

# Phát hiện và phân loại bệnh vông mạc tiểu đường trong ngữ cảnh sử dụng bộ dữ liệu APTOS

1<sup>st</sup> Lìu Thế Hiền

Khoa học dữ liệu (KHDL2021)

Đại học Công nghệ Thông tin

Thành phố Hồ Chí Minh, Việt Nam

21522062@gm.uit.edu.vn

2<sup>nd</sup> Nguyễn Đức Anh

Khoa học dữ liệu (KHDL2021)

Đại học Công nghệ Thông tin

Thành phố Hồ Chí Minh, Việt Nam

21520140@gm.uit.edu.vn

3<sup>rd</sup> Nguyễn Ngọc Yến Nhi

Khoa học dữ liệu (KHDL2021)

Đại học Công nghệ Thông tin

Thành phố Hồ Chí Minh, Việt Nam

21521231@gm.uit.edu.vn

**Tóm tắt nội dung**—Nghiên cứu này tập trung vào việc phát hiện và ngăn ngừa bệnh vông mạc tiểu đường, một trong những nguyên nhân hàng đầu gây mù lòa ở người trưởng thành trong độ tuổi lao động. Mục tiêu của nghiên cứu là phát triển một mô hình học máy để tăng tốc phát hiện bệnh. Các nội dung chính của nghiên cứu bao gồm phân loại mức độ nghiêm trọng của bệnh vông mạc tiểu đường trên từng hình ảnh, xây dựng một mô hình học máy để tự động phát hiện bệnh, đánh giá hiệu suất của mô hình và tối ưu hóa để đạt được kết quả tốt nhất. Nghiên cứu này nhằm cải thiện khả năng chẩn đoán sớm và can thiệp đối với người có nguy cơ mắc bệnh, đặc biệt là trong các khu vực có hạn chế tài nguyên.

**Index Terms**—phát hiện, phân loại, bệnh vông mạc tiểu đường

## I. GIỚI THIỆU

vông mạc tiểu đường là một nguyên nhân quan trọng gây suy giảm thị lực và mù lòa ở người trưởng thành, đặc biệt là những người mắc tiểu đường. Việc phát hiện và ngăn ngừa tình trạng này rất quan trọng để bảo vệ thị lực và đảm bảo sự phát triển toàn diện của cá nhân. Bệnh viện Mắt Aravind ở Ấn Độ nhận thức được tầm quan trọng của việc đối phó với vông mạc tiểu đường trong cộng đồng nông thôn, nơi khả năng tiếp cận dịch vụ y tế hạn chế. Để vượt qua những thách thức trong việc tiến hành kiểm tra mắt toàn diện trong các khu vực này, cần có một hệ thống tự động có khả năng phát hiện và đánh giá mức độ nghiêm trọng của vông mạc tiểu đường một cách hiệu quả.

Mục tiêu của nghiên cứu này là phát triển một mô hình học máy có khả năng nhận dạng và phân loại tự động vông mạc tiểu đường từ hình ảnh mắt. Bằng cách tận dụng sức mạnh của học máy, mô hình nhằm cải thiện khả năng chẩn đoán của Bệnh viện Mắt Aravind và có thể mở rộng ứng dụng của nó đến các chuyên gia chăm sóc mắt trong cộng đồng y tế rộng hơn thông qua Hội nghị Asia Pacific Tele-Ophthalmology Society (APTOS).

Nghiên cứu này tập trung vào ba khía cạnh chính. Thứ nhất, mức độ nghiêm trọng của vông mạc tiểu đường sẽ được phân loại và phân loại trên mỗi hình ảnh mắt, sử dụng một thang đánh giá từ 0 đến 4. Thứ hai, một mô hình học máy sẽ được xây dựng để tự động hóa quá trình phát hiện bệnh, cho phép nhận dạng vông mạc tiểu đường từ hình ảnh mắt nhanh chóng và chính xác. Cuối cùng, hiệu suất của mô hình phát triển sẽ được đánh giá một cách nghiêm ngặt bằng các chỉ số như accuracy, recall và confusion matrix. Hơn nữa, mô hình sẽ được điều chỉnh và tối ưu hóa để đạt được kết quả tốt nhất có thể.

Để đạt được các mục tiêu đã nêu, sẽ áp dụng một số phương pháp. Các kỹ thuật tiền xử lý dữ liệu sẽ được áp dụng để làm sạch và chuẩn hóa dữ liệu hình ảnh mắt, bao gồm cắt và cải thiện chi tiết hình ảnh. Các thuật toán học máy như K-nearest neighbors (KNN [2]), Naïve Bayes, Decision Tree[3] và ensemble models sẽ được khám phá để xây dựng một mô hình phát hiện bệnh hiệu quả. Phương pháp trích xuất đặc trưng dựa trên mạng nơ-ron tích chập (CNN [7]) sẽ được áp dụng để biểu diễn hình ảnh mắt dưới dạng vector đặc trưng. Hiệu suất của mô hình sẽ được đánh giá bằng các chỉ số đánh giá tiêu chuẩn và nỗ lực sẽ được tiến hành để tối ưu hiệu suất bằng cách điều chỉnh tham số và áp dụng các kỹ thuật xử lý tiên tiến. Bằng việc phát triển một mô hình học máy tự động để phát hiện vông mạc tiểu đường, nghiên cứu này nhằm góp phần cải thiện việc chẩn đoán sớm và can thiệp cho những người có nguy cơ mắc bệnh.

