên việc kết hợp hai bộ phận của cây sử dụng phương pháp kết hợp sớm. Nghiên cứu [6] đã thực hiện nhận dạng cây kết hợp từ hai đến sáu bộ phận cây.

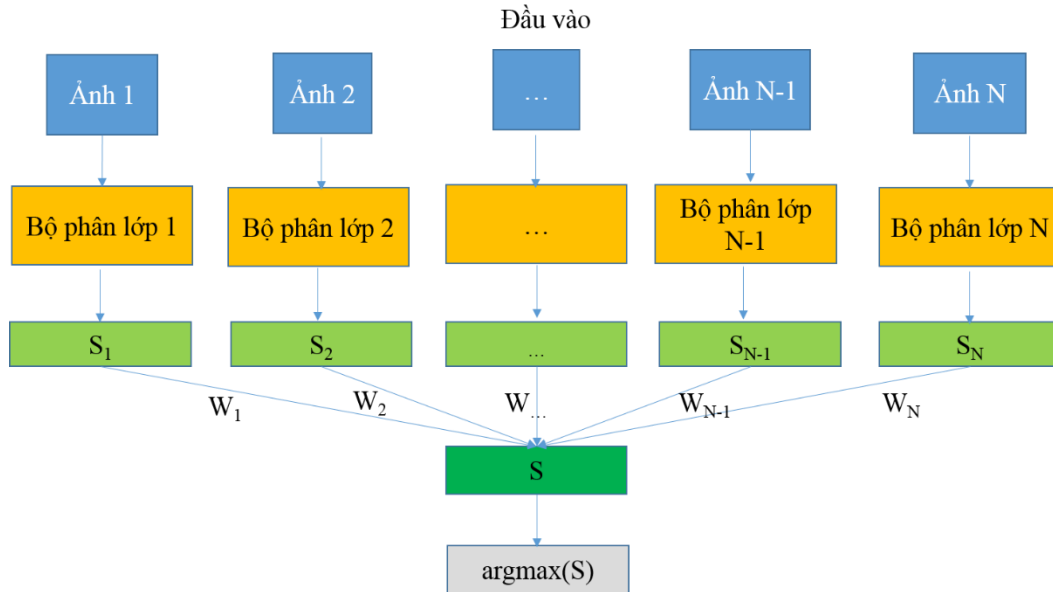
Với mục đích tiếp tục cải thiện khả năng kết hợp so với các phương pháp đã có, bài báo đề xuất một phương pháp kết hợp mới bằng việc sử dụng luật nhân kết hợp các độ tin cậy trả về có gán trọng số ưu tiên theo mỗi bộ phận/loài. Dựa trên độ chính xác nhận dạng cho mỗi bộ phận/loài, kết quả nhận dạng trên bộ phận/loài nào tốt hơn thì được gán trọng số ưu tiên cao hơn. Bài báo thực hiện kết hợp cho sáu bộ phận của cây trên cơ sở dữ liệu 50 loài cây được trích ra từ cơ sở dữ liệu LifeCLEF 2015 [2]. Phương pháp học sâu được áp dụng cho nhận dạng cây dựa trên mỗi bộ phận để nâng cao kết quả nhận dạng cây. Các kết quả thực nghiệm đã chỉ ra hiệu quả của phương pháp đề xuất.

2. Phương pháp kết hợp nhiều bộ phận cây

Dựa trên các phương pháp kết hợp muộn cơ bản được thực hiện trên các bài toán khác nhau và được truyền cảm hứng từ nghiên cứu [8] thực hiện kết hợp nhiều mô hình khác nhau là kết quả của việc áp dụng nhiều mạng nơ ron học sâu khác nhau, có gán trọng số theo các mô hình và theo loài cho bài toán nhận dạng cây, trong nghiên cứu này, chúng tôi đề xuất áp dụng trên các bộ phận khác nhau của cây, sau đó thực hiện việc gán trọng số theo bộ phận và theo loài.

