

AN EVALUATION OF EYE DISEASES CLASSIFICATION USING RESNET ON FUNDUS IMAGE DATASET COLLECTED FROM THAI BINH HOSPITAL

Vu Huy Luong^{1*}, Nguyen Thi Mai Phuong¹, Pham Van Ngoc¹, Nguyen Van Sinh², Tran Van Canh

¹TNU - University of Information and Communication Technology

²Thai Binh Provincial General Hospital, ³Thai Binh Provincial Eye Hospital

ARTICLE INFO		ABSTRACT
Received:	01/4/2023	The difference in the input retinal image size, the influence of the disease labelling process, the number of disease labels, and the disease recognition signs of the left and right retinal images are evaluated based on the ResNet model which has a depth of 101, 2048 channels. The seven datasets from Set-1 to Set-7 are created from the original dataset of 5000 fundus images with different labelling and preprocessing of the input images. The results show that the larger the input image is, the more accurate the results are and the signs of eye disease identification in both eyes are the same. In other words, the training dataset of the left eye can be used to classify diseases of the right eye. The results also show that the data set (Set-6), which contains 12 types of labels, is the most balanced and gives the most accurate classification results at 98.08%.
Revised:	05/5/2023	
Published:	08/5/2023	
KEYWORDS		
Artificial intelligence		
Convolutional neural network		
Machine learning		
Deep learning		
Fundus image		

ĐÁNH GIÁ PHÂN LOẠI CÁC BỆNH VỀ MẮT BẰNG RESNET TRÊN BỘ DỮ LIỆU HÌNH ẢNH CHỤP VÒNG MẠC THU THẬP TỪ BỆNH VIỆN THÁI BÌNH

Vũ Huy Lượng^{1*}, Nguyễn Thị Mai Phương¹, Phạm Văn Ngọc¹, Nguyễn Văn Sinh², Trần Văn Cảnh³

¹Trường Đại học Công nghệ Thông tin và Truyền thông - ĐHTT Thái Nguyên

²Bệnh viện đa khoa tỉnh Thái Bình, ³Bệnh viện mắt Thái Bình

THÔNG TIN BÀI BÁO		TÓM TẮT
Ngày nhận bài:	01/4/2023	Ảnh hưởng của kích thước ảnh võng mạc đầu vào, quá trình gán nhãn bệnh, số lượng nhãn bệnh, dấu hiệu nhận biết bệnh trên ảnh võng mạc mắt trái và mắt phải tới độ chính xác của việc chẩn đoán các bệnh về mắt được nghiên cứu và đánh giá trên mô hình học sâu ResNet với 101 lớp nơ-ron ẩn, 2048 kênh. Bảy bộ dữ liệu từ Set-1 đến Set-7 được tạo ra từ bộ dữ liệu gốc gồm 5000 ảnh chụp đáy mắt với cách gán nhãn và tiền xử lý ảnh đầu vào khác nhau. Kết quả cho thấy rằng, ảnh đầu vào càng lớn thì cho kết quả càng chính xác và các dấu hiệu nhận diện các bệnh mắt ở hai bên mắt là như nhau, hay nói cách khác có thể dùng bộ dữ liệu huấn luyện của mắt trái để phân loại bệnh ở mắt phải. Kết quả cũng cho thấy rằng, bộ dữ liệu có độ cân bằng nhất (Set-6) với 12 loại nhãn cho kết quả phân loại chính xác nhất là 98,08%.
Ngày hoàn thiện:	05/5/2023	
Ngày đăng:	08/5/2023	
TỪ KHÓA		
Trí tuệ nhân tạo		
Mạng nơ-ron tích chập		
Học máy		
Học sâu		
Ảnh võng mạc		

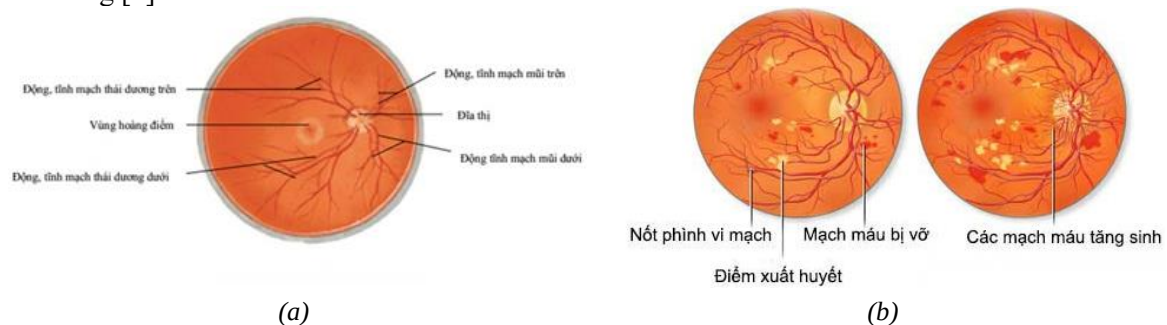
DOI: <https://doi.org/10.34238/tnu-jst.7644>

