

# ỨNG DỤNG MẠNG HỌC SÂU TRONG PHÁT HIỆN BỆNH CÂY LÚA NƯỚC

Nguyễn Hồ Khánh  
Khoa Điện tử viễn thông

Trường Đại học Công nghệ, ĐHQGHN  
21020562@vnu.edu.vn

**Tóm tắt**—Bài báo cáo này trình bày phương pháp phân loại bệnh trên cây lúa nước bằng cách so sánh hiệu quả giữa ba mạng học sâu: Mạng nơ-ron tích chập (CNN), mạng RBF toàn cục và mạng RBF địa phương. Các mô hình được huấn luyện và đánh giá trên tập dữ liệu hình ảnh bệnh lá lúa. Kết quả thực nghiệm cho thấy mỗi kiến trúc có ưu nhược điểm riêng trong việc nhận diện các loại bệnh, trong đó CNN đạt độ chính xác cao nhất, trong khi RBF địa phương cho thấy khả năng học nhanh và hiệu quả trên tập dữ liệu nhỏ.

**Từ khóa**—CNN, RBF, mạng học sâu

## I. GIỚI THIỆU

Lúa nước là cây lương thực chủ lực của nhiều quốc gia châu Á, trong đó có Việt Nam. Tuy nhiên, năng suất và chất lượng lúa thường bị ảnh hưởng nghiêm trọng bởi các loại bệnh lá như đạo ôn, cháy bìa lá hay vàng lá vi khuẩn. Việc phát hiện và phân loại chính xác bệnh trên lá lúa đóng vai trò quan trọng trong công tác giám sát và phòng trừ bệnh kịp thời.

Nhờ sự phát triển mạnh mẽ của trí tuệ nhân tạo, đặc biệt là các mô hình học sâu (deep learning), các giải pháp nhận dạng hình ảnh đã đạt độ chính xác cao trong nhiều lĩnh vực, bao gồm nông nghiệp thông minh. Trong nghiên cứu này, tôi tập trung xây dựng, so sánh và đánh giá hiệu quả của ba kiến trúc mạng học sâu: mạng nơ-ron tích chập (CNN), mạng RBF toàn cục (global RBF) và mạng RBF địa phương (local RBF) trong bài toán phân loại bệnh trên lá lúa dựa trên hình ảnh.

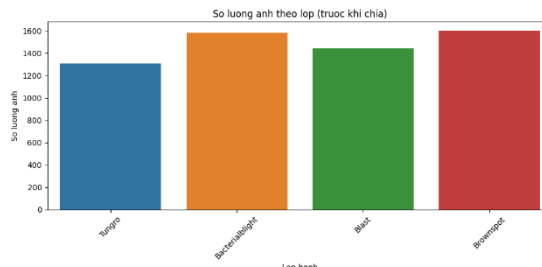
Báo cáo này đánh giá khả năng học, độ chính xác và tốc độ huấn luyện của từng mạng, từ đó đề xuất giải pháp phù hợp cho ứng dụng thực tế trong môi trường nông nghiệp thông minh.

## II. XỬ LÝ DỮ LIỆU

### A. Xây dựng, phân tích và mô tả bộ dữ liệu

Trong nghiên cứu này, chúng tôi sử dụng bộ dữ liệu “Rice Leaf Disease Image” được cung cấp trên nền tảng Kaggle. Bộ dữ liệu bao gồm hình ảnh của các lá lúa bị nhiễm bệnh, được phân loại thành bốn nhóm chính: **Bacterialblight**, **Blast**, **Brown Spot** và **Tungro**.

Tổng cộng, bộ dữ liệu bao gồm 5932 hình ảnh độ phân giải cao, được chia đều giữa các lớp. Các ảnh được chụp trong điều kiện ánh sáng tự nhiên, phong nền đa dạng, giúp phản ánh tốt các tình huống thực tế trong nông nghiệp. Dữ liệu được tổ chức theo cấu trúc thư mục, thuận tiện cho việc huấn luyện với các thư viện học sâu như TensorFlow hoặc PyTorch. Dưới đây là biểu đồ số lượng ảnh của từng lớp bệnh của bộ dữ liệu.



