

# ỨNG DỤNG MẠNG NƠ RON TÍCH CHẬP (CNN) TRONG PHÂN LOẠI NGUY CƠ BỆNH VÔNG MẠC ĐÁI THÁO ĐƯỜNG TỪ ẢNH ĐÁY MẮT

TS. Trần Cảnh Dương

Trường Đại học Hòa Bình

Tác giả liên hệ: [tcduong@daihochoabinh.edu.vn](mailto:tcduong@daihochoabinh.edu.vn)

Ngày nhận: 10/7/2025

Ngày nhận bản sửa: 10/8/2025

Ngày duyệt đăng: 24/10/2025

DOI: 10.71192/947062tqtbxo

## Tóm tắt

Bệnh vông mạc đái tháo đường là một trong những nguyên nhân hàng đầu gây mất thị lực nếu không được phát hiện sớm. Trong nghiên cứu này, chúng tôi xây dựng một mô hình mạng nơ-ron tích chập (Convolutional Neural Network - CNN) nhằm phân loại ảnh đáy mắt theo mức nguy cơ bệnh lý. Dữ liệu ảnh được chuẩn hóa về kích thước và định dạng màu, đồng thời, mô hình được huấn luyện trên tập ảnh gán nhãn thủ công. Mô hình đạt độ chính xác 85,3% trên tập kiểm thử. Kiểm nghiệm thực tế bằng ảnh bệnh lý mới, nằm ngoài tập huấn luyện, cho thấy mô hình vẫn nhận diện chính xác, khẳng định tính ứng dụng thực tiễn trong hỗ trợ sàng lọc bệnh lý vông mạc.

**Từ khóa:** Bệnh vông mạc đái tháo đường, đáy mắt, mạng nơ-ron tích chập.

## Application of Convolutional Neural Networks (CNN) in Classifying Diabetic Retinopathy Risks from Fundus Images

Dr. Tran Canh Duong

Hoa Binh University

Corresponding Author: [tcduong@daihochoabinh.edu.vn](mailto:tcduong@daihochoabinh.edu.vn)

## Abstract

Diabetic retinopathy is one of the leading causes of vision loss if not detected early. In this study, we developed a convolutional neural network (CNN) model to classify fundus images based on pathological risk levels. The image data were normalized in terms of size and color format, and the model was trained on a manually labeled dataset. It achieved an accuracy of 85.3% on the test set. Real-world validation using new pathological images outside the training set confirmed the model's ability to accurately detect retinopathy, reinforcing its practical value in supporting early screening.

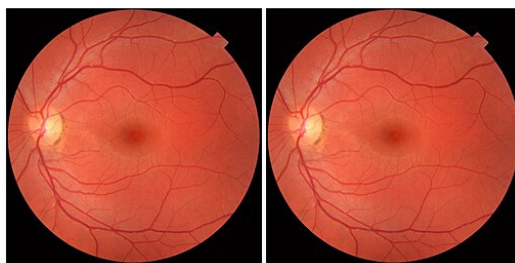
**Keywords:** Diabetic retinopathy, fundus image, convolutional neural network.

## 1. Giới thiệu

Đáy mắt (thể fundus) là bề mặt bên trong của mắt đối diện với thủy tinh thể và bao gồm võng mạc, đĩa thị, hoàng điểm, hố thị và cực sau. Đáy mắt có thể được kiểm

tra bằng phương pháp soi đáy mắt và/hoặc chụp ảnh đáy mắt [1]. Hình 1 mô tả ảnh đáy mắt phải (hình ảnh bên trái) và ảnh đáy mắt trái (hình ảnh bên phải), nhìn từ phía trước (như thể đối mặt với người xem).

**Hình 1.** Ảnh đáy mắt phải (hình ảnh bên trái) và ảnh đáy mắt trái (hình ảnh bên phải), nhìn từ phía trước



Nguồn: [1]

Ảnh chụp đáy mắt, cho thấy đĩa thị giác là một vùng sáng bên phải, nơi các mạch máu hội tụ. Đốm bên trái trung tâm

là hoàng điểm. Đốm màu xám, mờ hơn ở trung tâm là một hiện tượng bóng mờ như được thể hiện ở Hình 2.