## II. BỘ DỮ LIỆU APTOS

### A. Tổng quan

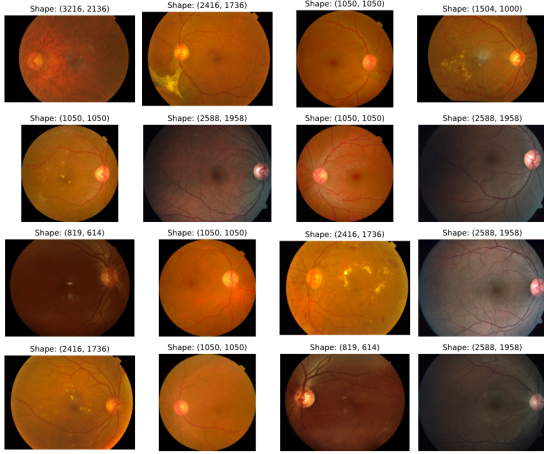
Nghiên cứu này sử dụng một tập dữ liệu gồm 3662 hình ảnh màu, mỗi hình được gán nhãn với mức độ nặng nhẹ của bệnh vông mạc tiểu đường từ 0 đến 4, đại diện cho các mức độ nặng khác nhau. Tuy nhiên, điều quan trọng cần lưu ý là tập dữ liệu này mất cân bằng, với lớp 0 có số lượng trường hợp tương tự như tổng số trường hợp trong tất cả các lớp khác được thể hiện qua Bảng I. Việc resampling dữ liệu không khả thi do tính chất y tế của tập dữ liệu này.

Bảng I  
TỔNG QUAN THỐNG KÊ CỦA BỘ DỮ LIỆU

Phân loại lớp	Số lượng
0	1805
1	370
2	999
3	193
4	295

Các hình ảnh trong tập dữ liệu có định dạng màu và có sự khác nhau về kích thước, tỷ lệ khung hình và phương pháp cắt ảnh như được minh họa trong Hình 1. Ngoài ra, do hình ảnh được thu thập từ nhiều nguồn khác nhau, cũng có sự khác biệt về màu sắc và độ sáng. Chiều cao tối thiểu và chiều rộng tối

thiểu của hình ảnh lần lượt là 358 và 474 điểm ảnh, trong khi chiều cao và chiều rộng tối đa lên đến 2848 điểm ảnh.



Hình 1. Một số thể hiện khác biệt trong kích thước, màu sắc của bộ dữ liệu

Để đảm bảo chất lượng hình ảnh, khuyến khích không thay đổi kích thước dưới 358 điểm ảnh.

### B. Thách thức

Tập dữ liệu đặt ra một số thách thức cần được xem xét cẩn thận trong quá trình phát triển mô hình học máy. Đầu tiên, vấn đề mất cân bằng lớp, với lớp 0 chiếm rất nhiều, có thể gây thiên vị và ảnh hưởng đến khả năng phân loại chính xác các mức độ nặng nhẹ của bệnh vông mạc tiểu đường. Do đó, cần chú trọng đến điểm recall của các lớp trong quá trình huấn luyện và đánh giá.

Thứ hai, sự khác nhau về kích thước hình ảnh, tỷ lệ khung hình và phương pháp cắt ảnh đòi hỏi các bước tiền xử lý cần thận để đảm bảo tính nhất quán và thích hợp cho các mô hình học máy khác nhau. Ta cần áp dụng phương pháp phù hợp để thay đổi kích thước, cắt ảnh và chuẩn hóa hình ảnh trong khi bảo tồn các đặc trưng quan trọng cho việc phát hiện bệnh. Hơn nữa, sự khác biệt về màu sắc và độ sáng giữa các hình ảnh đưa ra thách thức về việc chuẩn hóa và trích xuất đặc trưng. Cần áp dụng các phương pháp điều chỉnh độ tương phản và cải thiện hình ảnh để tăng độ chi tiết của các đặc trưng để cải thiện hiệu suất của mô hình học máy.

Cuối cùng, kích thước lớn của hình ảnh đòi hỏi các kỹ thuật xử lý hiệu quả để đảm bảo khả năng tính toán.

Giải quyết các thách thức này trong các giai đoạn tiền xử lý và mô hình hóa sẽ là yếu tố quan trọng để phát triển một mô hình học máy mạnh mẽ và chính xác cho việc phát hiện và phân loại mức độ nặng nhẹ của bệnh vông mạc tiểu đường.

