

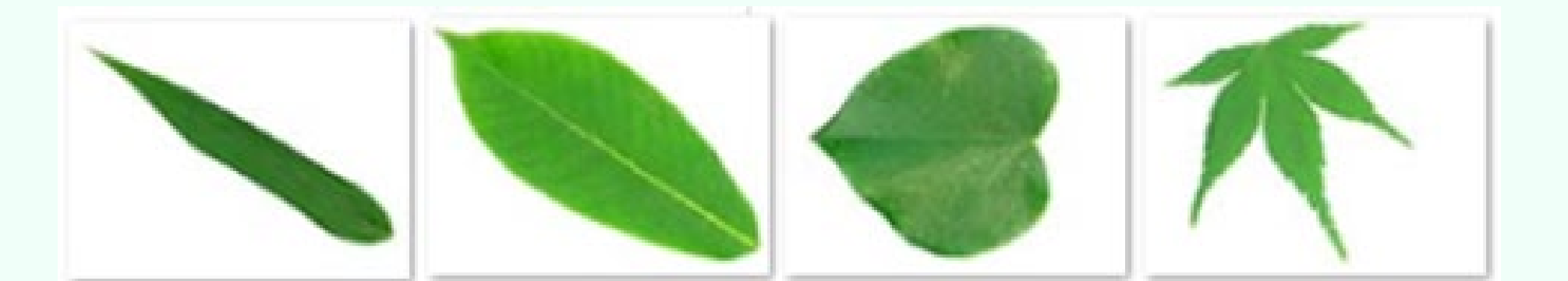
Nâng cao hiệu quả mô hình Convolutional Neural Network nhận diện loài cây qua hình ảnh lá cây sử dụng các lớp Denoising Convolutional Autoencoder và áp dụng Data Augmentation trong quá trình học

TS. Ngô Quốc Việt, Vương Lê Minh Nguyên, Lương Công Tâm

Chuyên ngành Khoa học máy tính trường Đại học Sư phạm Thành phố Hồ Chí Minh

Sơ lược

Nhận diện loài cây thông qua hình ảnh lá cây một cách tự động, cũng như nhiều chủ đề nhận diện hình ảnh trong lĩnh vực Sinh học, là một chủ đề thu hút được sự quan tâm sâu sắc từ những chuyên gia về Computer Vision và Intelligence System. Trong các nghiên cứu trước đây, nhiều nhóm nghiên cứu đã ứng dụng Machine Learning cho phương pháp nhận diện loài cây dựa vào những đặc tính của những chiếc lá của chúng như hình dạng, màu sắc, kết cấu, ..., những đặc tính không trùng nhau bởi bất kỳ một loại lá cây nào. Độ chính xác của các phương pháp trên đã được gia tăng đáng kể từ khi Deep Learning phát triển. Với khả năng mô phỏng cấu trúc hệ thần kinh của con người, các mô hình mạng lưới Convolutional Neural Network này có khả năng tự học lấy các đặc trưng của lá cây, từ đó có thể phân biệt được loài cây thông qua hình ảnh của một chiếc lá. Tuy nhiên, độ chính xác của Convolutional Neural Network có thể bị ảnh hưởng bởi các yếu tố như các điểm ảnh nhiễu, điều kiện ánh sáng xấu, bị khuất hoặc mất một vài đặc trưng quan trọng của chiếc lá. Để khắc phục vấn đề trên, chúng tôi sẽ xây dựng, thử nghiệm và so sánh một mô hình kết hợp giữa Denoising Convolutional Autoencoder và Convolution Neural Network với một mô hình mạng Convolutional Neural Network đơn thuần.



Flavia [1] là bộ dataset chúng tôi sử dụng để huấn luyện các mô hình DL trong đề tài này. Bộ dataset này gồm 1907 hình ảnh của 32 loài cây, tên và danh pháp khoa học của chúng.

Giới thiệu

Thực vật nói chung và cây cối nói riêng đều nắm giữ một vai trò rất quan trọng trong môi trường tự nhiên, chúng hiện đang bao phủ khoảng 30% bề mặt của Trái đất tính đến năm 2010 , là lá phổi tự nhiên với vai trò điều hòa lượng khí O2 và CO2. Trên thế giới hiện nay đã ghi nhận được có hơn 60,065 loài cây [1], nhưng có thể vẫn còn nhiều loài chưa được phát hiện. Với số lượng chủng loại lớn như thế thì việc nhận dạng chúng là quá phức tạp, thậm chí có một số loài có những đặc điểm gần giống nhau, rất khó để có thể phân biệt được.