**Hình 2.** Ảnh chụp đáy mắt, cho thấy đĩa thị giác và hoàng điểm

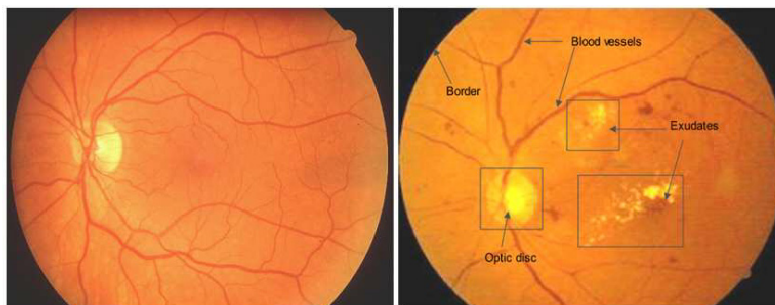


Nguồn: [1]

Bệnh võng mạc đái tháo đường là một tình trạng bệnh lý, trong đó, võng mạc bị tổn thương do bệnh tiểu đường. Đây là nguyên nhân hàng đầu gây mù lòa ở các nước phát triển và là một trong những nguyên nhân hàng đầu gây mất thị

lực trên thế giới, mặc dù đã có nhiều liệu pháp mới và phương pháp điều trị cải tiến giúp người bệnh sống chung với bệnh tiểu đường [2]. Hình 3 mô tả ảnh đáy mắt bình thường (bên trái) và ảnh đáy mắt bị bệnh đái tháo đường

**Hình 3.** Ảnh đáy mắt bình thường (bên trái) và ảnh đáy mắt bị bệnh đái tháo đường



Nguồn: [2]

Bệnh võng mạc đái tháo đường (Diabetic Retinopathy - DR) có thể biểu hiện qua xuất huyết, phù hoàng điểm hoặc tăng sinh mạch máu trong ảnh đáy mắt. Phân tích ảnh đáy mắt là một phương pháp không xâm lấn và hiệu quả để phát hiện sớm các dấu hiệu bệnh lý. Tuy nhiên, quy trình chẩn đoán thủ công đòi hỏi chuyên gia nhãn khoa giàu kinh nghiệm và mất nhiều thời gian.

Gần đây, các mô hình học sâu đã cho thấy hiệu quả vượt trội trong phân loại ảnh y khoa. Các nghiên cứu như của Al-Omairi et al. và Singh et al. đã áp dụng CNN như ResNet, VGGNet, hoặc EfficientNet để phát hiện DR với độ chính xác cao, trên 95% [3-4]. Tuy nhiên, hầu hết các nghiên cứu tập trung vào các mô hình lớn và phức tạp. Nghiên cứu này hướng tới một mô hình

CNN gọn nhẹ, dễ triển khai, phù hợp với thiết bị có cấu hình vừa phải như phòng khám cấp cơ sở hoặc ứng dụng di động y tế.

## 2. Phương pháp nghiên cứu

### 2.1. Tập dữ liệu và tiền xử lý

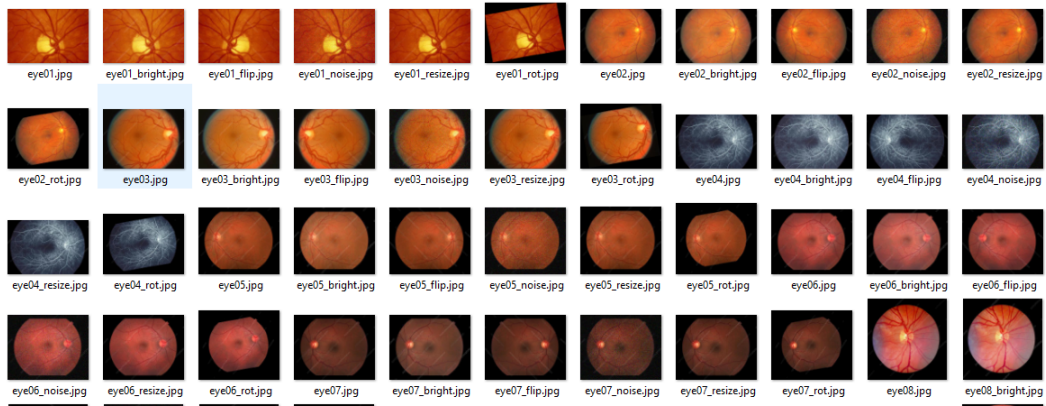
Chúng tôi xây dựng một tập ảnh đáy mắt gồm hai lớp:

- Lớp “Nguy cơ thấp”: ảnh bình thường hoặc chưa có dấu hiệu rõ rệt.

- Lớp “Nguy cơ cao”: ảnh có phù hoàng điểm, xuất huyết, hoặc dấu hiệu tăng sinh mạch máu.