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

1. Giới thiệu

1.1. Bối cảnh nghiên cứu

Theo tổ chức Y tế thế giới, có 2,2 tỉ người không nhìn thấy ánh sáng hoặc mất thị lực, trong đó ít nhất 1 tỉ người bị giảm thị lực có khả năng phục hồi. Phát hiện sớm các bệnh về mắt, khả năng phục hồi khi điều trị càng cao [1]. Sàng lọc bệnh sớm giúp tăng cường cơ hội điều trị kịp thời và ngăn ngừa suy giảm thị lực vĩnh viễn. Hình ảnh đáy võng mạc là một nguồn hữu ích để chẩn đoán các biến chứng võng mạc cho các bác sĩ nhãn khoa. Tuy nhiên, việc phát hiện bệnh về mắt bằng bác sĩ chẩn đoán hình ảnh có thể tốn nhiều công sức và thời gian. Sự thiếu hụt nguồn nhân lực y tế nói chung và nguồn nhân lực y tế chất lượng cao nói riêng so với nhu cầu thực tế đòi hỏi phải có giải pháp công nghệ hỗ trợ nhân viên y tế trong chẩn đoán và ra quyết định cận lâm sàng [2].



Hình 1. Ảnh chụp võng mạc (a) Võng mạc bình thường, (b) Võng mạc đái tháo đường

Võng mạc là một mô mỏng, mảnh và trong suốt được phát triển từ ngoại bì thần kinh (biểu mô ống thần kinh) [3]. Nó là mô cảm nhận ánh sáng nằm ở phía sau của mắt. Chụp ảnh võng mạc là một phương pháp chẩn đoán không xâm lấn nhằm xác định các vấn đề về mắt như bệnh võng mạc đái tháo đường, viêm mạc tiểu đường, bệnh võng mạc do cao huyết áp, bệnh thoái hóa hắc mạc liên quan đến tuổi tác và bệnh tăng nhãn áp. Chụp ảnh võng mạc sẽ tạo ra các hình ảnh cho phép quan sát võng mạc. Hình 1(a) mô tả ảnh chụp võng mạc của người bình thường và Hình 1(b) mô tả ảnh chụp võng mạc của người bị bệnh đái tháo đường.

1.2. Các nghiên cứu liên quan

Gần đây, việc sử dụng Học sâu để phân loại các hình ảnh đáy mắt thành hai loại (khỏe mạnh và bị bệnh) đã đạt được những kết quả khả quan. Tuy nhiên, việc phân loại các bệnh nhẹ và phân loại đa lớp vẫn là một thách thức mở, do đó Sarki và cộng sự trong [1] đã xây dựng một hệ thống phân loại cho hai kịch bản: (i) bệnh mắt đái tháo đường đa lớp dạng nhẹ, (ii) bệnh mắt đái tháo đường đa lớp. Phương pháp sử dụng hai mạng nơ-ron tích chập là VGG16 và InceptionV3. Mô hình đã được thử nghiệm trên nhiều bộ dữ liệu khác nhau. Độ chính xác cao nhất của mô hình là 88,3% cho VGG16 đối với phân loại đa lớp và 85,95% đối với phân loại đa lớp dạng nhẹ.

Trong nghiên cứu [4], mô hình được đề xuất gồm 2 lớp tích chập. Lớp đầu tiên sử dụng 32 bộ lọc để chiết xuất 32 đặc điểm khác nhau cùng với một hàm max pooling để giảm kích thước kết quả đầu ra. Lớp tích chập thứ hai gồm 64 bộ lọc và một hàm max-pooling. Kết quả thực nghiệm cho độ chính xác cao nhất là 99,89%. Tác giả trong [5] kết hợp ba mô hình CNN, Vgg16 và InceptionV3 thành một mô hình tích hợp và đánh giá hiệu năng của mô hình này so với ba mô hình đơn lẻ trong việc áp dụng vào bài toán chẩn đoán 03 bệnh về mắt. Nghiên cứu này chỉ ra rằng nên sử dụng các mô hình học sâu như AlexNet hay ResNet kết hợp với InceptionV3 để cho kết quả tốt hơn.

Để giảm độ phức tạp và tăng độ chính xác cho mô hình CNN trong việc chẩn đoán các bệnh lý về mắt thì Prasad và cộng sự trong [6] tích hợp tầng dropout và tầng dense cùng với tầng conv2D. Mô hình trong nghiên cứu này dùng mô hình CNN không phức tạp (5 lớp) và cung cấp

độ chính xác là 80%. Các tác giả trong [7] giới thiệu một mô hình phân loại và chẩn đoán các bệnh lý về mắt dựa trên sự kết hợp giữa CNN và RNN. Các ưu điểm của Transfer learning và Long Short-Term Memory (LSTM) đã được kết hợp trong các mô hình lai CNN-RNN đề xuất giúp cho mô hình này hoạt động hiệu quả trong việc phân loại các loại bệnh khác nhau. Một mô hình phân loại và chẩn đoán bệnh về mắt DCNet được đề xuất bởi Li và cộng sự trong [8] dựa trên sự kết hợp giữa CNN, mô-đun tương quan không gian SCM và một bộ phân loại. Các thử nghiệm trên các bộ dữ liệu công khai cho thấy DCNet đạt được hiệu suất tốt hơn so với các mô hình CNN cơ bản khác như ResNet.