Mỗi bộ phận cây sẽ được huấn luyện nhận dạng riêng biệt để sinh ra các mô hình nhận dạng. Mỗi mô hình nhận dạng này sẽ học được các đặc trưng riêng biệt cho mỗi bộ phận, do vậy, việc kết hợp các mô hình này sẽ cho ra kết quả nhận dạng tốt hơn. Dựa trên ý tưởng này, nghiên cứu

tiếp cận nhận dạng cây với nhiều mô hình nhận dạng dựa trên nhiều bộ phận cây, sau đó áp dụng phương pháp kết hợp muộn các thông tin đầu ra của các bộ phận lớp có sử dụng trọng số có ưu tiên để phân lớp ảnh các loài cây. Trong các bộ phận cây, có những bộ phận có khả năng phân biệt loài cao hơn bộ phận khác, thì bộ phận đó cần được gán trọng số cao hơn, tương tự như vậy giữa các loài cây, loài cây nào có khả năng nhận dạng tốt hơn thì được gán trọng số cao hơn. Trong bài báo này, chúng tôi áp dụng luật nhân [9] và phương pháp xác định trọng số ưu tiên được gán cho mỗi mô hình nhận dạng bộ phận cây hoặc theo loài cây. Véc tơ kết quả trả về được tạo bởi phép nhân các độ tin cậy trả về thực hiện lấy lũy thừa theo trọng số được gán.



Hình 1. Sơ đồ phương pháp đề xuất

Phương pháp đề xuất được trình bày trong hình 1. Câu truy vấn đầu vào gồm N ảnh của N bộ phận quan tâm. Trong các thực nghiệm, N biến đổi từ 2 đến 6. Đầu tiên, mỗi bộ phận cây tương ứng được huấn luyện riêng biệt cho nhận dạng cây dựa trên đơn bộ phận sẽ thu được bộ phận lớp tương ứng. Với mỗi ảnh đầu vào i , khi đi qua bộ phận lớp tương ứng thứ i sẽ trả về một véc tơ độ tin cậy tương ứng với các loài dự đoán trả về S_i , véc tơ này có độ dài bằng số loài C trong cơ sở dữ liệu. Phương pháp kết hợp sẽ lấy các véc tơ độ tin cậy này như là đầu vào, sau đó thực hiện gán các trọng số cho các giá trị trong véc tơ độ tin cậy theo bộ phận và theo loài, tiếp theo thực hiện kết hợp các độ tin cậy này để trả về một vector độ tin cậy mới S . Loài tương ứng với giá trị tin cậy trả về lớn nhất ($\text{argmax}(S)$) sẽ là loài dự đoán trả về.

Một số ký hiệu được sử dụng như sau:

- $q = \{I_1, I_2, \dots, I_N\}$ là câu truy vấn chứa N ảnh của N bộ phận;
- C : là số lớp của cơ sở dữ liệu;
- $s_i(I_k)$ là độ tin cậy của loài i khi sử dụng ảnh I_k từ bộ nhận dạng đơn bộ phận tương ứng, ở đó $1 \leq i \leq C, 1 \leq k \leq N$;
- c : lớp dự đoán của loài cho câu truy vấn q

Đề xuất 1: Kết hợp muộn sử dụng trọng số gán theo bộ phận cây

Sau khi huấn luyện các mô hình nhận dạng cây theo từng bộ phận cây, một vài mô hình có thể tốt hơn các mô hình khác. Do đó, chúng tôi gán trọng số độ ưu tiên của bộ phận k là độ chính xác trả về theo bộ phận đó, ký hiệu W_k . Để tối ưu các kết quả của các mô hình trên các bộ phận khác nhau khi thực hiện kết hợp, giả sử rằng tập các trọng số của N mô hình trên tập kiểm thử là A , $A = (W_1, W_2, \dots, W_N)$. Trọng số ưu tiên được gán theo phương trình (1):

$$WO_k = \frac{W_k}{\max(A)} \quad (1)$$

Phương pháp đề xuất thứ 1 thực hiện theo phương trình (2), các giá trị tin cậy $s_i(I_k)$ được thực hiện lũy thừa theo trọng số gán cho bộ phận thứ k . Sau đó, áp dụng luật nhân để kết hợp các giá trị thu được, kết quả sẽ trả về chỉ số của loài có giá trị trả về là lớn nhất.

$$c = \operatorname{argmax}_{i=1..C} (\prod_{k=1}^N s_i^{WO_k}(I_k)) \quad (2)$$

Đề xuất 2: Kết hợp muộn sử dụng trọng số gán theo bộ phận và trọng số gán theo loài.