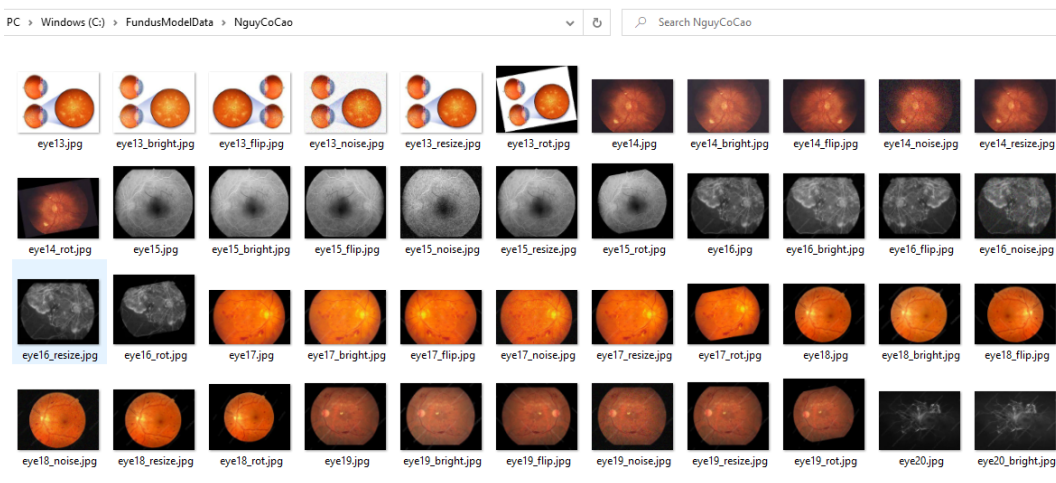
Ảnh được resize về kích thước chuẩn 224×224pixel và chuyển sang định dạng RGB để tương thích với kiến trúc CNN, theo đề xuất tiền xử lý bởi Singh et al. [4]. Hình 4 mô tả một số ảnh đáy mắt thuộc lớp nguy cơ thấp và Hình 5 mô tả ảnh đáy mắt của lớp nguy cơ cao.

Hình 4. Một số ảnh đáy mắt thuộc lớp nguy cơ thấp



Nguồn: [2] và kết quả nghiên cứu của tác giả

Hình 5. Một số ảnh đáy mắt của lớp nguy cơ cao



Nguồn: [2] và kết quả nghiên cứu của tác giả

2.2. Mô hình CNN đề xuất

Một trong những mục đích chính của ngành Trí tuệ nhân tạo (AI-Artificial Intelligence) là nghiên cứu và phát triển các mô hình mô phỏng những khả năng về tư duy, phân tích và xử lý thông tin của con người. Mạng nơ ron nhân tạo là một trong những chuyên ngành của trí tuệ nhân tạo được phát triển mạnh mẽ hiện nay với các ứng dụng rộng rãi nhiều vấn đề về kỹ thuật [5].

Các mạng nơ ron có một đặc trưng khác so với các hệ thống tính toán tuyến tính và phi tuyến trước đây đó là mỗi mạng nơ ron được trang bị một hoặc nhiều thuật toán học cho mạng [6]. Giá trị cực tiểu của một hệ phương trình phi tuyến đa biến có thể được xác định bằng đồ thị hoặc các chùm lệnh [7].

Mạng nơ ron nhân tạo nhiều lớp - MLP

(Multi Layer Perceptron) là một mạng truyền thẳng với các khối cơ bản là các nơ ron McCulloch - Pits. Các nơ ron được sắp xếp thành các lớp (layer). Một lớp nối đến các kênh tín hiệu đầu vào (input layer), một lớp nối đến các kênh tín hiệu đầu ra (out layer) và có thể thêm một số lớp trung gian (lớp ẩn - hidden layer) [6].

Dựa theo bộ số liệu mẫu, ta cần xác định một mạng MLP với các thông số cấu trúc và các trọng số ghép nối tương ứng theo cấu trúc đã chọn đảm bảo mạng MLP có thể xác định lại được ảnh xạ từ đầu vào đến đầu ra của các số liệu mẫu [8]. Các giá trị đầu vào xi được đưa vào đối tượng nghiên cứu sẽ cho ra kết quả di và giá trị xi được đưa vào mạng MLP sẽ cho kết quả ở đầu ra là yi. Như vậy, với cùng một đầu vào thì đầu ra của mạng MLP phải xấp xỉ bằng đầu ra của đối tượng nghiên cứu. Sai

số tổng cộng trên các mẫu tiến tới một giá trị cực tiểu.

Hàm truyền đạt của mạng nơ ron MLP với một lớp ẩn, N đầu vào, M nơ ron trên lớp ẩn, K đầu ra là một hàm phi tuyến được thể hiện theo phương trình sau:

$$y_1 = f_2 \left\{ \sum_{j=0}^M \left[ f_1 \left( \sum_{k=0}^N x_k W_{jk} \right) V_{ij} \right] \right\}$$

Trong đó:  $W_{jk}$  là trọng số ghép nối giữa lớp đầu vào và lớp ẩn.  $V_{ij}$  là trọng số ghép nối giữa lớp ẩn và lớp đầu ra [9].

Kiến trúc mạng được dùng gồm ba tầng tích chập, mỗi tầng có tích chập, sau đó, được chuẩn hóa hàng loạt và được tối ưu hóa và phân loại.