Liu và cộng sự trong [9] đề xuất một mô hình CNN để chẩn đoán bệnh khô mắt. Mô hình đề xuất có thể phát hiện ra khu vực bị phá hủy của lớp màng phim nước mắt khi nhuộm huỳnh quang. Các hình ảnh của lớp màng phim nước mắt (tear film) được đưa vào mô hình CNN để đo thời gian phá hủy phim nước mắt khi nhuộm huỳnh quang. Các tác giả trong [10] cũng đề xuất một phương pháp để phát hiện bệnh khô mắt dựa trên thời gian phá hủy phim nước mắt khi nhuộm huỳnh quang (Tear film breakup time - TBUT). Một mô hình CNN đã được giới thiệu để tự động chiết xuất các đặc điểm liên quan từ các khung hình trong TBUT video.

Berrimi và cộng sự trong [11] giới thiệu một mô hình phân loại các bệnh về mắt dựa trên học sâu và transfer learning. Một mô hình CNN gồm 3 lớp tích chập, theo sau đó là lớp Max pooling, lớp Dense, và lớp đầu ra được đề xuất. Mô hình này được so sánh hiệu năng với các mô hình đã được huấn luyện trước đó như InceptionV3 và VGG16. Mô hình đề xuất, mô hình Inception V3 và VGG-16 có độ chính xác lần lượt là 98,5%, 99,27% và 93,5% trên bộ dữ liệu kiểm tra. Các tác giả trong [12] đã đề xuất một phương pháp tương tự để giải quyết bài toán. Mô hình CNN được đề xuất bao gồm 5 lớp ẩn, mỗi lớp gồm hai lớp tích chập kết hợp với một lớp Max Pooling và theo sau nó là bộ tối ưu hóa Adam. Độ chính xác của mô hình đề xuất là 96,5%.

Một vấn đề có thể gặp phải khi làm việc với một dự án đòi hỏi số lượng lớp phân loại lớn là số lượng ảnh trong mỗi lớp không đều nhau. Bằng cách kết hợp 3 bộ dữ liệu khác nhau chứa hình ảnh của 12 loại bệnh về mắt, các tác giả trong [13] đã thu được kết quả khả quan. Mô hình CNN đề xuất trong nghiên cứu này gồm có 5 khối, mỗi khối gồm các lớp tích chập theo sau bởi một lớp maxpool. Mô hình có độ chính xác là 95% trong giai đoạn huấn luyện. Mô hình đã được thử nghiệm trên hình ảnh đáy mắt được ngay sau khi chụp từ bệnh nhân, đạt độ chính xác 92%. Mô hình được đánh giá bằng cách sử dụng nhiều số liệu đánh giá như Độ chính xác, TPR, Độ đặc hiệu FPR, F1-score cho các lớp đầu ra riêng lẻ. Tuy nhiên, nghiên cứu này chưa đánh giá tham số độ nhạy, đây là tham số rất quan trọng trong lĩnh vực y tế.

Các tác giả trong nghiên cứu [14] giới thiệu mô hình lai giữa CNN với Transfer Learning để hỗ trợ chẩn đoán bệnh võng mạc đáy mắt thoái hóa (DR). Nghiên cứu này không chỉ tập trung vào việc phát hiện bệnh DR mà còn phân tích các giai đoạn DR khác nhau. Các kiến trúc học sâu được đề xuất như CNN, CNN lai với ResNet, CNN lai với DenseNet 2.1 được sử dụng để chiết xuất các đặc điểm của mắt nhằm phân loại hiệu quả. Các mô hình đạt độ chính xác lần lượt là 96,22%, 93,18% và 75,61%.

Các nghiên cứu hiện tại mới chỉ tập trung vào việc sử dụng các hình ảnh để chẩn đoán bệnh mà không sử dụng thông tin tương quan hữu ích giữa mắt trái và mắt phải. Ngoài ra, các nghiên cứu này thường chỉ tập trung vào một hoặc một số loại bệnh về mắt tại một thời điểm. Do đó [15] đã đề xuất một mô hình CNN, cụ thể là DCNet, để chẩn đoán và phân loại các bệnh lý về mắt dựa trên cả hai mắt trái-phải của bệnh nhân. Kết quả thực nghiệm cho thấy mô hình có khả năng cho kết quả khả quan so với các phương pháp khác.

