

XÂY DỰNG HỆ THỐNG HỖ TRỢ VIỆC PHÂN LOẠI BỆNH VIÊM PHỔI BẰNG HÌNH ẢNH X-QUANG NGỰC DỰA TRÊN PHƯƠNG PHÁP HỌC SÂU

Bùi Xuân Tùng^{1*}, Huỳnh Thanh Danh¹, Trịnh Quang Minh¹,
Ngô Thị Lan¹ và Hoàng Văn Tư²

¹Trường Đại học Tây Đô, ²Bệnh viện Đa khoa Hoàn Mỹ Cửu Long
(*Email: bxtung@tdu.edu.vn)

Ngày nhận: 13/3/2022

Ngày phản biện: 09/4/2022

Ngày duyệt đăng: 29/4/2022

TÓM TẮT

Mục tiêu của nghiên cứu nhằm hỗ trợ bác sĩ chẩn đoán bệnh viêm phổi của trẻ em từ 1 tuổi đến 5 tuổi, thông qua ảnh X-quang ngực. Nguồn ảnh X-quang ngực được chọn dựa vào tập dữ liệu X-quang trên Kaggle (Kermany, 2018). Dữ liệu được tổ chức thành 2 thư mục và chứa các thư mục con cho từng loại hình ảnh. Tổng cộng là 5856 hình ảnh X-Ray, trong đó có 2 loại (viêm phổi và bình thường). Ngoài ra, nghiên cứu còn ghi nhận kết quả chẩn đoán của 20 ảnh X-quang ngẫu nhiên tại Bệnh viện Nhi Đồng Cần Thơ. Việc học sâu trong chẩn đoán bệnh dựa trên hình ảnh để thu được nhiều kết quả tốt. Tuy nhiên, để tăng độ chính xác trong việc phân loại chúng tôi đã dùng đối sánh đặc trưng và đối sánh mô hình. Trong nghiên cứu này, chúng tôi đề xuất đối sánh đặc trưng bao gồm: Color, Gist, Hog, Deep, Sift. Cuối cùng chúng tôi đối sánh 3 mô hình: CNN, SVM, KNN. Kết quả thực nghiệm mô hình SVM kết hợp với đặc trưng Deep là 96.72% vượt trội so với 2 môn hình còn lại là CNN 95,1%, KNN 95,2%. Từ kết quả này chúng tôi ứng dụng phân loại hình ảnh X-quang ngực để hỗ trợ việc chẩn đoán bệnh viêm phổi.

Từ khóa: K láng giềng gần nhất, mạng nơron tích chập, máy học vectơ hỗ trợ

Trích dẫn: Bùi Xuân Tùng, Huỳnh Thanh Danh, Trịnh Quang Minh, Ngô Thị Lan và Hoàng Văn Tư, 2022. Xây dựng hệ thống hỗ trợ việc phân loại bệnh viêm phổi bằng hình ảnh X-quang ngực dựa trên phương pháp học sâu. Tạp chí Nghiên cứu khoa học và Phát triển kinh tế Trường Đại học Tây Đô. 15: 227-244.

*Ths. Bùi Xuân Tùng – Giảng viên Khoa Kỹ thuật - Công nghệ, Trường Đại học Tây Đô

1. GIỚI THIỆU

Các loại ảnh: X-quang, chụp cộng hưởng từ, chụp cắt lớp được sử dụng rộng rãi trong phân tích lâm sàng. Bệnh SARS-CoV-2 xuất hiện cuối năm 2019 gây ra hậu quả lớn và kéo dài đến nay. Những nỗ lực để ngăn chặn bệnh bằng cách sử dụng các hình ảnh y tế là rất cần thiết. Việc bác sĩ sử dụng ảnh X-quang ngực để chẩn đoán bệnh, là phương pháp phổ biến. Hơn nữa Trường X-quang của Hoa Kỳ đã nghiên cứu trên ảnh X-quang phổi có thể được sử dụng để giảm thiểu nguy cơ lây nhiễm chéo (ACR, 2020). Theo NHS & NHS England (2000), X-quang phổi tạo ra hàng triệu lượt quét được thực hiện trên toàn cầu mỗi năm.

Kỹ thuật cải tiến việc học sâu mang lại cho ngành y tế các kết quả đầy hứa hẹn trong dự đoán các bệnh trong y tế (Kuhlmann et al., 2018; Kwon et al., 2018), khám phá về kháng sinh (Stokes et al., 2020), phân tích tự động về hồ sơ sức khỏe (Shickel et al., 2018). Hơn nữa, cải tiến phân loại ảnh, phân đoạn ảnh (Ma et al. 2020; Jiamin et al., 2020) cho ra kết quả đầy hứa hẹn cho sự phát triển của hình ảnh y tế. Một số phương pháp học sâu đã được đề xuất cho phát hiện và chẩn đoán bệnh, phân loại ung thư da (Esteva et al., 2017) hoặc kiến trúc bộ mã hóa, giải mã sâu cho hình ảnh 3D phân đoạn y sinh (Tseng et al., 2017). Gần đây nhất, Nguyễn Văn Hùng (2020) đã xây dựng hỗ trợ chẩn đoán và đề xuất các hướng điều trị bệnh phổi tắc nghẽn mạn tính qua thu thập dữ liệu tại Bệnh viện Đa khoa Cần Thơ về bệnh phổi này. Bên cạnh đó, nghiên cứu giải thích kết quả mạng

neutron tích chập phân loại hình ảnh X-quang ngực hỗ trợ chẩn đoán bệnh viêm phổi (Nguyễn Thanh Hải, 2020).