## III. NGUYÊN CỨU PHƯƠNG PHÁP LÝ THUYẾT

### A. Tiền xử lý dữ liệu

Nhằm đưa dữ liệu về một tập dữ liệu có cấu trúc. Nhóm nghiên cứu sẽ thay đổi kích thước hình ảnh thành kích thước cố định 300x300, đồng thời chuyển đổi chúng sang ảnh xám. Sau đó, các hình ảnh sẽ được làm phẳng, tạo ra một tập dữ liệu với định dạng bảng quen thuộc gồm 3662 hàng và 90000 cột thuộc

tính. Mỗi cột thuộc tính tương ứng với một điểm ảnh trong hình ảnh.

Vì giá trị điểm ảnh có khoảng từ 0 đến 255, các tính toán liên quan đến những giá trị lớn này có thể làm chậm các thuật toán học máy. Để giảm thiểu điều này, nhóm nghiên cứu sẽ thực hiện chuẩn hóa dữ liệu, giảm phạm vi giá trị điểm ảnh trong khi vẫn duy trì tỷ lệ với dữ liệu gốc. Bước chuẩn hóa này giúp cải thiện hiệu suất tính toán của các thuật toán học máy.

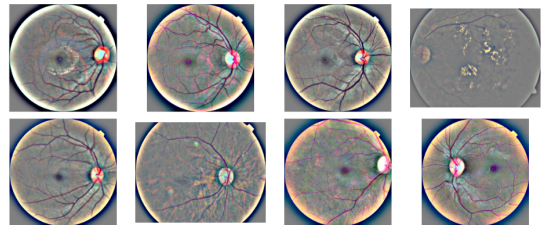
Bảng II  
SỐ LƯỢNG CÁC LỚP TRÊN TẬP HUẤN LUYỆN VÀ ĐÁNH GIÁ

Diagnosis	Train	Diagnosis	Test
	Count		Count
0	1265	0	540
1	272	1	98
2	697	2	302
3	138	3	55
4	191	4	104

Bảng II thể hiện sự phân chia tập dữ liệu, cụ thể tập dữ liệu sẽ được chia thành hai tập: tập huấn luyện và tập đánh giá, sử dụng tỷ lệ chia lần lượt là 70-30. Ngoài ra, tham số ngẫu nhiên 42 được sử dụng để đảm bảo sự chia dữ liệu cố định qua các lần thử khác nhau.

Ngoài các bước tiền xử lý, nhóm nghiên cứu sẽ áp dụng một số phương pháp tiền xử lý đã được sử dụng trong bảng xếp hạng của các cuộc thi trước đây. Các phương pháp này bao gồm:

- **Circle crop** Kỹ thuật này liên quan đến việc cắt bỏ các vùng tối không cần thiết trong hình ảnh, đảm bảo rằng hình ảnh thay đổi kích thước không bị méo hoặc bị biến dạng.
- **Phương pháp tiền xử lý của Ben Graham's:** Phương pháp này, được giới thiệu bởi Ben Graham trong một cuộc thi trước đây, nhằm cải thiện điều kiện ánh sáng và nhấn mạnh chi tiết trong hình ảnh như minh họa trong hình 2. Nó giúp làm nổi bật các đặc trưng quan trọng. Vì phương pháp tiền xử lý của Ben Graham đã bao gồm circle crop, nhóm đã quyết định áp dụng cả hai kỹ thuật để đảm bảo tính nhất quán với phương pháp gốc. Tuy nhiên để đảm bảo hiệu suất của bất kỳ kỹ thuật tiền xử lý nào. Nhóm quyết định thử nghiệm song song cùng lúc 3 phương pháp xử lý trong suốt quá trình nghiên cứu. Bao gồm, ảnh chưa qua 2 bước tiền xử lý trên, ảnh đã qua circle crop và ảnh sử dụng kỹ thuật tiền xử lý của Ben Graham.



Hình 2. Một số thể hiện trong kỹ thuật Ben Graham

## B. Mô hình đề xuất

Do đây là một dạng bài toán phân loại đa lớp nên nhóm sẽ ưu tiên sử dụng một số thuật toán phân loại đa lớp như SVC [1], Gaussian [4], KNN [2], AdaBoost [6], DecisionTree, RandomForest [5]. Đây đều là những mô hình học máy cho các kết quả rất tốt trên các mô hình đa lớp.