Mặc dù hiệu năng tổng thể của mô hình bộ phận k có thể tốt hơn mô hình bộ phận l , nhưng điều đó không chắc rằng mô hình k tiến hành tốt hơn trên tất cả các loài so với mô hình l bởi vì kiến trúc của mô hình l có thể thích hợp cho một số đặc trưng thu được trên một số loài so với mô hình k . Do vậy, trong đề xuất thứ 2, chúng tôi thực hiện gán thêm trọng số theo loài, với mỗi loài i trong mô hình k gán một trọng số ưu tiên WS_i^k , đây chính là độ chính xác nhận dạng đúng loài i trong mô hình k trên tập kiểm thử. Phương pháp đề xuất 2 được thực hiện theo phương trình (3) trong đó mỗi giá trị độ tin cậy $s_i(I_k)$ sẽ được lấy lần lượt lũy thừa tương ứng theo trọng số gán cho bộ phận k như ở đề xuất 1 và trọng số gán theo loài i trong mô hình thứ k . Sau đó, áp dụng luật nhân với tích của các độ tin cậy tương ứng. Đề xuất này tận dụng được cả những ưu điểm của việc sử dụng trọng số theo bộ phận và theo loài:

$$c = \operatorname{argmax}_{i=1..C} (\prod_{k=1}^N s_i^{WO_k}(I_k) \times s_i^{WS_i^k}(I_k)) \quad (3)$$

Để so sánh hiệu quả của phương pháp đề xuất chúng tôi thực hiện so sánh với các phương pháp kết hợp lai bền vững RHF [6]– phương pháp này đã đạt kết quả vượt trội so với các phương pháp kết hợp cơ bản là phương pháp dựa trên luật lớn nhất, luật tổng và luật nhân.

Với mỗi bộ phận, chúng tôi áp dụng mạng GoogLeNet để huấn luyện. Đây là mạng học sâu được sử dụng nhiều trong bài toán nhận dạng cây và cho các kết quả rất tích cực. GoogLeNet được đề xuất bởi Szegedy và các cộng sự đã chiến thắng trong cuộc thi nhận dạng các đối tượng trên cơ sở dữ liệu hình ảnh lớn, đa dạng ImageNet năm 2014 [10]. GoogLeNet có kiến trúc sâu hơn và rộng hơn so với nhiều kiến trúc mạng khác như AlexNet, VGGNet. Mạng này sử dụng kiến trúc mô đun lặp (inception module), mô đun này lấy các đặc trưng đầu vào và thực hiện nhân chập với các mẫu 1×1 , 3×3 , 5×5 và lớp giảm chiều cực đại (max pooling), sau đó, thực hiện kết nối các đặc trưng đầu ra để trở thành mô đun đầu vào cho lớp tiếp theo. Mạng này bao gồm 6,8 triệu tham số, 22 lớp với 9 mô đun lặp, 2 lớp nhân chập, 2 lớp chuẩn hóa, 5 lớp giảm chiều, một lớp kết nối đầy đủ, một lớp tuyến tính với hàm kích hoạt Softmax như là một bộ phân lớp.