Các nghiên cứu trên đã nêu ra các phương pháp sử dụng học sâu để phân loại các bệnh về mắt. Tuy nhiên chưa có nghiên cứu nào so sánh đánh giá về mức độ chính xác của mô hình học sâu với sự thay đổi khác nhau của kích thước ảnh của bộ dữ liệu, sự ảnh hưởng quá trình gán nhãn, số lượng nhãn, sự khác nhau trong việc phân loại bệnh theo bên mắt trái và mắt phải sử dụng cùng một bộ dữ liệu gốc. Trong nghiên cứu này, MMClassification là một phần của bộ công cụ OpenMMLab sẽ được sử dụng để đánh giá sự ảnh hưởng của kích thước ảnh, cách gán nhãn

dữ liệu ảnh hưởng như thế nào đến độ chính xác của mô hình học sâu. Mạng học sâu được lựa chọn để đánh giá trong nghiên cứu này là mạng Residual Learning (ResNet) [16].

2. Phương pháp nghiên cứu

2.1. Bộ dữ liệu và phương pháp gán nhãn dữ liệu

Bộ dữ liệu sử dụng trong nghiên cứu này gồm tối đa 94 loại nhãn ảnh được tạo ra từ 5000 ảnh được cung cấp bởi khoa Mắt – Bệnh viện Đa khoa tỉnh Thái Bình. Bộ dữ liệu được chụp và gán nhãn bởi các bác sĩ chuyên khoa cấp 1 và cấp 2 công tác tại khoa Mắt của Bệnh viện qua các năm. Trước khi được phép sử dụng cho mục đích nghiên cứu, tác giả phải đảm bảo không thu thập thông tin định danh của bệnh nhân. Dữ liệu được thu thập từ cả chụp đáy mắt không huỳnh quang và chụp mạch huỳnh quang với nhiều độ phân giải, góc chụp khác nhau trên máy VISUACAM của ZEISS.

Bảng 1. Ảnh chụp võng mạc gốc có độ phân giải 3000x3000 sẽ được tiền xử lý thay đổi kích cỡ trước khi thực hiện cắt và gán nhãn tạo ra các bộ dữ liệu dùng để học và kiểm thử mô hình học sâu đã học, các bộ dữ liệu được tạo ra đều là các bộ dữ liệu không đối xứng

Bộ dữ liệu	Số lượng nhãn	Kích thước ảnh	Ảnh được cắt ngẫu nhiên	Trái vs Phải	Nhãn đa lớp
Set-1	94	256x256	224x224	X	X
Set-2	72	256x256	224x224		X
Set-3	72	720x720	710x710		X
Set-4	72	900x900	880x880		X
Set-5	25	900x900	880x880		X
Set-6	12	900x900	880x880		
Set-7	12	>900x900			

Do ảnh được lấy trong cơ sở dữ liệu có sẵn của máy chụp cũng như đặc điểm cơ cấu bệnh tật chứ không thông qua chọn lọc loại bệnh cụ thể nên có sự không đồng đều về số lượng ảnh của mỗi nhãn bệnh. Bộ dữ liệu thuộc loại có mẫu không đối xứng. Cụ thể số mẫu ảnh của các nhãn bệnh khác nhau là không đồng đều, những bệnh hiếm gặp sẽ có số lượng mẫu ít hoặc rất ít. Số ảnh chụp được gán nhãn không mắc bệnh chiếm tỉ lệ lớn nhất.