**SVC (Support Vector Machine)** là một thuật toán phân loại phổ biến, phân chia các điểm dữ liệu vào các lớp khác nhau bằng cách sử dụng siêu phẳng. Độ chính xác của SVC phụ thuộc vào bộ dữ liệu và lựa chọn các tham số như hàm kernel và tham số điều chuẩn. Việc huấn luyện mô hình có thể tốn nhiều thời gian tính toán, đặc biệt là đối với các bộ dữ liệu lớn. Thời gian dự đoán của mô hình SVC thường nhanh sau khi mô hình đã được huấn luyện. [1]

**Gaussian (Naive Bayes)** là một thuật toán phân loại dựa trên định lý Bayes, giả định rằng các đặc trưng tuân theo phân phối chuẩn. Thuật toán này đơn giản và hiệu quả, đặc biệt là đối với các bộ dữ liệu có số chiều cao. Độ chính xác của Gaussian[4] Naive Bayes có thể cạnh tranh, đặc biệt khi giả định về tính chuẩn hóa đặc trưng đúng. Việc huấn luyện mô hình Gaussian[4] Naive Bayes thường nhanh do tính đơn giản của nó. Thời gian dự đoán của Gaussian[4] Naive Bayes thường nhanh. [4]

**KNN (K-Nearest Neighbors)** là một thuật toán phân loại phi tham số, gán nhãn lớp dựa trên phiếu biểu quyết của các điểm láng giềng. Độ chính xác của KNN có thể bị ảnh hưởng bởi lựa chọn số láng giềng (k) và phương pháp đo khoảng cách sử dụng. Việc huấn luyện mô hình KNN tương đối nhanh do chủ yếu liên quan đến việc lưu trữ dữ liệu huấn luyện. Thời gian dự đoán của KNN có thể chậm hơn so với các mô hình khác do cần tính toán khoảng cách đến từng mẫu huấn luyện. [2]

**AdaBoost (Adaptive Boosting)** AdaBoost là một phương pháp học theo nhóm kết hợp nhiều bộ phân loại yếu thành một bộ phân loại mạnh. Nó lặp đi lặp lại việc huấn luyện các bộ phân loại yếu, tăng cường trọng số cho các mẫu bị phân loại sai ở mỗi vòng lặp. Độ chính xác của AdaBoost có thể cao và nó có xu hướng xử lý tốt các bộ dữ liệu phức tạp. Việc huấn luyện mô hình AdaBoost có thể mất thời gian hơn so với các bộ phân loại đơn lẻ do quá trình lặp. Thời gian dự đoán của AdaBoost thường nhanh sau khi mô hình đã được huấn luyện. [6]

**Decision Tree** là một thuật toán phân loại đơn giản và dễ hiểu, xây dựng một mô hình cây dựa trên việc chia các đặc trưng. Độ chính xác của mô hình Decision Tree phụ thuộc vào các yếu tố như độ sâu của cây và chất lượng các điểm chia. Việc huấn luyện mô hình Decision Tree thường nhanh do chỉ cần tìm các điểm chia tốt nhất tại mỗi nút. Thời gian dự đoán của mô hình DecisionTree nhanh sau khi cây được xây dựng. [3]

**RandomForest** là một phương pháp học theo nhóm kết hợp nhiều cây quyết định để đưa ra dự đoán. Nó xây dựng nhiều cây bằng cách lấy mẫu dữ liệu và ngẫu nhiên chọn các đặc trưng ở mỗi điểm chia. Độ chính xác của RandomForest có thể cao, và nó có xu hướng xử lý nhiễu và việc overfitting tốt hơn so với cây quyết định đơn lẻ. Việc huấn luyện mô hình RandomForest có thể mất thời gian hơn so với cây quyết định đơn lẻ do xây dựng nhiều cây. [5]

Không dừng lại ở đó nhóm còn tiến hành trên các mô hình học sâu, cụ thể là mạng tích chập (CNN [7]). Cụ thể nhóm xây dựng một mô hình tích chập đơn giản như sau:

Layer (type:depth-idx)	Output Shape	Param #
SimpleCNN	[5]	--
Sequential: 1-1	[1, 64, 35, 35]	--
└─Conv2d: 2-1	[1, 64, 298, 298]	1,792
└─BatchNorm2d: 2-2	[1, 64, 298, 298]	128
└─ReLU: 2-3	[1, 64, 298, 298]	--
└─MaxPool2d: 2-4	[1, 64, 149, 149]	--
└─Conv2d: 2-5	[1, 64, 147, 147]	36,928
└─BatchNorm2d: 2-6	[1, 64, 147, 147]	128
└─ReLU: 2-7	[1, 64, 147, 147]	--
└─MaxPool2d: 2-8	[1, 64, 73, 73]	--
└─Conv2d: 2-9	[1, 64, 71, 71]	36,928
└─BatchNorm2d: 2-10	[1, 64, 71, 71]	128
└─ReLU: 2-11	[1, 64, 71, 71]	--
└─MaxPool2d: 2-12	[1, 64, 35, 35]	--
└─AdaptiveAvgPool2d: 1-2	[1, 64, 1, 1]	--
└─Linear: 1-3	[5]	325
-----		
Total params: 76,357		
Trainable params: 76,357		
Non-trainable params: 0		
Total mult-adds (G): 1.14		
-----		
Input size (MB): 1.08		
Forward/backward pass size (MB): 118.22		
Params size (MB): 0.31		
Estimated total size (MB): 119.61		
-----		

Hình 3. Kiến trúc mô hình CNN đơn giản

Khác với các mô hình trên phân loại các điểm ảnh có phần thiên về thống kê và hình học. Mạng tích chập thật sự phân loại các đặc trưng ở trên một bức ảnh để đưa ra kết luận. Sau đó, nhóm tiến hành thử nghiệm trên các mô hình đã được huấn luyện nổi tiếng như EfficientNet B3 và ResNet18.