3. Kết quả thực nghiệm

Bảng 1. Cơ sở dữ liệu 50 loài với 6 bộ phận

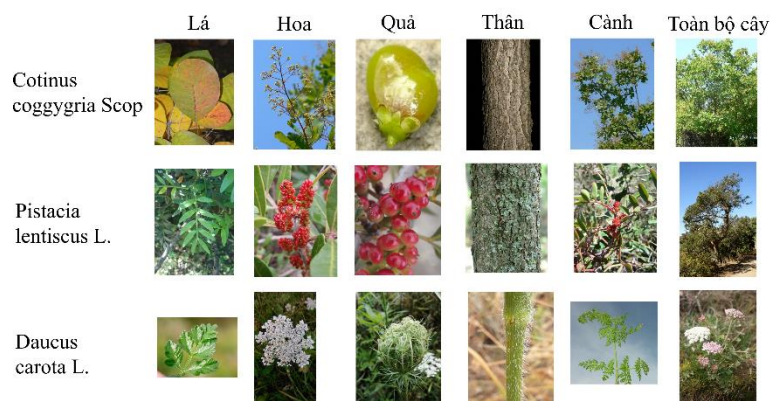
	Tập huấn luyện	Tập kiểm thử	Tổng
Lá	1930	776	2706
Hoa	1650	673	2323
Cành	1388	553	1941
Toàn bộ cây	825	341	1166
Quả	3821	500	4321
Thân	2912	500	3412

Trong bài báo này, chúng tôi sử dụng độ chính xác để đo hiệu năng của hệ thống.

Độ chính xác = số mẫu nhận dạng đúng/Tổng số mẫu nhận dạng.

Chúng tôi thực hiện trên tập con cơ sở dữ liệu LifeCLEF 2015 [2] với 50 loài cây phổ biến bao gồm đầy đủ 6 bộ phận của cây. Với mục đích gia tăng dữ liệu cho việc huấn luyện khi sử dụng mạng học sâu, chúng tôi thực hiện thu thập thêm các ảnh từ internet để bổ sung vào cơ sở dữ liệu này. Cơ sở dữ liệu gồm các ảnh cây thuộc vào sáu bộ phận của cây, chúng tôi ký hiệu các bộ phận tương ứng ở trong ngoặc đơn như sau: lá (Le), hoa (Fl), toàn bộ cây (En), cành (Br), quả

(Fr) và thân (St). Bảng 1 cung cấp số liệu số lượng ảnh cho mỗi bộ phận cây. Hình 2 biểu diễn hình ảnh thuộc 6 bộ phận của 3 loài khác nhau trong cơ sở dữ liệu, các ảnh trong cơ sở dữ liệu đều có nền phức tạp.

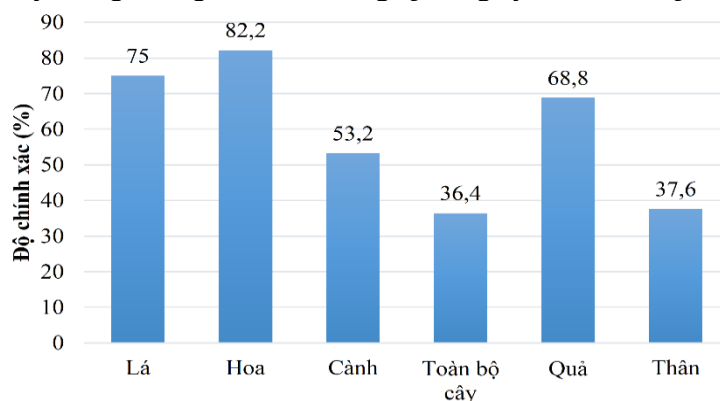


Hình 2. Ảnh các bộ phận của một số loài cây trong cơ sở dữ liệu

3.1. Nhận dạng cây dựa trên một bộ phận

Các kết quả thực nghiệm được tiến hành trên máy chủ với cấu hình: 2,20 GHz CPU, 16 GB RAM, GeForce GTX 1080 Ti GPU, thực hiện trên Caffé. Chúng tôi áp dụng kỹ thuật học chuyển đổi (transfer learning), sử dụng bộ trọng số tiền huấn luyện trên cơ sở dữ liệu rất lớn và đa dạng ImageNet làm bộ trọng số khởi tạo cho quá trình huấn luyện. Kỹ thuật này đã được chỉ ra là hiệu quả hơn so với việc huấn luyện từ đầu với bộ trọng số khởi tạo ngẫu nhiên. Các ảnh đầu vào có kích thước 224x224. Dựa trên việc nghiên cứu các kết quả liên quan, cấu hình máy, và các thực nghiệm, chúng tôi lựa chọn các tham số cho mạng như sau: learning_rate: 0,001, batch_size: 32, weight_decay: 0,0002, dropout: 0,4, số bước lặp là 50.000.