Hình 1: Số lượng ảnh từng lớp

### B. Tiền xử lý dữ liệu

#### 1. Làm sạch và chuẩn hóa dữ liệu

Tất cả hình ảnh trong tập dữ liệu được chuyển đổi về kích thước chuẩn 224×224 pixel còn đối với hai mạng RBF sẽ sử dụng kích thước 64×64 pixel và định dạng RGB. Ngoài ra, các giá trị điểm ảnh được chuẩn hóa về khoảng [0,1] thông qua phép chia cho 255. Việc này giúp tăng tốc độ hội tụ trong quá trình huấn luyện mô hình học sâu.

#### 2. Tăng cường dữ liệu

Để đối phó với tình trạng tập dữ liệu không lớn, kỹ thuật tăng cường dữ liệu được áp dụng trên tập huấn luyện. Cụ thể, các phép biến đổi hình học bao gồm xoay ngẫu nhiên (tối đa 20 độ), tịnh tiến theo chiều ngang và dọc (20%), cắt xén (shear 0.2 rad), thu phóng (zoom), và lật ngang ảnh. Các phép biến đổi này được thực hiện bằng thư viện **ImageDataGenerator**, giúp mô hình tiếp xúc với nhiều dạng biến thiên khác nhau của bệnh lá mà không cần thu thập thêm dữ liệu thực tế.

#### 3. Chia dữ liệu

Dữ liệu được chia thành ba tập riêng biệt:

- Tập huấn luyện** chiếm 80% dữ liệu, dùng để mô hình học.
- Tập kiểm định (validation)** chiếm 10%, nhằm điều chỉnh siêu tham số và tránh quá khớp.
- Tập kiểm tra (test)** chiếm 10%, dùng để đánh giá khách quan hiệu năng cuối cùng. Việc chia tách được thực hiện ngẫu nhiên bằng thư viện random và theo đúng cấu trúc.

Thông tin về các lớp phân loại cũng được lưu trữ dưới dạng tệp class\_indices.json để thuận tiện cho việc ánh xạ đầu ra mô hình sang nhãn bệnh tương ứng.

### III. XÂY DỰNG MÔ HÌNH MẠNG

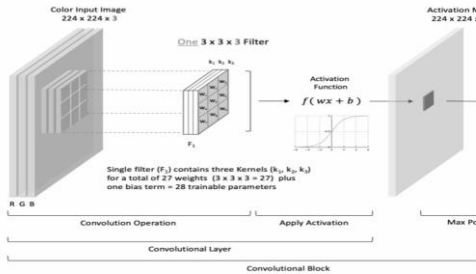
Trong báo cáo này, tôi thực hiện xây dựng cấu trúc 3 mạng là: CNN, RBF tiêu chuẩn và RBF địa phương để thực hiện bài toán phân loại bệnh trên cây lúa. Dưới đây là trình bày về ba cấu trúc mạng trên.

#### A. Cấu trúc mạng

##### 1. Mạng CNN (Convolutional Neural Network)

Mạng CNN được xây dựng theo hướng điển hình, gồm ba khối tích chập với cấu trúc lặp lại:

- Conv2D với số lượng kênh lần lượt là 32, 64, và 128, kích thước kernel  $3 \times 3$ , hàm kích hoạt phi tuyến ReLU.
- BatchNormalization nhằm chuẩn hóa đầu ra mỗi tầng, giúp tăng ổn định huấn luyện.
- MaxPooling2D để giảm chiều không gian và trích xuất đặc trưng bền vững.