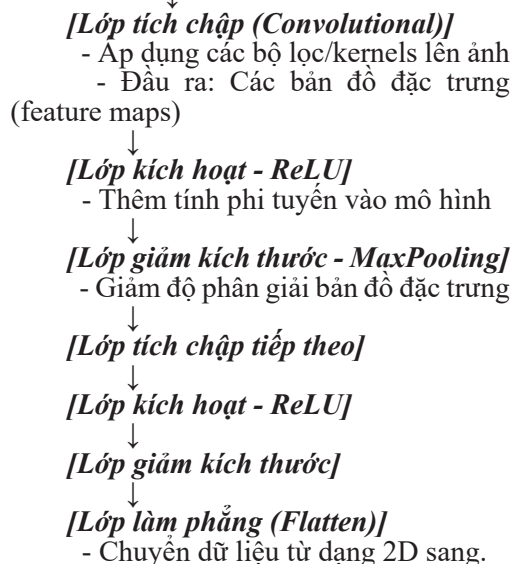
Thiết lập huấn luyện theo thuật toán tối ưu Adam; Kích thước lô (Batch size): 16; Epochs: 10; Tỷ lệ học tập (Learning rate): 0.001.

Mô hình huấn luyện bằng hàm trainNetwork của MATLAB trên tập dữ liệu đã tăng cường. Thiết kế này được

tham khảo từ cấu trúc CNN gọn nhẹ trong nghiên cứu của Sinha et al. (2024) [10].

Mô hình tích chập CNN có các lớp được biểu diễn như sau:

**Hình ảnh đầu vào (ví dụ: 28x28x1)**



**Hình 6.** Các lớp của mô hình tích chập CNN



*Nguồn: Kết quả nghiên cứu của tác giả*

**3. Kết quả mô phỏng**

Để có kết quả mô phỏng, một chương trình phần mềm được thực hiện. Sau đây là một phần chương trình:

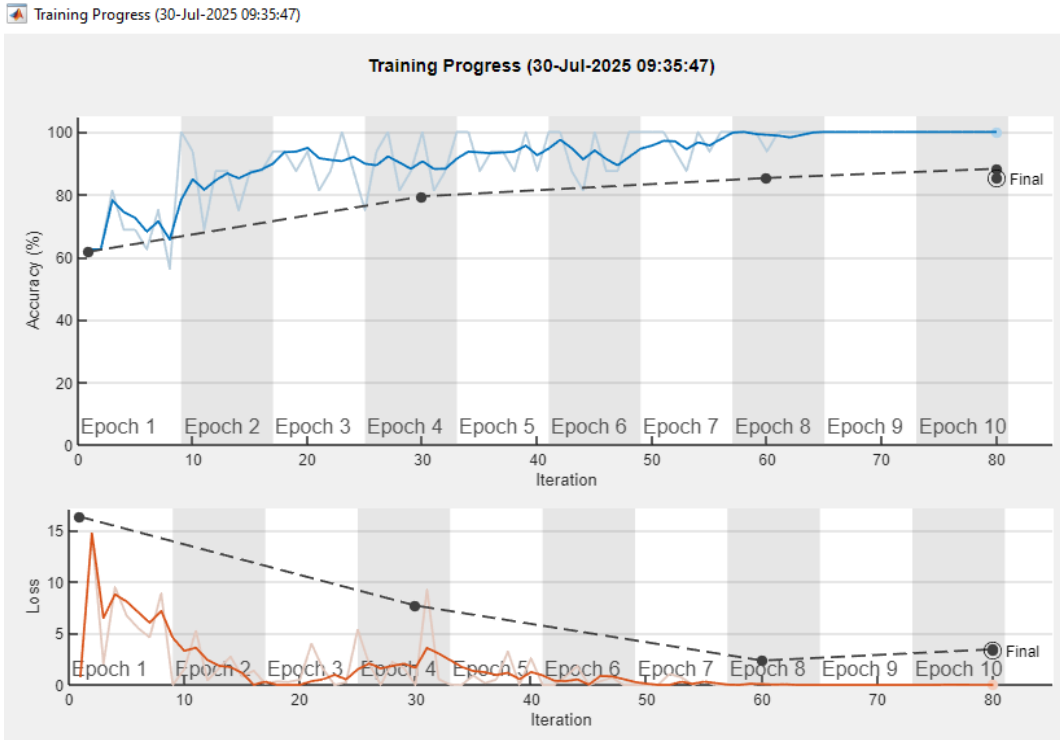
```
FUNDUS MODEL CNN - PHÂN
LOẠI NGUY CƠ VÕNG MẠC
% Tác giả: Trần Cảnh Dương
% Mô tả: Chuẩn hóa dữ liệu ảnh
fundus, huấn luyện mô hình CNN, dự đoán
& đánh giá
clc; clear; close all; %1. Tạo
imageDatastore và kiểm tra nhãn
```

```
dataFolder = 'C:\FundusModelData';
% Gồm NguyCoThap và NguyCoCao
imds = imageDatastore(dataFolder, ...
'IncludeSubfolders', true, ...
'LabelSource', 'foldernames');
disp("Tổng số ảnh theo lớp:");
disp(countEachLabel(imds));
%2. Chia dữ liệu huấn luyện
và kiểm thử [trainImds, testImds] =
splitEachLabel(imds, 0.8, 'randomized');
% 3. Resize ảnh và chuyển ảnh
grayscale sang RGB trainData =
```

```
augmentedImageDatastore([224 224 3],
trainImds, ...
    'ColorPreprocessing', 'gray2rgb');
testData = augmentedImageDatastore([224
224 3], testImds, ...
    'ColorPreprocessing', 'gray2rgb');
.....
.....
```