Các kết quả đạt được cho nhận dạng dựa trên đơn bộ phận khi áp dụng mạng GoogLeNet cho kết quả từ cao xuống thấp như sau: hoa (82,2%), lá (75%), quả (68,8%), cành (53,2%), thân (37,6%), toàn bộ cây (36,4%) được chỉ ra trong hình 3. Hoa là bộ phận cho kết quả nhận dạng cao nhất là 82,2% vì bộ phận hoa có nhiều đặc điểm bề ngoài (màu sắc, hình dạng) có tính chất phân biệt cao giữa các loài. Các giá trị này được sử dụng làm trọng số W_k sử dụng cho phương pháp đề xuất 1, đề xuất 2. Giá trị trung bình trên mỗi loài cho mỗi mô hình nhận dạng trên mỗi bộ phận được sử dụng để làm trọng số gán cho mỗi loài sử dụng cho đề xuất 2. Nhận dạng cây dựa trên một bộ phận là trường hợp đặc biệt của nhận dạng cây dựa trên nhiều bộ phận, ở đây số bộ phận $N = 1$, điều này tương đương với việc không áp dụng kỹ thuật kết hợp mụten.



Hình 3. Kết quả nhận dạng cây trên từng bộ phận

3.2. Nhận dạng cây dựa trên kết hợp nhiều bộ phận của cây

Chúng tôi tiến hành thực hiện cho hai phương pháp đề xuất. Các kết quả thực nghiệm được tiến hành kết hợp cho các cặp từ 2 đến 6 bộ phận của cây cho việc nhận dạng cây dựa trên nhiều ảnh bộ phận của cây. Các kết quả được chỉ ra ở trong Bảng 2 với phần in đậm là kết quả đạt được tốt nhất theo hàng, có tổng tất cả 57 trường hợp kết hợp. Các kết quả của phương pháp RHF chúng tôi lấy từ nghiên cứu [6]. Từ bảng kết quả này chúng tôi có một số nhận xét sau:

Đề xuất 1 vượt trội hơn phương pháp RHF với 42/57 trường hợp với độ vượt trội lớn nhất là 3%, ví dụ kết hợp bộ phận hoa và thân (Fl+St) ở phương pháp RHF đạt độ chính xác là 87,4% trong khi phương pháp đề xuất 1 đạt kết quả 90,4%.

Đề xuất 2 vượt trội hơn phương pháp RHF với 56/57 trường hợp với độ vượt trội lớn nhất là 4,4%, ví dụ cho trường hợp này là kết hợp lá, cành và thân (Le+Br+St) ở phương pháp RHF đạt độ chính xác là 88,2% còn phương pháp đề xuất 2 đạt độ chính xác 92,6%. Bài báo [6] chỉ ra RHF là tốt hơn so với phương pháp luật nhân, luật lớn nhất, luật tổng, do vậy suy ra phương pháp đề xuất 1, đề xuất 2 cũng tốt hơn các phương pháp kết hợp cơ bản này.

Đề xuất 2 lại vượt trội hơn đề xuất 1 với 54/57 trường hợp với độ vượt trội lớn nhất là 4,4% cho trường hợp kết hợp Le+Br+St. Trong ba phương pháp thực hiện so sánh thì phương pháp đề xuất 2 là cho kết quả tốt nhất.