Chủ đề nhận biết loài cây này không phải là một chủ đề quá mới, nó đã được thực hiện bởi rất nhiều phương pháp khác nhau, từ các phương pháp ứng dụng ML, xử lý ảnh (Image Processing) cho tới các phương pháp ứng dụng DL, NN, CNN... Các phương pháp trên đã từng bước cải thiện được độ chính xác khi nhận diện bằng thông tin đặc trưng trên lá cây. Tuy nhiên vẫn còn gặp một số vấn đề khi gặp các trường hợp ảnh kém chất lượng như lách sáng, nhiễu, lá bị hư hỏng một phần nhỏ.

Nhằm khắc phục những thiếu sót trên, chúng tôi đưa ra một phương pháp kết hợp mô hình Denoising Convolutional Autoencoder với Convolutional Neural Network (DCA-CNN) đồng thời kết hợp sử dụng Data Augmentation trong quá trình học. Để kiểm tra tính khả thi khi sử dụng kết hợp lớp DCA với CNN, chúng tôi sẽ sử dụng cùng một cấu trúc CNN cho 2 mô hình, một mô hình sẽ được liên kết với lớp DCA, sau đó tiến học cho cả 2 cùng học và thực hiện dự đoán trên cùng một train, test sau đó thống kê và đưa ra kết luận.

Sau cùng, chúng tôi đã xây dựng một chương trình ứng dụng giao diện nhỏ, sử dụng mô hình DCA-CNN để nhận diện loại cây của hình ảnh chiếc lá mà người dùng truyền vào.

Phương pháp

Mô hình DCA-CNN kết hợp
- Xây dựng và huấn luyện một mô hình Convolutional Neural Network để phân loại cây qua hình ảnh lá cây và sao chép thành 2 mô hình có cùng cấu trúc.

Model: "model_1"		
Layer (type)	Output Shape	Param #

input_1 (InputLayer)	(None, 384, 512, 3)	0
conv2d_1 (Conv2D)	(None, 192, 256, 64)	1792
conv2d_2 (Conv2D)	(None, 96, 128, 64)	36928
max_pooling2d_1 (MaxPooling2	(None, 48, 64, 64)	0
conv2d_3 (Conv2D)	(None, 24, 32, 128)	73856
conv2d_4 (Conv2D)	(None, 12, 16, 128)	147584
max_pooling2d_2 (MaxPooling2	(None, 6, 8, 128)	0
dropout_1 (Dropout)	(None, 6, 8, 128)	0
flatten_1 (Flatten)	(None, 6144)	0
dense_1 (Dense)	(None, 2048)	12584960
dense_2 (Dense)	(None, 256)	524544
dense_3 (Dense)	(None, 32)	8224

Total params: 13,377,888		
Trainable params: 13,377,888		
Non-trainable params: 0		

- Tiếp tục xây dựng và huấn luyện một mô hình Denoising Convolutional Autoencoder với mục đích khôi phục lại hình ảnh đầu ra, đồng thời khử nhiễu cho ảnh nếu có