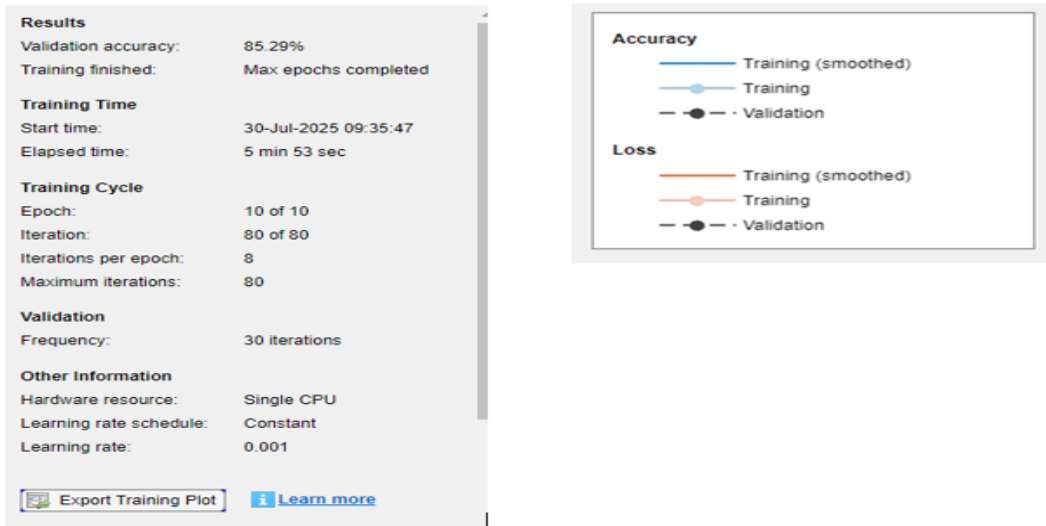
```
predictedLabels = classify(net,
testData); trueLabels = testImds.Labels;
figure;
confusionchart(trueLabels,
predictedLabels);
title('Ma trận nhầm lẫn CNN –
Fundus Risk Classification');
```