Bảng 2. Độ chính xác tại hạng 1 (tính theo %) khi kết hợp các bộ phận khác nhau

Các bộ phận kết hợp	RHF	Đề xuất 1	Đề xuất 2	Các bộ phận kết hợp	RHF	Đề xuất 1	Đề xuất 2
Le+Fl	95,8	95,6	96,0	Fl+En+St	89,6	91,0	92,0
Le+Br	84,6	85,2	86,2	Fl+Fr+St	94,0	95,0	96,0
Le+En	80,6	79,4	80,6	Br+En+Fr	88,2	89,4	89,2
Le+Fr	88,0	87,8	89,4	Br+En+St	76,2	74,4	77,8
Le+St	80,6	82,8	84,2	En+Fr+St	84,4	85,8	86,2
Fl+Br	91,6	91,8	93,4	Br+Fr+St	87,2	86,8	89,2
Fl+En	84,2	86,0	88,2	Le+Fl+Br+En	97,6	97,4	98,6
Fl+Fr	94,6	94,6	95,8	Le+Fl+Br+Fr	98,4	98,2	98,8
Fl+St	87,4	90,4	89,8	Le+Fl+Br+St	97,6	98,0	98,4
Br+En	64,2	63,8	66,4	Le+Fl+En+Fr	98,4	98,4	98,4
Br+Fr	82,2	83,4	84,8	Le+Fl+En+St	97,0	97,4	98,0
Br+St	67,6	66,6	68,4	Le+Fl+Fr+St	98,0	98,0	98,4
En+Fr	79,6	79,8	78,0	Le+Br+En+Fr	95,4	96,2	97,6
En+St	54,0	54,4	55,6	Le+Br+En+St	91,0	91,8	93,4
Fr+St	74,6	77,4	77,6	Le+Br+Fr+St	94,8	95,4	96,8
Le+Fl+Br	96,8	96,6	97,8	Le+En+Fr+St	93,2	94,4	96,4
Le+Fl+En	96,8	96,6	97,4	Fl+Br+En+Fr	96,8	97,6	98,0
Le+Fl+Fr	98,0	98,0	98,2	Fl+Br+En+St	93,8	95,2	95,6
Le+Fl+St	96,0	96,0	96,4	Fl+Br+Fr+St	95,8	96,8	97,8
Le+Br+En	88,8	89	89,8	Br+En+Fr+St	90,2	91,8	92,2
Le+Br+Fr	94,4	93,6	95,0	Fl+En+Fr+St	95,2	97,0	97,2
Le+Br+St	88,2	88,2	92,6	Le+Fl+Br+En+Fr	98,6	98,2	98,6
Le+En+Fr	93,0	93	94,0	Le+Fl+Br+En+St	97,6	98,0	98,8
Le+En+St	85,2	86,8	89,2	Le+Br+En+Fr+St	96,4	96,4	98,2
Le+Fr+St	91,2	91,4	92,6	Fl+Br+En+Fr+St	97,4	98,2	98,4
Fl+Br+En	92,4	93,6	94,6	Le+Fl+Br+Fr+St	98,0	98,6	99,0
Fl+Br+Fr	96,4	96,8	97,6	Le+Fl+En+Fr+St	98,2	98,4	99,0
Fl+Br+St	91,8	93,8	94,8	Le+Fl+Br+En+Fr+St	98,8	99,2	99,2
Fl+En+Fr	95,2	95,6	96,6				

Kết quả bảng 2 và hình 3 cho thấy việc kết hợp các bộ phận cho kết quả nhận dạng tốt hơn nhận dạng đơn bộ phận và khi càng kết hợp nhiều bộ phận thì kết quả nhận dạng càng cao. Điều này đúng cho cả ba phương pháp trong bảng 2.

Bảng 3 chỉ ra với mỗi giá trị tổ hợp N bộ phận, kết quả kết hợp đạt được tốt nhất tương ứng, trong bảng 3 này kết quả tốt nhất đạt được cho phương pháp đề xuất 2. Phân tích chi tiết cho các kết quả của phương pháp đề xuất 2 như sau. Đối với việc kết hợp 2 bộ phận, kết quả đạt tốt nhất khi kết hợp hai bộ phận lá và hoa (Le+Fl) đạt kết quả là 96%, trong khi nhận dạng chỉ sử dụng các ảnh hoa đạt kết quả là 82,2% và nhận dạng chỉ sử dụng ảnh lá đạt kết quả là 75,0% tại hạng 1 cho dữ liệu kiểm thử. Kết quả nhận dạng tốt nhất cho 3 bộ phận đạt được là 98,2% cho kết hợp lá, hoa, quả (Le+Fl+Fr), 4 bộ phận là 98,8% (Le+Fl+Br+Fr), 5 bộ phận là 99% (Le+Fl+Br+Fr+St và Le+Fl+En+Fr+St) và cuối cùng kết quả trên sáu bộ phận là 99,2%.