**ResNet-18** là một kiến trúc mạng neural convolutional (CNN) rất phổ biến và là một trong số các mô hình ResNet (Residual Network). Mô hình được xây dựng dựa trên một ý tưởng cốt lõi là "skip connection", cho phép thông tin truyền qua mạng một cách dễ dàng. Thay vì chỉ sử dụng các lớp convolutional truyền thẳng (feed-forward) để học các đặc trưng từ đầu vào đến đầu ra, ResNet-18 sử dụng các "residual block" để tạo ra các kết nối nhảy qua các lớp. Điều này giúp tránh hiện tượng mất mát thông tin (vanishing gradient) và giúp mạng học các biểu diễn sâu hơn. [9]

**EfficientNet B3** là một kiến trúc mạng neural convolutional (CNN) thuộc họ EfficientNet, được giới thiệu bởi Mingxing Tan và đồng nghiệp vào năm 2019. Kiến trúc EfficientNet được xây dựng dựa trên một ý tưởng chính là tự động điều chỉnh độ phức tạp của mạng để đạt được hiệu suất tốt nhất trong khi giữ nguyên các thông số như kích thước, chiều sâu và chiều rộng của mạng. Điều này được thực hiện thông qua việc sử dụng một siêu tham số gọi là "compound coefficient" để tự động điều chỉnh kích thước của mạng. Điểm mạnh của EfficientNet B3 là sự cân bằng giữa hiệu suất và độ phức tạp của mạng, giúp tối ưu hóa việc huấn luyện và triển khai trên các nền tảng có tài nguyên hạn chế. [8]

## IV. THỰC NGHIỆM

### A. Chuẩn bị tập dữ liệu

Như đã đề cập ở phần phương pháp lý thuyết. Nhóm tiến hành thay đổi kích thước ảnh về 3x300x300 sau đó chuyển ảnh màu thành ảnh xám. Kế đến là làm phẳng các hình ảnh, chia dữ liệu thành tập huấn luyện và tập đánh giá với tỉ lệ 70-30. Và đối với các mô hình học sâu, khác với các mô hình học máy, mô hình học sâu với PyTorch được hỗ trợ chạy trên GPU với tốc

độ rất cao. Lúc này nhóm chỉ tiến hành thay đổi kích thước và giữ nguyên ảnh màu. Mạng tích chập hoạt động trên dữ liệu 2 chiều nên ta không cần làm phẳng. Sau đó, chia dữ liệu thành tập huấn luyện và tập đánh giá với tỉ lệ 70-30. Với các mô hình học sâu nhóm sử dụng Batch Gradient Descent để tăng số lần cập nhật tham số trong 1 epoch cho mô hình, với tham số cụ thể nhóm sử dụng là batch size = 32.

### B. Kết quả thực nghiệm

Trong thời gian huấn luyện được thể hiện ở Bảng III, nhóm phát hiện rằng với một lượng dữ liệu lớn một số mô hình đòi hỏi rất nhiều thời gian chỉ cho 1 lần huấn luyện như AdaBoost [6] (hơn 1h training), DecisionTreeClassifier (gần 12' training), SVM (gần 10' training). Do đây là một bài toán về y tế, nên cả hai độ đo là độ chính xác (độ đo vô cùng quan trọng với các bài toán về lĩnh vực sức khỏe) và tốc độ xử lý của mô hình đều cần phải tập trung quan tâm ở một mức độ cao.

Bảng III  
THỜI GIAN HUẤN LUYỆN CỦA CÁC MÔ HÌNH HỌC MÁY

Model	Raw	Circle Crop	Ben's Processing
SVC	553.63	399.83	582.85
GaussianNB	3.85	2.67	3.61
KNeighborsClassifier	0.49	0.52	0.45
AdaBoostClassifier	3301.03	3622.42	3417.42
DecisionTreeClassifier	651.70	700.65	719.41
RandomForestClassifier	100.13	114.95	124.57

Sau khi nhóm tiến hành huấn luyện mô hình dựa trên kết quả tốt nhất ta thấy được RandomForest kết hợp với bước tiền xử lý Ben Graham cho ra kết quả tốt nhất so với các mô hình còn lại, được thể hiện ở Bảng IV. Các phương pháp tiền xử lý hoạt động tốt với một số mô hình và không hiệu quả với các mô hình khác. Điều này phụ thuộc vào nhiều yếu tố như cách mà mô hình được thiết kế, các yêu cầu đặc biệt của dữ liệu.