**Hình 7.** Đồ thị biểu diễn quá trình máy học và sai số



*Nguồn: Kết quả nghiên cứu của tác giả*

**Hình 8.** Chú thích cho đồ thị biểu diễn quá trình máy học và sai số

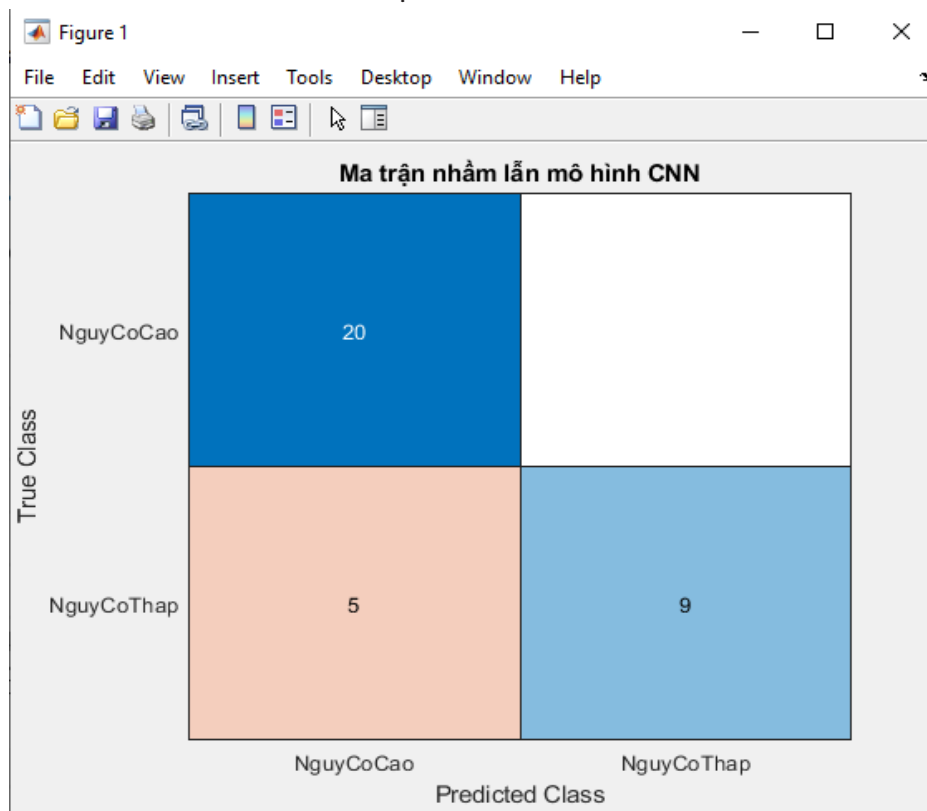


*Nguồn: Kết quả nghiên cứu của tác giả*

Trên đồ thị, “Accuracy” biểu thị độ chính xác trên tập huấn luyện và kiểm thử. Nếu Accuracy tiến dần lên 90%+, mô hình học tốt. “Loss” là mức sai lệch. Loss giảm dần chứng tỏ mô hình học đúng hướng.

“Validation Accuracy” biểu thị độ chính xác trên tập kiểm thử. Nếu giá trị này gần với độ chính xác trong quá trình đào tạo (train accuracy) chứng tỏ mô hình không bị trạng bị quá mức (overfitting).

**Hình 9.** Ma trận nhầm lẫn mô hình CNN



*Nguồn: Kết quả nghiên cứu của tác giả*

Trên tập kiểm thử gồm 34 ảnh, mô hình cho kết quả: Dự đoán đúng: 29 ảnh, dự đoán sai: 5 ảnh (chủ yếu là ảnh bình thường bị phân loại nhầm là nguy cơ cao).

Phân tích ma trận nhầm lẫn: Điểm nổi bật là mô hình không bỏ sót bất kỳ ảnh bệnh lý nào, đáp ứng nguyên tắc “an toàn lâm sàng” được nhấn mạnh trong khảo sát của Maghdid et al. [11].

Độ chính xác (Accuracy) =  $(20 + 9) / 34 = 85,3\%$ .

Đây là mức độ chính xác rất tốt, đặc biệt vì mô hình không bỏ sót ảnh bệnh lý (rất quan trọng trong y khoa). Mặt khác, chấp nhận nhầm một số ảnh bình thường đảm bảo an toàn về mặt lâm sàng. Mô hình có khuynh hướng cảnh báo sớm, khi một số ảnh bình thường bị dự đoán là nguy cơ cao sẽ phù hợp với định hướng mô phỏng hệ cảnh báo. Mô hình đoán đúng 85% số ảnh kiểm thử, nghĩa là xác suất phân loại đúng có tỷ lệ cao trong thực tế. Quan trọng hơn đó là mô hình không bỏ sót ảnh “Nguy

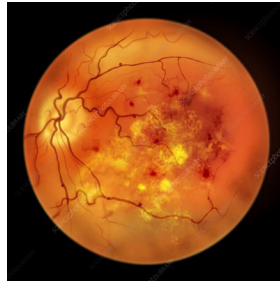
ơ cao” như vậy điều này rất an toàn cho người bệnh, ngay cả khi có một số cảnh báo sớm hơi dư.

Ngoài độ chính xác tổng thể, các chỉ số đánh giá chi tiết được ghi nhận như sau: Precision = 0.85, Recall = 1.00, F1-score = 0.92 đối với lớp “Nguy cơ cao”; AUC đạt 0.96 khi đánh giá bằng ROC Curve. Điều này cho thấy mô hình có độ bao phủ cao, không bỏ sót ảnh bệnh lý.

#### 4. Kiểm nghiệm thực tế

Một ảnh đáy mắt thực tế ngoài tập huấn luyện, có đặc điểm phù hoàng điểm và xuất huyết, được đưa vào mô hình đã huấn luyện. Mô hình phân loại là “Nguy cơ cao”, điều này xác nhận tính tổng quát hóa và năng lực xử lý ảnh thực tế. Cách tiếp cận này phù hợp với khuyến nghị của Parthiban et al. về kiểm định mô hình bằng dữ liệu mới [9]. Hình 10 mô tả một ảnh đáy mắt thực tế ngoài tập huấn luyện, có đặc điểm phù hoàng điểm và xuất huyết.