Bảng 3. Các kết hợp các bộ phận đạt kết quả tốt nhất tương ứng đối với N bộ phận

N	RHF	Đề xuất 1	Đề xuất 2
2	Le+Fl: 95,8	Le+Fl: 95,6	Le+Fl: 96,0
3	Le+Fl+Fr: 98,0	Le+Fl+Fr: 98,0	Le+Fl+Fr: 98,2
4	Le+Fl+Br+Fr: 98,4 Le+Fl+En+Fr: 98,4	Le+Fl+En+Fr: 98,4	Le+Fl+Br+Fr: 98,8
5	Le+Fl+Br+En+Fr: 98,6	Le+Fl+Br+Fr+St: 98,6	Le+Fl+Br+Fr+St: 99,0 Le+Fl+En+Fr+St: 99,0
6	Le+Fl+Br+En+Fr+St: 98,8	Le+Fl+Br+En+Fr+St: 99,2	Le+Fl+Br+En+Fr+St: 99,2

Từ kết quả này cho thấy kết quả nhận dạng hai bộ phận tốt nhất vượt trội so với một bộ phận là 14,8%, kết quả nhận dạng 3 bộ phận vượt trội so với hai bộ phận là 2,2%, kết quả nhận dạng 4 bộ phận vượt trội so với kết hợp 3 bộ phận là 0,6%, kết quả nhận dạng 5 bộ phận vượt trội so với 4 bộ phận là 0,2 % và kết quả nhận dạng 6 bộ phận vượt trội hơn so với kết quả nhận dạng 5 bộ phận là 0.2%. Càng kết hợp nhiều bộ phận thì cho kết quả nhận dạng càng cao, trong đó với $N=2$ và $N = 3$ thì độ chính xác gia tăng đáng kể và sau đó sự cải thiện tăng chậm với $N > 3$.

Bảng 4. Thời gian trung bình (tính theo giây) thực hiện để phân lớp 1 câu truy vấn chứa N ảnh

N	Phương pháp đề xuất		RHF	
	GPU	CPU	GPU	CPU
1	0.51	2.54	0.51	2.54
2	0.56	3.51	0.98	3.93
3	0.60	5.06	1.07	5.53
4	0.64	6.54	1.21	7.11
5	0.70	8.01	1.38	8.69
6	0.74	9.52	1.54	10.32

Các kết quả kết hợp giữa các bộ phận thường được cải thiện nhiều khi kết hợp với bộ phận có kết quả nhận dạng cao trong nhận dạng đơn bộ phận hay bộ phận được gán trọng số cao, ví dụ như bộ phận hoa, lá. Phương pháp đề xuất 1 và đề xuất 2 là thực hiện đơn giản hơn và lại cho kết quả tốt hơn phương pháp RHF, trong đó phương pháp đề xuất 2 là cho kết quả tốt nhất vì phương pháp này cho phép gán trọng số ưu tiên theo bộ phận và theo loài nên phát huy được các ưu điểm để cho kết quả nhận dạng cao. Cách tiếp cận này là linh hoạt, phù hợp với thực tế bởi vì một cây không phải luôn tồn tại sẵn tất cả các bộ phận của cây tại một thời điểm. Khi kết hợp luôn ưu tiên kết hợp các bộ phận có trọng số cao trước.