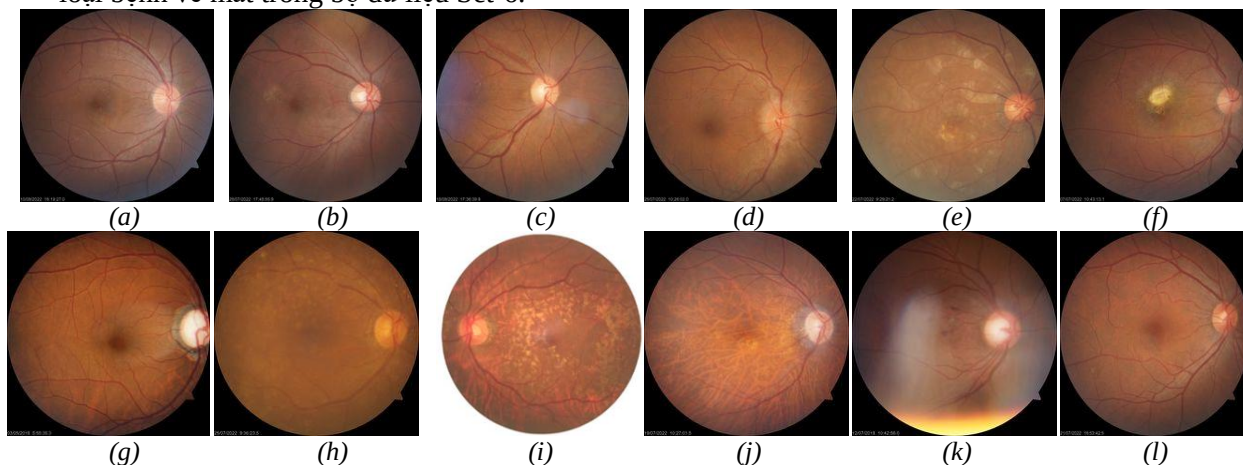
Bộ dữ liệu nguyên bản gồm các ảnh lớn (kích thước 3000x3000), có độ phân giải cao, sẽ được tiền xử lý bằng cách giảm kích thước ảnh về ba cỡ 256x256, 720x720 và 900x900. Sau đó ảnh sẽ được cắt ngẫu nhiên thành các ảnh 224x224, 710x710 và 880x880. Set-1 và Set-2 là hai bộ dữ liệu có ảnh giống nhau nhưng khác nhau cách gán nhãn. Set-1 có phân biệt nhãn mắt trái và mắt phải, Set-2 không phân biệt nhãn mắt trái và mắt phải. Từ Set-3 đến Set-5 sẽ lần lượt loại bỏ những nhãn mà chiếm dưới 10%, 15% và 20% của nhãn có số lượng ảnh lớn nhất (nhãn mắt bình thường). Kích thước ảnh đầu vào từ Set-4 đến Set-6 là 900x900. Set-6 loại bỏ hoàn toàn các nhãn đa lớp. Set-7 có cùng bộ ảnh với Set-6 nhưng có kích thước ảnh đầu vào lớn hơn.

Dữ liệu sẽ được gán nhãn để tạo thành các bộ dữ liệu trong Bảng 1. Mỗi một bộ dữ liệu gán nhãn sẽ được chia ra làm ba phần: Dữ liệu huấn luyện (training) chiếm 70%, dữ liệu đánh giá huấn luyện (validating) chiếm 10% và dữ liệu kiểm tra (testing) chiếm 20% trên tổng số mỗi loại nhãn.

Kết quả dùng để đánh giá độ chính xác của các mô hình học sâu trên các bộ dữ liệu khác nhau (Set-1 tới Set-6) sẽ được tính toán dựa trên phần dữ liệu kiểm tra (testing) của từng bộ dữ liệu. Phần dữ liệu kiểm tra này được thiết kế tách rời và không dùng trong khi thực hiện huấn luyện mô hình học sâu.

Mô hình học sâu sẽ được cho học các bộ dữ liệu huấn luyện và đánh giá lần lượt từ Set-1 đến Set-7 tương đương với bộ dữ liệu không đối xứng với 94 nhãn, 72 nhãn, 25 nhãn, và 12 nhãn. Mô

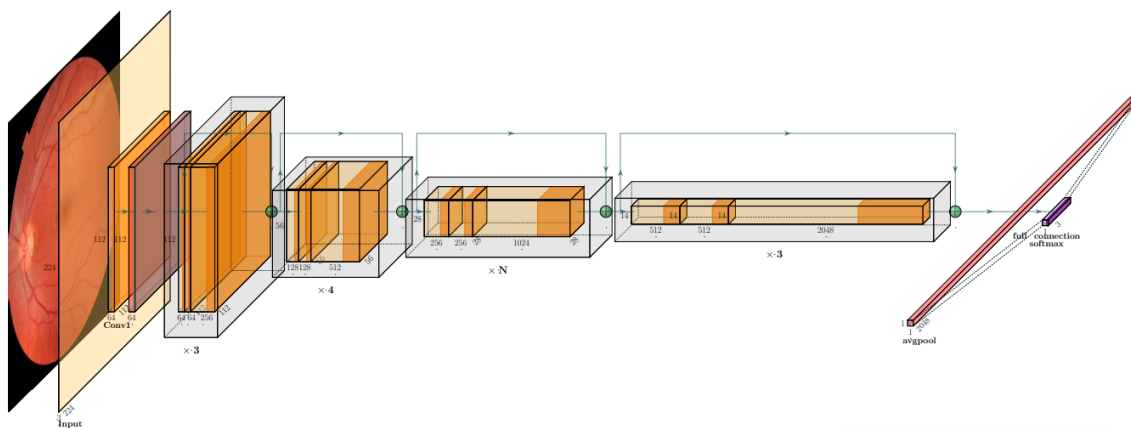
hình thu được sẽ được sử dụng để đánh giá sự khác biệt về độ chính xác trong phân loại các loại bệnh về mắt trên bộ dữ liệu dùng để kiểm tra. Hình 2 mô tả cách đánh 12 nhãn dữ liệu cho các loại bệnh về mắt trong bộ dữ liệu Set-6.