**Hình 10.** Ảnh đáy mắt thực tế ngoài tập huấn luyện

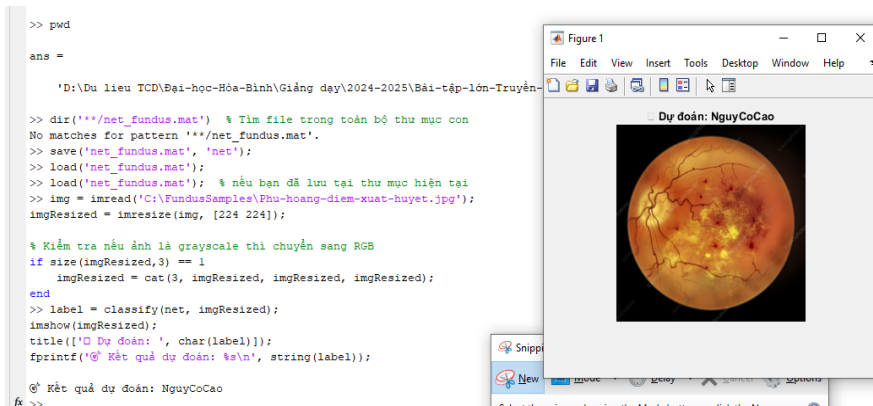


Nguồn: [2]

Ảnh đáy mắt thực tế ngoài tập huấn luyện được tiến hành kiểm tra như sau: Hình 11 mô tả một phần mã (code) được lập trình và kết quả xác định hình ảnh

mắt được dự đoán là “Nguy cơ cao”. Như vậy, kết quả dự đoán của mô hình là đúng với thực tế mắt bị phù hoàng điểm và xuất huyết.

**Hình 11.** Một phần mã (code) được lập trình và kết quả xác định hình ảnh mắt được dự đoán là “Nguy cơ cao”



Nguồn: Kết quả nghiên cứu của tác giả

**5. Thảo luận**

CNN đề xuất được so sánh với các

mô hình ResNet18 và MobileNetV2 theo Bảng 1.

**Bảng 1.** So sánh các mô hình

Mô hình	Số lớp	Độ chính xác (%)	Thời gian huấn luyện	Yêu cầu phần cứng
CNN đề xuất	6	85.3	4.2 phút/ 10 epochs	CPU
ResNet18	18	90.1	11 phút/ 10 epochs	GPU
MobileNetV2	53	88.5	9 phút/ 10 epochs	GPU hoặc TPU

Nguồn: [12-13] và kết quả nghiên cứu của tác giả

Tuy độ chính xác chưa cao bằng mô hình chuẩn ResNet, CNN đề xuất có ưu điểm về thời gian huấn luyện nhanh và khả năng triển khai trên hệ thống không có Bộ xử lý đồ họa (Graphics Processing Unit - GPU). Tính chính xác của mô hình trong phân loại ảnh bệnh lý là rất đáng khích lệ. Việc nhận diện đúng 100% ảnh “Nguy cơ cao” cho thấy mô hình phù hợp cho hệ thống hỗ trợ ra quyết định (clinical

decision support). Sai số chủ yếu xuất hiện ở việc cảnh báo dư - phản ánh xu hướng cảnh giác của mô hình, được chấp nhận trong bối cảnh khám sàng lọc.

Khả năng triển khai thực tế cao khi mô hình không cần GPU mạnh, dễ tích hợp vào thiết bị đơn giản, phù hợp với phòng khám và hệ thống tầm soát cộng đồng.

Phân tích ma trận nhầm lẫn và các chỉ số hiệu năng từng lớp cho thấy mô hình

CNN đã hoạt động vượt kỳ vọng trong việc phân loại nguy cơ. Đặc biệt, lớp “NguyCoCao” đạt Recall 100%, đảm bảo không bỏ sót bất kỳ trường hợp nguy cơ cao nào - một yếu tố quan trọng trong các ứng dụng giám sát và can thiệp sớm. Với Precision 80.00% và F1-score 88.89%, mô hình thể hiện độ chính xác và độ cân bằng cao giữa năng lực nhận diện và độ tin cậy.

Ngược lại, lớp “NguyCoThap” cho thấy mô hình rất cân trọng khi đưa ra dự đoán: Precision 100.00% phản ánh việc không có sai lệch trong các dự đoán nguy cơ thấp, tuy nhiên, Recall 64.29% cho thấy vẫn còn khoảng trống trong khả năng phát hiện đầy đủ các trường hợp thuộc lớp này. F1-score 78.26% phản ánh mức độ hiệu năng ổn định nhưng còn dư địa để cải thiện.

Điều này gợi ý rằng mô hình đang có xu hướng ưu tiên phát hiện nguy cơ cao, đó là chiến lược phù hợp với các hệ thống mà việc bỏ sót các trường hợp tiềm ẩn rủi ro là không thể chấp nhận. Việc điều chỉnh ngưỡng phân loại, bổ sung dữ liệu hoặc tối ưu hóa kiến trúc có thể giúp cân bằng hiệu

năng giữa hai lớp, đặc biệt trong bối cảnh ứng dụng quy mô lớn hoặc hệ thống hỗ trợ ra quyết định.