Chúng tôi tiến hành đo thời gian chạy của phương pháp đề xuất khi thực hiện trên CPU và GPU. Các ảnh từ tập kiểm thử được sử dụng để đo thời gian trung bình cần thiết để phân lớp một câu truy vấn gồm cặp N ảnh tương ứng cho N bộ phận cây. Bảng 4 cho thấy thời gian thực hiện sẽ tăng tuyến tính khi số lượng ảnh trong câu truy vấn tăng lên. Đối với phương pháp đề xuất, thời gian thực hiện trên CPU cao gấp từ 5 đến 13 lần so với thời gian thực hiện trên GPU. Với kết

quả thực hiện trên GPU thì phương pháp đề xuất của chúng tôi có thể trả kết quả phân loại đáp ứng được yêu cầu thời gian thực. Với kết quả trả về ở bảng 4 cũng chỉ ra rằng phương pháp đề xuất là nhanh hơn so với phương pháp RHF. Các kết quả thực nghiệm cho thấy phương pháp đề xuất mới hiệu quả cả về độ chính xác và thời gian tính toán.

4. Kết luận

Chúng tôi đã đề xuất phương pháp kết muôn các mô hình thu được từ nhận dạng cây dựa trên các bộ phận cây bằng việc áp dụng luật nhân và gán trọng số cho các độ tin cậy trả về theo bộ phận và theo loài. Chúng tôi gán trọng số ưu tiên theo nguyên tắc, kết quả nhận dạng theo bộ phận nào/loài nào cho kết quả cao sẽ được gán trọng số cao hơn. Các kết quả thực nghiệm đã chỉ ra rằng việc sử dụng phương pháp kết hợp muôn cho nhiều bộ phận sẽ cho kết quả tốt hơn nhiều so với nhận dạng dựa trên một bộ phận; tuy nhiên, sự cải thiện sẽ tăng chậm khi kết hợp quá nhiều bộ phận. Phương pháp đề xuất của chúng tôi là hiệu quả hơn phương pháp RHF và các phương pháp kết hợp cơ bản khác. Chúng tôi tin rằng, nghiên cứu này là rất hứa hẹn để xây dựng ứng dụng nhận dạng cây trong môi trường thực.

Lời cảm ơn

Nghiên cứu này được hỗ trợ từ đề tài nghiên cứu khoa học cấp cơ sở với mã số T2020-07-13.

TÀI LIỆU THAM KHẢO/ REFERENCES

- [1] W. Jana and P. Mäder, "Plant species identification using computer vision techniques: A systematic literature review," *Archives of Computational Methods in Engineering*, vol. 25, no. 2, pp. 507-543, 2018.
- [2] H. Goëau, P. Bonnet, and A. Joly, "Lifeclef plant identification task 2015," in CEUR-WS (Ed.), *CLEF: Conference and Labs of the Evaluation forum*, vol. 1391 of CLEF2015 Working notes, Toulouse, France, 2015.
- [3] H. Goëau, P. Bonnet, and A. Joly, "Plant identification in an open-world (lifeclef 2016)," *CLEF working notes 2016*, pp. 428-439.
- [4] H. Goëau, P. Bonnet, and A. Joly, "Plant identification based on noisy web data: the amazing performance of deep learning (lifeclef 2017)," *CEUR Workshop Proceedings*, 2017.
- [5] J. Kittler, M. Hatef, R. P. Duin, and J. Matas, "On combining classifiers," *IEEE transactions on pattern analysis and machine intelligence*, vol. 20, no. 3, pp. 226-239, 1998.
- [6] T. T. N. Nguyen, T. L. Le, and H. Vu, "Do we need multiple organs for plant identification?" 2020 International Conference on Multimedia Analysis and Pattern Recognition (MAPR), IEEE, 2020, pp. 1-6.
- [7] A. He, and X. Tian, "Multi-organ plant identification with multi-column deep convolutional neural networks," *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2016, pp. 002020-002025.
- [8] H. T. Vo, G.-H. Yu, T. V. Dang, and J. -Y. Kim, "Late fusion of multimodal deep neural networks for weeds classification," *Computers and Electronics in Agriculture*, vol. 175, 2020, Art. no. 105506.
- [9] A. Jain, K. Nandakumar, and A. Ross, "Score normalization in multimodal biometric systems," *Pattern recognition*, vol. 38, no. 12, pp. 2270-2285, 2005.
- [10] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1-9.