Bảng IV  
ĐỘ CHÍNH XÁC TRÊN TẬP ĐÁNH GIÁ CỦA CÁC MÔ HÌNH HỌC MÁY

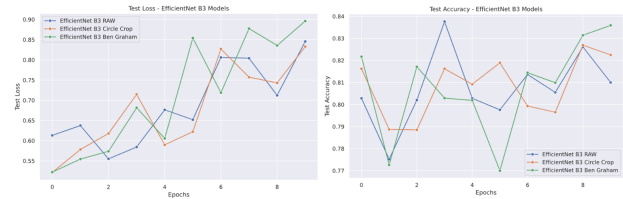
Model	Raw	Circle Crop	Ben's Processing
SVC	0.70	0.71	0.71
GaussianNB	0.52	0.51	0.49
KNeighborsClassifier	0.69	0.67	0.67
AdaBoostClassifier	0.47	0.60	0.65
DecisionTreeClassifier	0.67	0.66	0.63
RandomForestClassifier	0.73	0.73	0.73

Sau khi tiến hành huấn luyện sơ bộ, nhóm tiến hành tinh chỉnh tham số. Tuy nhiên, để giảm chi phí và thời gian tính toán nhóm chỉ tiến hành trên phương pháp có độ chính xác cao nhất ở các mô hình tham số mặc định. Do đó, với mô hình RandomForest [5] nhóm tiến hành tinh chỉnh trên các tham số như: n\_estimators (số cây), max\_depth (độ sâu của cây), criterion (hàm lỗi). Sau khi, tinh chỉnh nhóm thu được một bộ tham số với các tham số như sau: criterion='gini', max\_depth=20, n\_estimators=150. Sau đó thử lại với tập đánh giá độ chính xác lớn hơn xấp xỉ 0.004%.

Bảng V  
SO SÁNH ĐỘ CHÍNH XÁC CỦA CÁC MÔ HÌNH HỌC SÂU

Mô hình	Ảnh chưa xử lý		Circle Crop		Circle Crop và Ben Graham's	
	Train	Test	Train	Test	Train	Test
CNN [7]	0.71	0.72	0.71	0.71	0.76	0.72
ResNet18	0.83	0.78	0.83	0.77	0.82	0.78
EfficientNet B3	0.96	0.82	0.95	0.82	0.97	0.82

Bên cạnh đó ở Bảng V, mô hình CNN cơ bản cho kết quả thậm chí tệ hơn so với một số mô hình học máy trên. Tuy nhiên với các mô hình được huấn luyện sẵn, lại cho ra kết quả vượt trội so với mô hình học máy. Lưu ý, đối với các mô hình đã được huấn luyện sẵn nhóm không sử dụng tham số có sẵn của mô hình. Mà hoàn toàn huấn luyện lại toàn bộ tham số trong mô hình. Do đó, nhóm cũng không cần thiết phải sử dụng lại các bước tiền xử lý của các mô hình đó.



Hình 4. Độ chính xác và độ lỗi trên tập đánh giá của mô hình EfficientNet B3

Ở biểu đồ thể hiện trong Hình 4, với mô hình EfficientNet B3 các phương pháp đều có độ chính xác trong khoảng 80% và tuy chỉ 10 epochs nhưng mô hình loss của mô hình có xu hướng tăng kể từ epoch số 0. Trong đó, Circle Crop mang lại độ ổn định cao nhất.

### C. Phân tích kết quả và bàn luận

**1. Kết hợp hướng pháp tiền xử lý Ben Graham và mô hình RandomForest với tham số criterion=gini, max\_depth=20, n\_estimators=150:** Đối với các bài toán y tế không những độ chính xác mà khi ta xem qua ma trận lỗi phần mà ta tập trung sẽ là phần mà các trường hợp có bệnh nhưng mô hình lại đoán là không có bệnh (False Negative ở class 0 so với các class còn lại). Ở phần này có tổng cộng 67 trường hợp tức 12% (cũng là 1 – precision ở lớp 0) so với toàn bộ các trường hợp có bệnh.

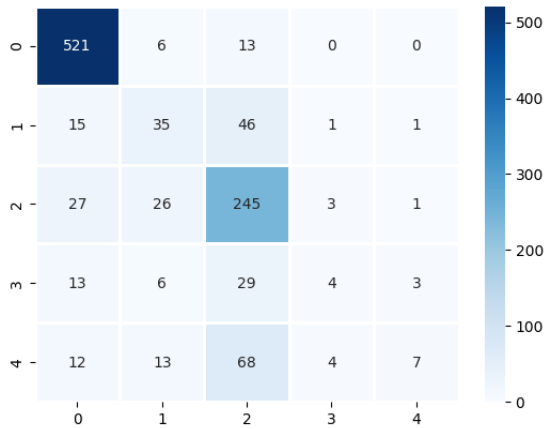
Bảng VI  
KẾT QUẢ TRÊN TẬP ĐÁNH GIÁ CỦA MÔ HÌNH RANDOM FOREST

	precision	recall	f1-score	support
0	0.88	0.97	0.92	540
1	0.37	0.36	0.36	98
2	0.60	0.78	0.68	302
3	0.33	0.07	0.12	55
4	0.64	0.07	0.12	104
accuracy			0.73	1099