Hình 2. Các nhãn của bộ dữ liệu Set-6. Trong đó, (a) Bìnhthuong, (b) Bongbieubisacto, (c) Mangtruocvongmac, (d) Phugaithi, (e) Phuvongma, (f) Seovongmac, (g) Teogaithi, (h) Thoaihoadrusenhubien, (i) Thoaihoatuoitac, (j) Thoaihoavongmac, (k) Xuathuyetvongmac và (l) Xuattietcong

2.2. Công cụ sử dụng để đánh giá

OpenMMLab xây dựng hệ thống thuật toán thị giác máy tính nguồn mở. Nó nhằm mục đích: cung cấp các thư viện chất lượng cao; giảm bớt khó khăn trong việc triển khai thuật toán; tạo các chuỗi công cụ hiệu quả; xây dựng nền tảng vững chắc cho nghiên cứu và phát triển thị giác máy tính; thu hẹp khoảng cách giữa nghiên cứu học thuật và ứng dụng công nghiệp. Trong nghiên cứu này, MMClassification được sử dụng để phân loại các bệnh về mắt.



Hình 3. Cấu trúc của một mô hình ResNet có số lớp nơ-ron ẩn là 101

MMClassification là một bộ công cụ phân loại hình ảnh mã nguồn mở dựa trên PyTorch, là một phần của dự án OpenMMLab. Mạng học sâu được lựa chọn trong nghiên cứu này là mạng Residual Learning (ResNet) [16]. ResNet được lựa chọn vì nó cho phép sử dụng số lượng lớn các lớp ẩn, có khả năng học sâu với số lượng nhân lớn và nhiều tham số phức tạp. Ngoài ra, các tác giả trong [17] đã tiến hành so sánh 10 mô hình CNN khác nhau và kết quả thực nghiệm cho thấy ResNet-101 cho kết quả tốt nhất. Hình 3 biểu diễn một mô hình mạng học sâu ResNet có số lớp

neuron là 101, số kênh là 2048 và gồm 4 giai đoạn. Bộ dữ liệu ảnh dùng để huấn luyện cho mô hình học sâu ở Hình 3 có kích thước ảnh là 224x224.

Resnet-101 được lựa chọn trong nghiên cứu này thay vì các mô hình Resnet với độ sâu khác bởi vì trong nghiên cứu Resnet trên bộ dữ liệu ảnh ImageNet-1k, bộ dữ liệu bao gồm 1000 lớp khác nhau, mỗi lớp có ít nhất 1000 ảnh (đối tượng trong bộ dữ liệu chủ yếu là động vật và đồ vật), chỉ ra rằng Resnet-101 đồng thời cho kết quả tốt nhất so sánh với các Resnet có độ sâu khác [16].

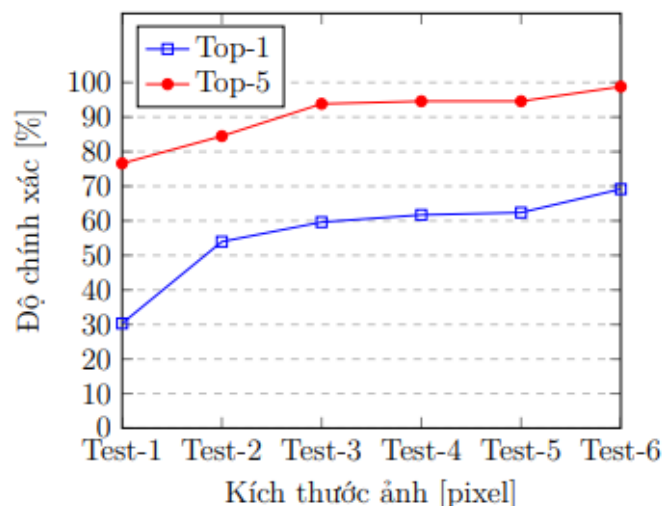
3. Kết quả nghiên cứu

Mô hình học sâu sử dụng (ResNet) được huấn luyện và đánh giá huấn luyện thông qua bảy bộ dữ liệu (từ Set-1 đến Set-7). Các mô hình được huấn luyện trên máy tính chạy hệ điều hành Ubuntu 20.2, cấu hình máy AMD® Ryzen 7 5800 x 8-core processor × 16 4.7GHz, RAM 64Gb, ổ cứng 195TB. Độ chính xác của các mô hình học sâu trên phần dữ liệu kiểm tra (testing) của mỗi bộ dữ liệu từ Set-1 đến Set-7. Phần dữ liệu kiểm tra không tham gia vào quá trình huấn luyện các mô hình học sâu. Kết quả được trình bày trong Bảng 2 và Hình 4.

Bảng 2. Độ chính xác của các mô hình học sâu trên bộ dữ liệu ảnh đầy mắt với các kiểu gân nhãn và tiền xử lý ảnh khác nhau. Kết quả được tính toán dựa trên phần dữ liệu kiểm tra (testing) của mỗi bộ dữ liệu. Phần dữ liệu kiểm tra không tham gia vào quá trình huấn luyện các mô hình học sâu.