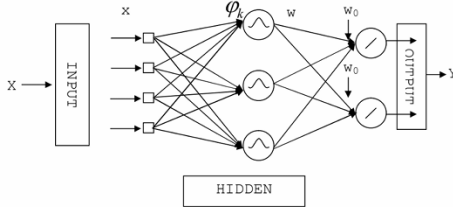


Hình 2: Cấu trúc tầng tích chập CNN

##### 2. Mạng RBF

Cấu trúc mạng RBF:

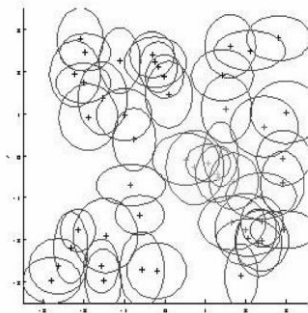
- Là một mạng truyền thẳng hai lớp
- Lớp ẩn sử dụng các hàm cơ sở hướng tâm (radial basis functions).
- Lớp đầu ra là tổng tuyến tính của các giá trị từ lớp ẩn



Hình 3: Cấu trúc mạng RBF

##### 3. Mạng RBF địa phương

Tương tự, mạng RBF có cấu trúc mạng tương đương mạng RBF. Điểm khác biệt của RBF địa phương là tại mỗi điểm dự đoán sẽ xây dựng một mô hình RBF riêng biệt gần điểm đó. Tại mỗi điểm, ta sẽ fit vectơ trong vùng lân cận thay vì đường trung bình.



Hình 4: RBF địa phương

#### B. Quá trình huấn luyện mạng

##### 1. Huấn luyện mạng CNN

Đầu tiên, mạng CNN bao gồm ba tầng tích chập với số lượng kênh lần lượt là 32, 64 và 128. Tích hợp trong tầng tích chập là hai hàm BatchNormalization và MaxPooling2D. Cuối cùng là tầng phân loại gồm Flatten, Dense và Dropout.

Thứ hai, để cải thiện hiệu quả huấn luyện và tránh overfitting, thiết kế hàm callback bao gồm hai hàm chính EarlyStopping và ReduceLROnPlateau.

Cuối cùng, mô hình được huấn luyện với dữ liệu của tập train và đánh giá hiệu suất sau mỗi epoch dựa trên tập val.

##### 2. Huấn luyện mạng RBF

Quy trình huấn luyện mạng RBF bao gồm hai giai đoạn chính:

- Xác định trọng số từ input đến RBF Layer (tầng tùy chỉnh do tự xây dựng). Thường là lựa chọn tâm (centroids) và độ rộng beta của các hàm Gaussian.

- Xác định trọng số từ hidden layer đến output

- Mạng RBF xấp xỉ một hàm số  $f(x)$  theo công thức:

$$f(x) = \sum_{i=1}^n w_i \phi(r) \text{ với } r = \|x - c_i\|$$

- $x \in \mathbb{R}^n$ : vec-tơ đầu vào

- $c_i$ : tâm của hàm cơ sở (giá trị trọng số ở hidden layer – "centroid")

- $w_i$ : trọng số kết nối từ hidden layer đến lớp đầu ra

- $\phi(r)$ : hàm kích hoạt hướng tâm (radial basis function)

- $r$ : Khoảng cách từ  $x$  đến  $c_i$

- Ví dụ xét hàm kích hoạt Gaussian:  $\phi(r) = \exp\left(-\frac{r^2}{2\sigma_i^2}\right)$

- $\sigma_i$ : tham số điều chỉnh độ rộng của Gaussian (giống "độ mượt").