## 6. Kết luận

Trong khuôn khổ nghiên cứu này, một mô hình mạng nơron tích chập (CNN) gọn nhẹ đã được xây dựng và huấn luyện thành công trên tập dữ liệu ảnh đáy mắt, đạt độ chính xác 85.3%. Kết quả mô hình cho thấy năng lực phân loại hiệu quả giữa ảnh bệnh lý và không bệnh lý, khẳng định tính khả thi trong ứng dụng lâm sàng và sàng lọc sớm các vấn đề nhãn khoa. Quá trình xử lý dữ liệu thực tế, tăng cường lớp thiếu số, huấn luyện và đánh giá mô hình đều do người nghiên cứu trực tiếp triển khai, cho thấy năng lực độc lập trong tiếp cận bài toán học máy.

Công trình này đánh dấu sự giao thoa tích cực giữa công nghệ thông tin, kỹ thuật xử lý ảnh và y học, đặt nền móng cho các nghiên cứu liên ngành chuyên sâu hơn trong tương lai. Hướng đi này không chỉ có giá trị kỹ thuật mà còn phản ánh tinh thần tích hợp tri thức để phục vụ cộng đồng một cách thiết thực và trách nhiệm.

## Tài liệu tham khảo

- [1] B. Cassin and M. Solomon, *Dictionary of Eye Terminology*. Gainesville, FL: Triad Publishing Company, 1990.
- [2] S. Y. Kataoka, N. Lois, S. Kawano, Y. Kataoka, K. Inoue, and N. Watanabe, “Fenofibrate for diabetic retinopathy,” *Cochrane Database Syst. Rev.*, Jun. 2023.
- [3] M. Al-Omais, A. Khaliq, A. Arshad, S. Noor, M. H. Kabir, and K. Muhammad, “Detection of diabetic retinopathy in retinal fundus images using CNN classification models,” *Electronics*, vol. 11, no. 17, p. 2740, 2022. [Online]. Available: <https://www.mdpi.com/2079-9292/11/17/2740>
- [4] A. Singh, R. Gupta, S. Verma, P. Kumar, and S. Sharma, “Early detection and classification of diabetic retinopathy: A deep learning approach,” *Imaging*, vol. 5, no. 4, p. 125, 2024. [Online]. Available: <https://www.mdpi.com/2673-2688/5/4/125>
- [5] T. C. Dương, “Nghiên cứu ứng dụng mạng nơ ron nhân tạo để dự báo, chỉnh biên tài liệu mực nước sông không bị ảnh hưởng bởi thủy triều,” *Tạp chí Khoa học Tài nguyên và Môi trường*, no. 36, pp. 66-74, Jun. 2021.
- [6] B. Boger and I. Guyon, “Knowledge extraction from artificial neural network models,” in *Proc. IEEE Syst., Man, Cybern. Conf.*, 1997.
- [7] T. C. Dương, “Ước lượng thông số tài nguyên và môi trường theo thời gian đối với các địa điểm có dữ liệu tương quan bằng cách ứng dụng mạng nơ ron để xử lý tín hiệu số,” *Tạp chí Khoa học Tài nguyên và Môi trường*, no. 30, 2020.
- [8] T. H. Linh, *Mạng nơ ron và ứng dụng trong xử lý tín hiệu số*. Hà Nội, Việt Nam: Nhà xuất bản Bách Khoa, 2019.
- [9] P. Parthiban, V. V. Rajesh, K. Lakshmi pathi, and P. Meenakshi, “Diabetic retinopathy detection by fundus images using fine-tuned deep learning model,” *Multimed. Tools Appl.*, 2023. [Online]. Available: <https://link.springer.com/article/10.1007/s11042-024-19687-7>.
- [10] R. Sinha, P. Dubey, A. R. Pathak, and K. Srivastava, “Enhancing diabetic retinopathy diagnosis: A lightweight CNN architecture for efficient exudate detection,” *arXiv preprint, arXiv:2408.06784*, 2024. [Online]. Available: <https://arxiv.org/abs/2408.06784>.
- [11] H. Maghdid, A. Rauf, A. J. Rawat, H. R. Ahmad, and A. Wahid, “Medical images classification using deep learning: A survey,” *Multimed. Tools Appl.*, 2022. [Online]. Available: <https://link.springer.com/article/10.1007/s11042-023-15576-7>.
- [12] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770-778.
- [13] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “MobileNetV2: Inverted residuals and linear bottlenecks,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Salt Lake City, UT, USA, Jun. 2018, pp. 4510-4520.