Mô hình	Params(M)	Flop(G)	Bộ dữ liệu	Top-1(%)	Top-5(%)
ResNet-101-1	25,56	4,12	Test-1	38,30	76,60
ResNet-101-2	25,56	4,12	Test-2	53,96	84,50
ResNet-101-3	25,56	4,12	Test-3	59,60	93,86
ResNet-101-4	25,56	4,12	Test-4	61,70	94,59
ResNet-101-5	25,56	4,12	Test-5	62,27	94,59
ResNet-101-6	25,56	4,12	Test-6	69,16	98,08
ResNet-101-7	25,56	4,12	Test-7		

Test-1 đến Test-7 là phần dữ liệu kiểm tra (testing) được tách ra từ các bộ dữ liệu tương ứng từ Set-1 đến Set-7. Test-1 đến Test-7 chiếm 20% tổng số dữ liệu trên mỗi nhãn của bộ dữ liệu gốc. Bộ dữ liệu kiểm tra không tham gia vào quá trình huấn luyện mô hình học sâu Resnet-101.



Hình 4. Biểu đồ độ chính xác của các mô hình học sâu Resnet-101 trên bộ dữ liệu ảnh đầy mắt với các kiểu gân nhãn và tiền xử lý ảnh khác nhau. Kết quả được tính toán dựa trên phần dữ liệu kiểm tra (testing) của mỗi bộ dữ liệu. Phần dữ liệu kiểm tra không tham gia vào quá trình huấn luyện các mô hình học sâu.

Số liệu trong Bảng 2 và Hình 4 cho thấy rằng, kích thước ảnh sau tiền xử lý càng lớn thì độ chính xác càng cao, ví dụ Set-2 cho độ chính xác 84,50% trong khi Set-3 cho độ chính xác 93,86%. Qua quá trình thực nghiệm cho thấy chỉ có thể để kích thước ảnh tối đa 900x900 trước khi đưa vào huấn luyện (sử dụng cấu hình nêu ở phần 2.1). Nếu kích cỡ ảnh lớn hơn sẽ gây ra tình trạng thiếu bộ nhớ và không thể tiếp tục huấn luyện.

Kết quả cũng cho thấy Set-1 (gán nhãn mắt trái/phải) cho kết quả 76,60% trong khi Set-2 (không phân biệt mắt trái/phải) cho kết quả là 84,50%. Như vậy, có thể do dấu hiệu bệnh ở hai mắt (nhãn bệnh) là giống nhau, gây ra sự nhầm lẫn trong phân loại nhãn ở Set-1 dẫn đến kết quả kém hơn Set-2.

Quá trình loại bỏ những nhãn có số lượng mẫu bé cho thấy rằng độ cân bằng dữ liệu của mỗi nhãn ảnh hưởng tới độ chính xác của đầu ra. Cụ thể Set-2 có độ chính xác là 84,5% trong khi Set-4, Set-5 và Set-6 đều có độ chính xác lớn hơn 94%. Kết quả thực nghiệm cũng cho thấy rằng gán nhãn kiểu đa nhãn cũng tác động tới kết quả chẩn đoán và phân loại. Cụ thể, Set-5 có độ chính xác là 94,59% trong khi Set-6 có độ chính xác là 98,08%. Kết quả thực nghiệm cho thấy, mô hình cho kết quả khả quan khi phân loại nhiều loại bệnh khác nhau. Các bộ dữ liệu (từ Set-1 tới Set-7) đều có số lượng bệnh lớn hơn hoặc bằng 12, lớn hơn số lượng bệnh phân loại của các công trình nghiên cứu liên quan.

4. Kết luận

Nghiên cứu này đã tiến hành so sánh và đánh giá mức độ chính xác của mô hình học sâu đối với sự thay đổi khác nhau của kích thước ảnh của bộ dữ liệu đầu vào và sự ảnh hưởng của quá trình gán nhãn số lượng nhãn. Mô hình học sâu ResNet với độ sâu (depth) 101 đã được áp dụng để tiến hành các thực nghiệm trên. Kết quả cho thấy rằng ảnh đầu vào càng lớn thì cho kết quả càng chính xác và các dấu hiệu nhận diện bệnh mắt ở hai bên mắt là như nhau. Hay nói cách khác có thể dùng bộ dữ liệu huấn luyện của mắt trái để phân loại bệnh ở mắt phải. Kết quả cũng cho thấy rằng bộ dữ liệu cân bằng nhất (Set-6) với 12 loại nhãn cho kết quả tốt nhất với độ chính xác là 98,08%. Trong các nghiên cứu sau, các công nghệ học sâu khác cũng như các phương pháp tiền xử lý ảnh phức tạp sẽ được nghiên cứu để quá trình phân loại nhãn bệnh có tỉ lệ chính xác cao hơn.