Và cũng cần lưu ý rằng 73% accuracy đối với bộ dữ liệu này cũng không phải là một con số lý tưởng vì dữ liệu của ta vốn đã mất cân bằng trên lớp 0 (Nếu ta dự đoán hoàn toàn kết quả



là 0 thì cũng đã đạt được 50% acc), số liệu được đề cập trong Bảng VI và Hình 5.

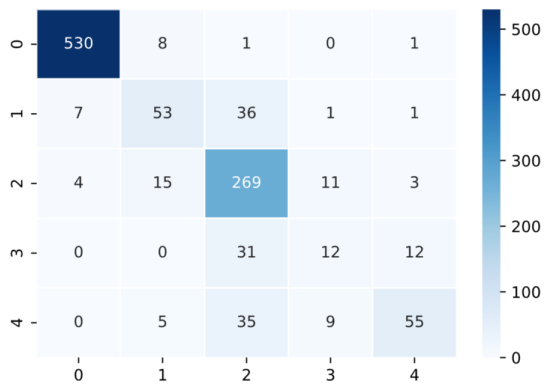


Hình 5. Ma trận lỗi trên tập đánh giá của mô hình Random Forest

**2. EfficientNet B3:** Ở Bảng VII và Hình 6 Với mô hình EfficientNet B3 sử dụng Ben Graham điểm accuracy khá cao 84% còn số False Negative ở class 0 so với các lớp còn lại là 11 trường hợp tức 0.02% trong số các trường hợp bị bệnh thấp hơn 56 trường hợp so với các mô hình học máy. Với ma trận lỗi ta thấy hầu hết các dự đoán sai đều đến từ lớp 2.

Bảng VII  
KẾT QUẢ TRÊN TẬP ĐÁNH GIÁ CỦA MÔ HÌNH EFFICIENTNET B3

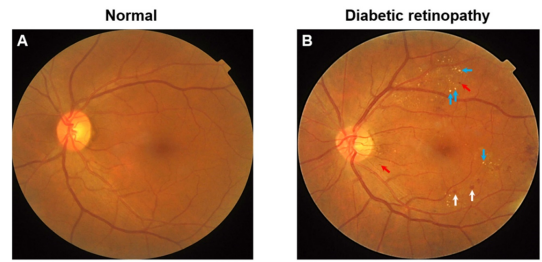
	precision	recall	f1-score	support
0	0.98	0.98	0.98	540
1	0.65	0.54	0.59	98
2	0.72	0.89	0.80	302
3	0.36	0.22	0.27	55
4	0.76	0.53	0.63	104
accuracy			0.84	1099



Hình 6. Ma trận lỗi trên tập đánh giá của mô hình EfficientNet B3

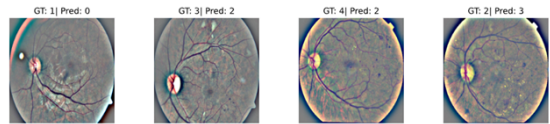
Và để đánh giá hình ảnh ta vẫn cần một vài kiến thức chuyên môn. Nhóm có nghiên cứu được rằng để chuẩn đoán bệnh võng mạc tiểu đường, thì các hình ảnh võng mạc thường có các điểm mạch

máu bị rò rỉ và các điểm bị xuất huyết. Một số còn xuất hiện các dịch vàng như thể hiện ở Hình 7.



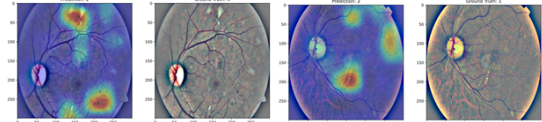
Hình 7. Hình ảnh so sánh giữa võng mạc bình thường và võng mạc bị ảnh hưởng bởi tiểu đường

Từ đó, nếu xem xét qua các hình ảnh được dự đoán sai ở Hình 8. Ngoài các dự đoán sai về mức độ nặng nhẹ của bệnh, ta sẽ quan tâm nhiều về các dự đoán sai về việc có bệnh hay không. Ví dụ ở bức hình thứ nhất, việc tìm ra các điểm máu bị rò rỉ thậm chí còn khó đối với mắt người bình thường.