Model: "model_6"		
Layer (type)	Output Shape	Param #

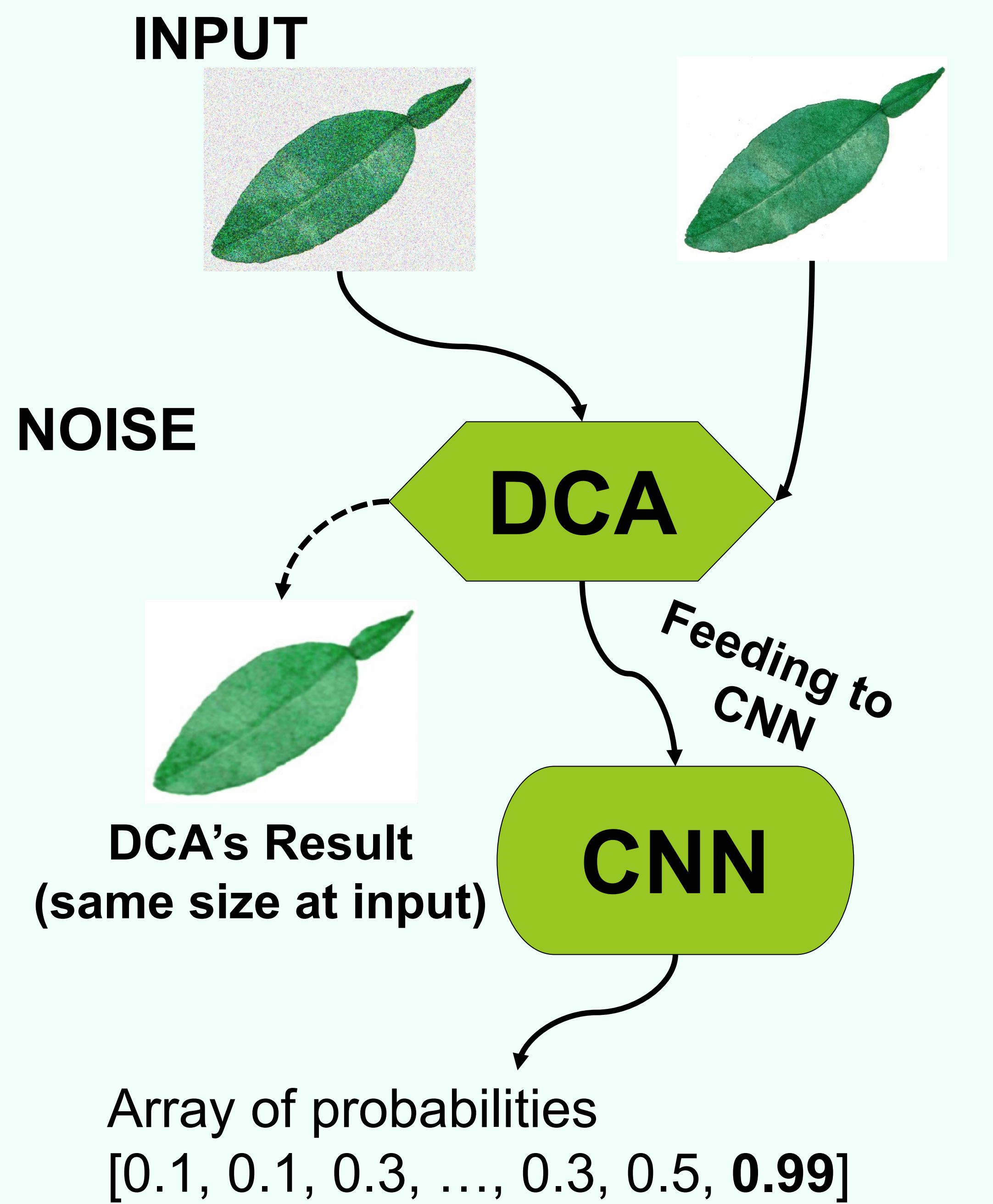
input_7 (InputLayer)	(None, 384, 512, 3)	0
conv2d_15 (Conv2D)	(None, 384, 512, 16)	1216
max_pooling2d_9 (MaxPooling2	(None, 192, 256, 16)	0
conv2d_16 (Conv2D)	(None, 190, 254, 32)	4640
conv2d_transpose_10 (Conv2DT	(None, 192, 256, 16)	4624
up_sampling2d_6 (UpSampling2	(None, 384, 512, 16)	0
conv2d_transpose_11 (Conv2DT	(None, 384, 512, 3)	1203