- Quá trình huấn luyện là quá trình tìm giá trị phù hợp của các biến  $w_{ij}, c_{ij}, \sigma_j$

#### C. Trích xuất đặc trưng

##### 1. Đặc trưng mạng CNN

Với CNN, đặc trưng được học một cách tự động thông qua các tầng tích chập và chuẩn hóa. Mạng bao gồm ba khối tích chập với số lượng kênh tăng dần (32, 64, 128), mỗi khối gồm: Conv2D, BatchNormalization, và MaxPooling. Sau khi được làm phẳng, các đặc trưng đầu ra được đưa qua lớp Dense để biểu diễn thông tin bệnh. Đây là cách tiếp cận học đặc trưng đặc trưng nội tại mạnh mẽ, cho phép mô hình tự trích xuất các đặc điểm phức tạp như màu sắc vết bệnh, hình dạng tổn thương, và hoa văn lá.

Đối với các tầng tích chập, việc lựa chọn số layer tăng dần từ 32 lên 128 giúp cho mô hình sẽ học được các đặc trưng đơn giản trước như cạnh, góc, .... Sau đó, mô hình sẽ học các đường gân của lá và lớp cuối cùng sẽ học sâu những đặc trưng của bệnh, vùng tổn thương của cây lúa. Trong mô hình này, tôi chỉ sử dụng 3 tầng tích chập để tránh mô hình học quá sâu và overfitting.

##### 2. Đặc trưng mạng RBF

Do RBF không tự học đặc trưng từ ảnh thô, nên đặc trưng đầu vào cần được xử lý trước. Ảnh được làm phẳng từ kích thước 3D (64x64x3) thành vector 1 chiều. Sau đó, phương pháp phân cụm KMeans được áp dụng để tìm 300 tâm cụm (center) trong không gian đặc trưng. Khoảng cách giữa các ảnh và các tâm được đưa vào một lớp RBF tùy chỉnh để tính đầu ra theo công thức Gaussian. Một hệ số  $\beta$  duy nhất được tính từ khoảng cách trung vị giữa các tâm cụm để kiểm soát độ rộng của hàm cơ sở. Đặc trưng đầu ra từ lớp RBF sau đó được đưa vào các lớp Dense để phân loại.

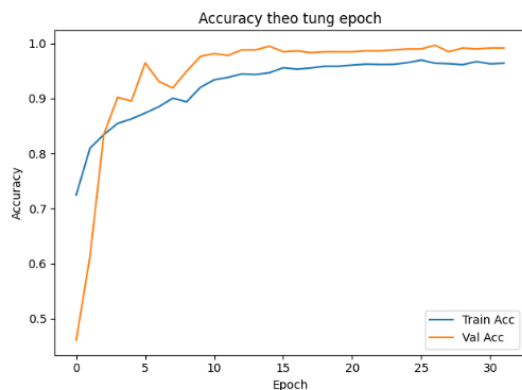
### 3. Đặc trưng mạng RBF địa phương

Tương tự mạng RBF tiêu chuẩn, ảnh đầu vào được làm phẳng và dùng KMeans để tìm center. Tuy nhiên, khác biệt cốt lõi là mỗi center có một hệ số  $\beta$  riêng, được tính toán dựa trên khoảng cách trung bình tới  $k$  ( $k=5$ ) hàng xóm gần nhất. Điều này giúp RBF địa phương thích ứng tốt hơn với độ phân bố không đều của không gian đặc trưng, tăng tính linh hoạt trong mô hình hóa dữ liệu phi tuyến tính. Lớp RBF được thiết kế lại để xử lý từng  $\beta$  tương ứng với mỗi center, cải thiện khả năng phân biệt bệnh có biểu hiện tương tự nhau.