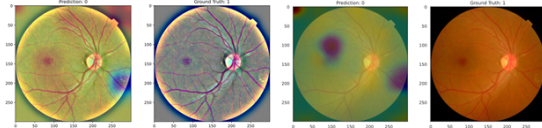
Hình 8. Một số hình ảnh dự đoán sai của mô hình EfficientNet B3 với tiền xử lý Ben Graham

Và một cái hay của các mô hình CNN, là chúng ta hoàn toàn có thể biết được chúng học những gì và từ đâu lại kết luận như vậy. Ví dụ như Hình 9 tuy mô hình dự đoán sai về mức độ bệnh. Nhưng chúng lại hoàn toàn nhìn được các điểm xuất huyết trong võng mạc. Do đó, trước khi ta kết luận cho một bệnh nhân bất kỳ ta có thể tham khảo xem liệu mô hình có tìm thấy bất kỳ điểm máu rò rỉ hay bất thường nào không.



Hình 9. Bản đồ nhiệt cho kết luận của mô hình EfficientNet B3

Thậm chí ở dự đoán sai ở Hình 10, tuy kết luận không có bệnh nhưng mô hình cũng hoàn toàn nhìn thấy được dịch vàng trong võng mạc (Khu vực màu đỏ tương trưng cho việc đây là khu vực ảnh hưởng đến kết luận của mô hình, màu xanh lá cây tượng trưng cho việc không liên quan đến kết luận của mô hình và xanh dương là tương trưng cho việc vùng này chắc chắn không nằm trong lớp được kết luận). Và ở đây ta cũng thấy được vùng màu đỏ ở các góc của ảnh điều này có thể do mô hình học được những thuộc tính không cần thiết của các vùng sáng hoặc một trường hợp tệ hơn đó là một số ảnh có hình tròn hoặc bầu dục có thể ảnh hưởng đến kết luận của mô hình.



Hình 10. Bản đồ nhiệt trong với lớp dự đoán 0 của mô hình EfficientNet B3

## V. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Sau thời gian nghiên cứu và thực hiện nhóm đã hoàn thành được mô hình chuẩn đoán bệnh võng mạc tiểu đường thông qua các ảnh võng mạc, đáp ứng mục tiêu đặt ra ban đầu.

Nhóm cũng đưa ra kết luận rằng sẽ sử dụng mô hình Deep Learning cụ thể là EfficientNet B3[8] để giải quyết bài toán này. Đồng thời, nhóm suy đoán rằng việc áp dụng Machine Learning mang lại kết quả không tốt bằng Deep Learning là do các thuật toán Machine Learning có thiên hướng sử dụng phương pháp thống. Tuy nhiên trong bài toán này một vết xuất huyết nhỏ và các động tĩnh mạch (cái giống sợi gân) có màu sắc gần như không có nhiều sự khác biệt, chưa kể các chi tiết trong ảnh lại vô cùng nhỏ nếu không phải là một mô hình để phục vụ cho hình ảnh thì rất khó để có thể tin tưởng vào mô hình chỉ với những con số accuracy.

Nhóm cũng lên kế hoạch, nhằm khắc phục phần mất cân bằng dữ liệu với augmentation cụ thể là các phương pháp như lật hình xoay hình trên các trường thiếu dữ liệu. Và tiến hành thử nghiệm trên nhiều mô hình SOTA khác trong học máy.

## TÀI LIỆU

- [1] Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning, 20(3), 273-297.
- [2] Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13(1), 21-27.
- [3] Breiman, L., Friedman, J. H., Olshen, R. A., Stone, C. J. (1984). Classification and Regression Trees. CRC Press.
- [4] Domingos, P., & Pazzani, M. (1997). On the optimality of the simple Bayesian classifier under zero-one loss. Machine learning, 29(2-3), 103-130.
- [5] Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.
- [6] Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. Journal of computer and system sciences, 55(1), 119-139.
- [7] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).
- [8] Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In International Conference on Machine Learning (pp. 6105-6114). PMLR.
- [9] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
- [10] Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., & Torralba, A. (2016). Learning deep features for discriminative localization. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2921-2929).
- [11] Rajalakshmi, R., Subashini, R., & Lakshmi Priya, M. (2017). Classification of diabetic retinopathy using deep learning neural networks. In 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT) (pp. 1422-1426). IEEE.
- [12] Ramachandran, N., Hong, S. C., Sime, M. J., Wilson, G. A., Simpson, A. J., & Dickinson, M. E. (2018). Diabetic retinopathy screening using deep neural network. Clinical and experimental ophthalmology, 46(4), 412-416.

Công việc	Nhi	Hiền	Anh
Xác định mục tiêu, phân tích bộ dữ liệu	x	x	x
Tìm hiểu các thuật toán, phương pháp tiền xử lý, cấu trúc mạng phù hợp	x	x	x
Tiền xử lý dữ liệu	x	x	x
Xây dựng mô hình và các thuật toán Machine Learning	x	x	x
Xây dựng Neural Network cho bài toán Deep Learning và sử dụng các mô hình pre-trained		x	x
Đánh giá mô hình và kết luận	x	x	
Viết báo cáo	x	x	x
Thiết kế slide			x
Thuyết trình	x		

Hình 11. Bảng phân công thực hiện đồ án