Total params: 11,683		
Trainable params: 11,683		
Non-trainable params: 0		

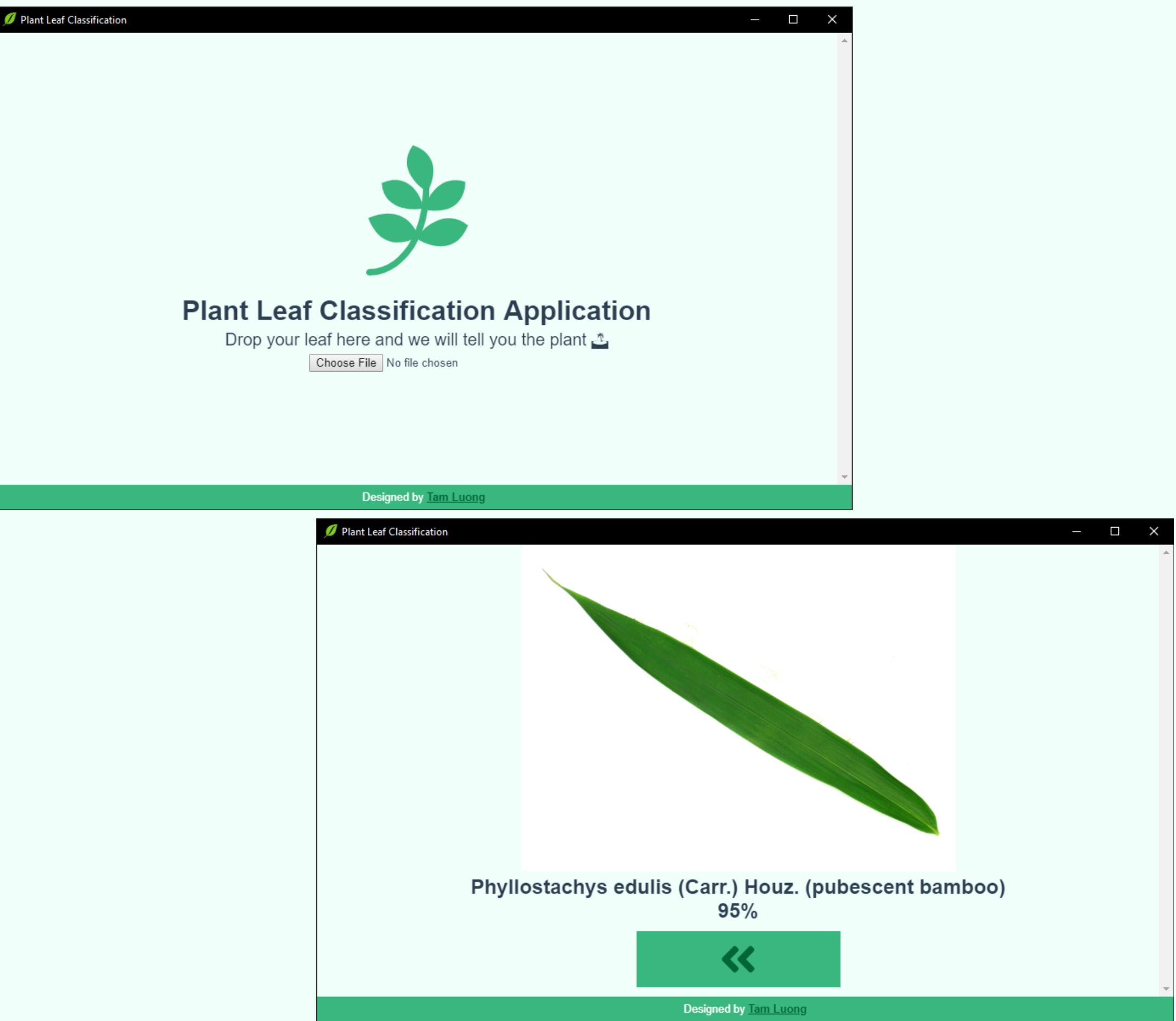
- Sau đó thực hiện nối tiếp mô hình Denoising Convolutional Autoencoder đã được huấn luyện với mô hình Convolutional Neural Network đã sao chép trước đó và cho huấn luyện tiếp mô hình cuối. Quá trình huấn luyện này chỉ thay đổi trọng số của các lớp trong mạng Convolutional Neural Network, các trọng số của Denoising Convolutional Autoencoder phải được giữ nguyên, không thay đổi.

- Quá trình train các mô hình Convolutional Neural Network được áp dụng thêm phương pháp Data Augmentation để làm đa dạng tập dữ liệu huấn luyện và xác thực.

- Luồng thực hiện của mô hình



Ứng dụng nhận diện cây bằng hình ảnh lá cây
- Ứng dụng được xây dựng trên các framework nổi tiếng: Electron, Vue.JS, Webpack, Tensorflow.js và Opencv.js