Tác giả đã sử dụng mô hình ResNet-101, mô hình được đánh giá tốt nhất được chỉ ra tại nghiên cứu [16] và [17], để đánh giá sự tác động của kích thước ảnh vồng mạc đầu vào, quá trình gán nhãn bệnh, số lượng nhãn bệnh, dấu hiệu nhận biết bệnh trên ảnh vồng mạc mắt trái và mắt phải tới độ chính xác của việc chẩn đoán các bệnh về mắt. Tác giả sẽ tiến hành đánh giá bộ dữ liệu này trên các mô hình học sâu khác nhau trong các nghiên cứu tiếp theo.

TÀI LIỆU THAM KHẢO/ REFERENCES

- [1] R. Sarki, K. Ahmed, H. Wang, and Y. Zhang, "Automated detection of mild and multi-class diabetic eye diseases using deep learning," *Health Information Science and Systems*, vol. 8, no. 1, pp. 1-9, 2020, doi: 10.1007/s13755-020-00125-5.
- [2] C. Mai, "Together to solve the shortage of medical personnel," (in Vietnamese), Oct. 2022. [Online]. Available: <https://dangcongsan.vn/giao-duc/chung-tay-giai-quyet-tinh-trang-thieu-hut-nhan-su-nganh-y-te-621422.html>. [Accessed April 15, 2023].
- [3] M. P. Gupta, A. A. Herzlich, T. Sauer, and C. C. Chan, "Retinal Anatomy and Pathology," *Developments in Ophthalmology*, vol. 55, pp. 7-17, 2016, doi: 10.1159/000431128.
- [4] O. Bernabé, E. Acevedo, A. Acevedo, R. Carreño, and S. Gómez, "Classification of Eye Diseases in Fundus Images," in *IEEE Access*, vol. 9, pp. 101267-101276, 2021, doi: 10.1109/ACCESS.2021.3094649.
- [5] M. Smaida and S. Yaroshchak, "Bagging of Convolutional Neural Networks for Diagnostic of Eye Diseases," in *International Conference on Computational Linguistics and Intelligent Systems*, vol. 2604, pp. 715-729, 2020.

-
- [6] K. Prasad, P. S. Sajith, M. Neema, L. Madhu, and P. N. Priya, "Multiple eye disease detection using deep neural network," in *TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON)*, 2019, pp. 2148–2153.
 - [7] M. Londhe, "Classification of eye diseases using hybrid CNN-RNN models," MSc Research Project, Data Analytics, School of Computing - National College of Ireland, 2021.
 - [8] C. Li, J. Ye, J. He, S. Wang, Y. Qiao, and L. Gu, "Dense correlation network for automated multi-label ocular disease detection with paired colour fundus photographs," in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, 2020, pp. 1–4.
 - [9] T.-Y. Su, Z.-Y. Liu, and D.-Y. Chen, "Tear film break-up time measurement using deep convolutional neural networks for screening dry eye disease," *IEEE Sensors Journal*, vol. 18, no. 16, pp. 6857–6862, 2018.
 - [10] A. H. Vyas, M. A. Mehta, K. Kotecha, S. Pandya, M. Alazab, and T. R. Gadekallu, "Tear film breakup time-based dry eye disease detection using convolutional neural network," *Neural Computing and Applications*, 2022, doi: <https://doi.org/10.1007/s00521-022-07652-0>.
 - [11] M. Berrimi and A. Moussaoui, "Deep learning for identifying and classifying retinal diseases," in *2020 2nd International Conference on Computer and Information Sciences (ICCIS)*, 2020, pp. 1–6.
 - [12] R. Bhadra and S. Kar, "Retinal disease classification from optical coherence tomographical scans using multilayered convolution neural network," in *2020 IEEE Applied Signal Processing Conference (ASPCON)*, 2020, pp. 212–216.
 - [13] S. Karthikeyan, P. K. Sanjay, R. J. Madhusudan, S. K. Sundaramoorthy, P. K. K. Namboori, "Detection of multi-class retinal diseases using artificial intelligence: An expeditious learning using deep CNN with minimal data," *Biomed Pharmacol J.*, vol. 12 no. 3, pp.1577–1586, 2019.
 - [14] R. Yasashvini, M. S. R. Vergin, R. Panjanathan, G. Jasmine, and L. Anbarasi, "Diabetic retinopathy classification using CNN and hybrid deep convolutional neural networks," *Symmetry*, vol. 14, 2022, Art. no. 1932.
 - [15] J. He, C. Li, J. Ye, Y. Qiao, and L. Gu, "Multi-label ocular disease classification with a dense correlation deep neural network," *Biomedical Signal Processing and Control*, vol. 63, 2021, Art. no. 102167.
 - [16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on computer vision and pattern recognition*, 2016, pp. 770–778.
 - [17] A. A. Ardakani, A. R. Kanafi, U. R. Acharya, N. Khadem, and A. Mohammadi, "Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks," *Computers in Biology and Medicine*, vol. 121, 2020, Art. no. 103795.