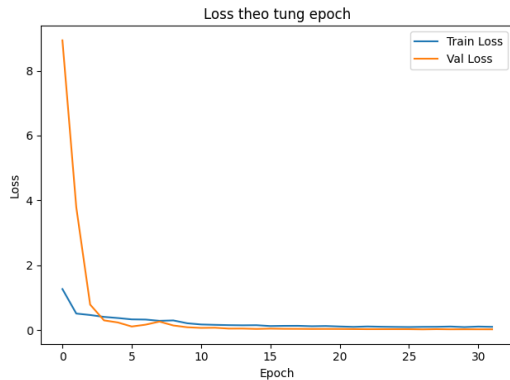
## IV. KẾT QUẢ VÀ ĐÁNH GIÁ

### A. Kết quả

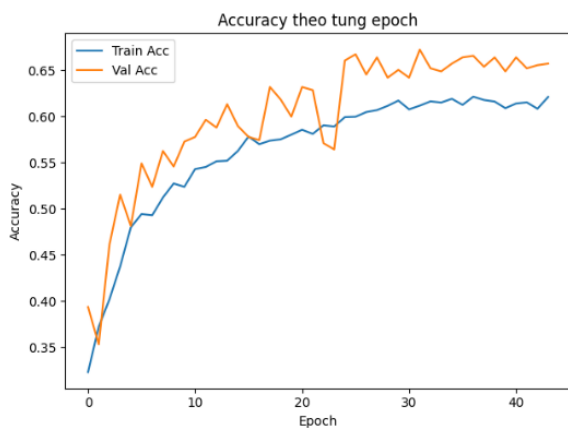
Sau khi thực hiện huấn luyện mô hình của ba mạng học CNN, RBF và RBF địa phương tôi thu được kết quả của ba mô hình như sau:



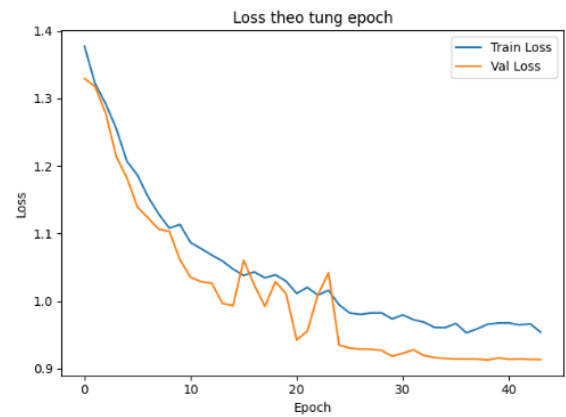
Hình 5: Accuracy CNN



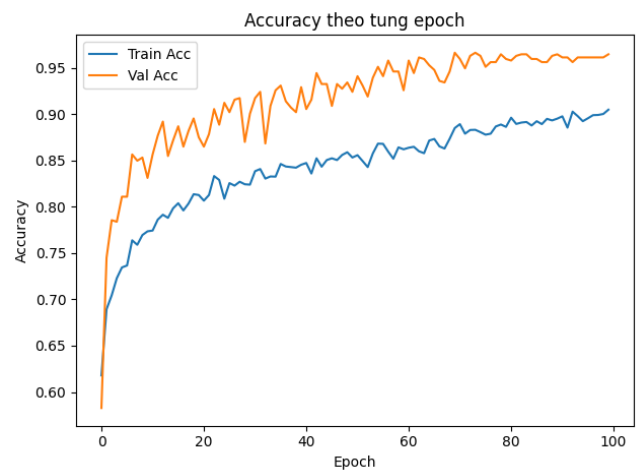
Hình 6: Loss CNN



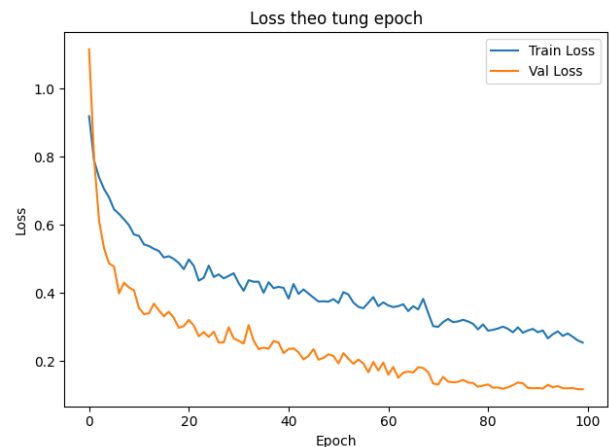
Hình 7: Accuracy RBF



Hình 8: Loss RBF



Hình 9: Accuracy RBF địa phương



Hình 10: Loss RBF địa phương

Dựa vào kết quả Accuracy, Loss và thực tế khi chạy thử mô hình thu được sau khi huấn luyện mô hình. Có thể thấy, trong bài toán này việc sử dụng mô hình mạng CNN là tốt nhất. Dưới đây là lý do tôi đánh giá mạng CNN là tốt nhất:

- Hiệu suất cao nhất: Accuracy validation ~ 99% và loss rất thấp. Khi kiểm tra thực tế có 3 ảnh sai trong tập test so với (205 của RBF và 25 của RBF địa phương).
- Sự ổn định: Đường loss và accuracy mượt, ít dao động nhất. Điều này cho thấy mô hình học tốt và ổn định.
- Không overfitting: Độ chính xác validation luôn cao hơn hoặc gần bằng tập train, không bị dao động và rơi sau một vài epoch.

## B. Đánh giá

Đối với ý nghĩa thực tế của mô hình mạng CNN:

- **Ứng dụng tốt vào phân loại bệnh lá lúa:** Với độ chính xác validation gần 99%, mô hình CNN có thể dùng để triển khai thật vào các hệ thống giám sát bệnh tự động.
- **Tối ưu hóa tài nguyên:** Trong khi CNN yêu cầu nhiều tài nguyên huấn luyện hơn, kết quả của nó là vượt trội đáng kể so với RBF nên xứng đáng để sử dụng trong thực tế.

Dựa vào kết quả thu được, tôi đưa ra những điểm mạnh và điểm yếu của ba mô hình mạng như sau:

### Mạng CNN:

- **Điểm mạnh:**
  - Độ chính xác cao, học tốt đặc trưng ảnh.
  - Ổn định, khả năng tổng quát hóa tốt.
- **Điểm yếu:**
  - Tốn tài nguyên tính toán.
  - Cần nhiều dữ liệu để tối ưu.

### Mạng RBF:

- **Điểm mạnh:**
  - Đơn giản, huấn luyện nhanh.
  - Không cần nhiều layer.
- **Điểm yếu:**
  - Accuracy thấp, dễ dao động.
  - Không học tốt phi tuyến tính phức tạp của ảnh.

### Mạng RBF địa phương:

- **Điểm mạnh:**
  - Cải thiện rõ rệt so với RBF thường.
  - Độ chính xác khá cao (~96% val acc).
  - Huấn luyện nhanh hơn CNN.
- **Điểm yếu:**
  - Chưa đạt mức hiệu suất cao như CNN.
  - Hiệu quả phụ thuộc cách chọn tâm (center) và  $\sigma$ .

## V. KẾT LUẬN

Dựa vào kết quả của huấn luyện mô hình và những đánh giá về ba mạng như đã đề cập ở trên. Tôi đưa ra những hướng cải tiến và đề xuất cho các mạng như sau:

### Với CNN:

- Data Augmentation thêm (nếu chưa dùng): xoay, lật, thay đổi ánh sáng...
- Early stopping để tránh overfitting khi dữ liệu nhỏ.
- Có thể sử dụng các mô hình mạng pretrained như ResNet hoặc MobileNet nếu tài nguyên hạn chế.

### Với RBF:

- Cải tiến chọn center bằng kỹ thuật như KMeans++ hoặc dùng clustering có trọng số.
- Thực hiện tăng dữ liệu đầu vào và chia lại tỉ lệ tập train, val, test.
- Có thể kết hợp cùng mạng CNN để trích xuất đặc trưng rồi qua lớp RBF.
- Sử dụng trọng số để học thay vì chỉ sử dụng validation.