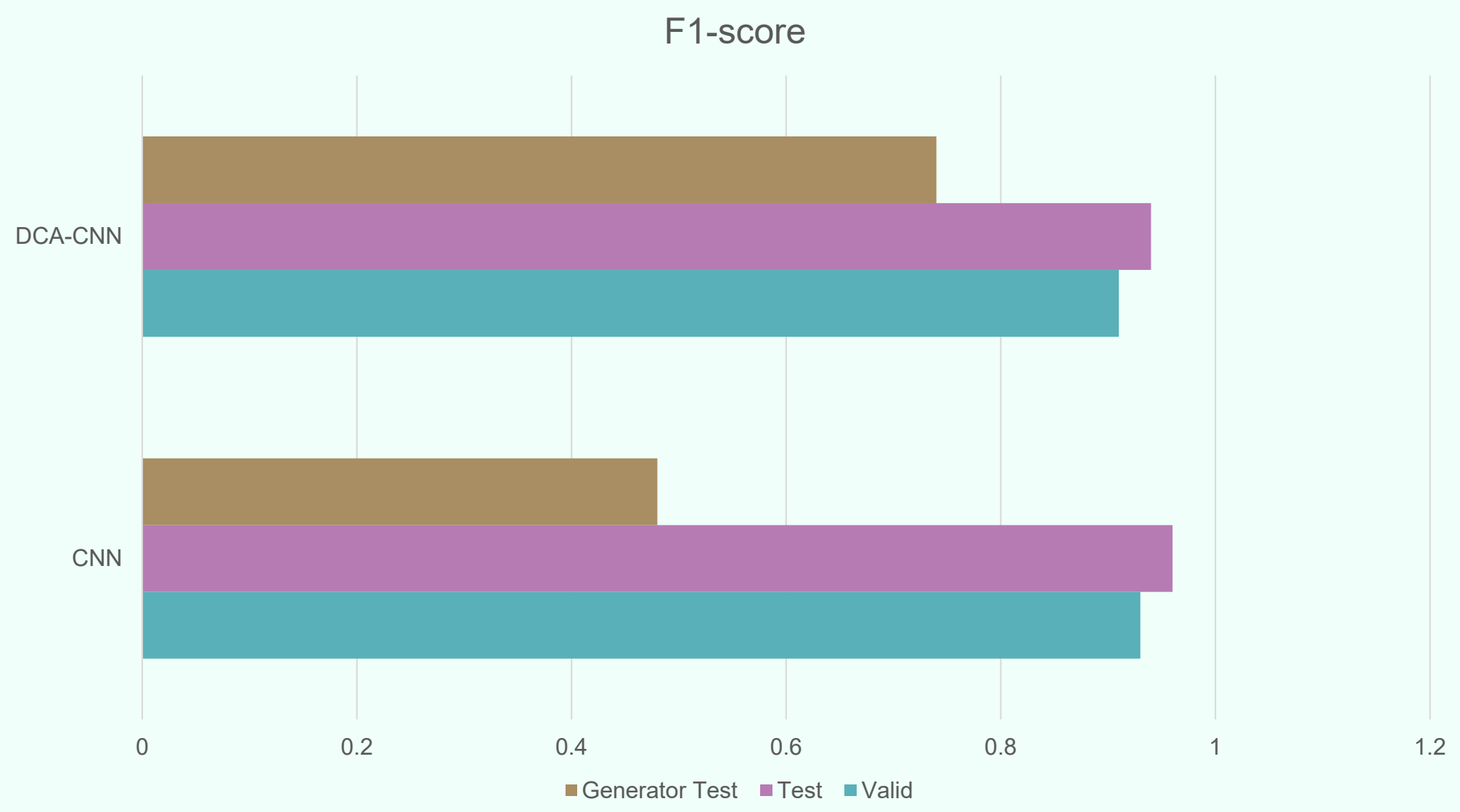
Kết quả thử nghiệm

Bộ dữ liệu được chia thành 2 phần, 85% hình ảnh được sử dụng cho huấn luyện (train – valid set) và 15% hình ảnh dùng cho kiểm thử (test set). Tập train – valid lại tiếp tục được chia thành 2 phần, 85% để huấn luyện mô hình, và 15% còn lại để xác thực tính hiệu quả của quá trình học.

Ta có kích thước các tập dữ liệu như sau:

- Tập train: 1377 hình ảnh
- Tập valid: 243 hình ảnh
- Tập test: 287 hình ảnh
- Tập sinh ra từ test generator: 957 hình ảnh

Kết quả kiểm thử:



Kết luận

Qua kết quả thu được từ thí nghiệm, ta nhận thấy các mô hình CNN rất dễ dẫn đến overfitting trên tập dữ liệu mà nếu không có phương pháp kiểm thử hợp lý thì khi ứng dụng thực tế, mô hình dự đoán của chúng ta sẽ không đủ khả năng khái quát hóa để đạt độ chính xác cao như khi còn trong thí nghiệm.

Denoising Autoencoder là một mô hình hữu hiệu để giảm overfitting khi kết hợp với các mô hình CNN dùng trong thị giác máy tính nói chung và phân loại thực vật thông qua hình ảnh lá cây nói riêng.

Tất cả các công việc của chúng tôi có thể được tải về tại đây: <https://github.com/nguyenvlm/PlantLeafClassification>