Đề tài nghiên cứu đóng góp về mặt lý thuyết và có thể triển khai trong thực tiễn với mục tiêu xây dựng một phần mềm dự đoán ảnh X-quang của bệnh nhi với đánh giá kết quả có độ chính xác cao và thời gian chẩn đoán nhanh.

2. PHƯƠNG PHÁP NGHIÊN CỨU

Trong nghiên cứu này đối sánh 5 đặc trưng, trên ảnh X-quang ngực của trẻ từ 1 tuổi đến 6 tuổi được đề xuất: Color, Gist, Sift, Hog, Deep. Việc đối sánh này dùng huấn luyện 2 mô hình SVM, KNN. Tiếp theo sẽ đối sánh chúng với mô hình CNN. Cuối cùng chọn ra mô hình tốt nhất để hỗ trợ xác định bệnh viêm phổi dựa trên ảnh X-quang sử dụng máy học. Dựa vào dữ liệu đó, chúng tôi xây dựng phần mềm hỗ trợ việc chẩn đoán trong thực tiễn. Khâu then chốt nhất của đề tài là xây dựng mô hình chẩn đoán bệnh và chẩn đoán bệnh án mới.

2.1. Nguồn dữ liệu

Nghiên cứu dựa vào tập dữ liệu X-quang trên Kaggle (Kermany, 2018). Dữ liệu được tổ chức thành 2 thư mục (train, test) và chứa các thư mục con cho từng loại hình ảnh (Pneumonia/Normal). Có 5856 hình ảnh X-Ray và 2 loại (Viêm phổi/Bình thường) đã được các bác sĩ chuyên khoa dự đoán. Kết quả đó đã được liệt kê thành 2 thư mục Pneumonia - Bệnh và Normal - Bình thường. Kết quả này đã được bên thứ 3 kiểm duyệt độ chính xác. Ngoài ra chúng tôi còn thu thập 20 ảnh X-quang bệnh nhi gồm 13

ảnh bệnh và 7 ảnh bình thường. Dữ liệu này tại Bệnh viện Nhi Đồng Cần Thơ.

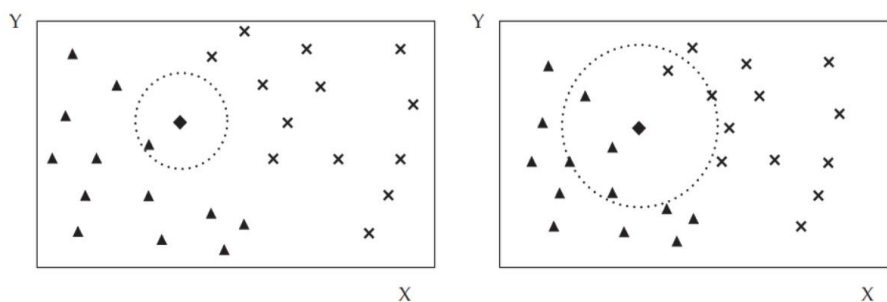
2.2. Phương pháp nghiên cứu

2.2.1. Các mô hình nghiên cứu

Mô hình KNN (K-Nearest Neighbor)

Thuật toán k láng giềng gần nhất là một kỹ thuật học có giám sát. Dùng để

phân loại quan sát mới bằng cách tìm điểm tương đồng giữa quan sát mới này với dữ liệu sẵn có. Quá trình tìm k láng giềng gần nhất thường sử dụng độ đo khoảng cách. Việc chọn giá trị k sẽ ảnh hưởng đến hiệu quả phân lớp của mô hình KNN. Hai hình bên dưới mô tả quá trình chọn k.



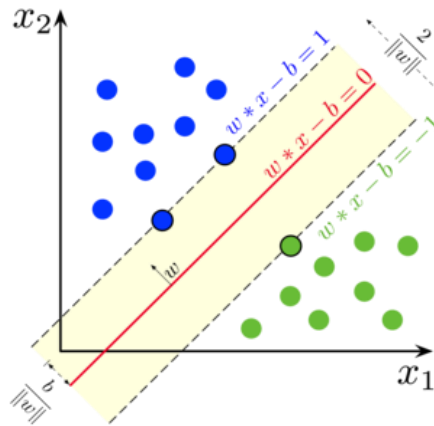
Hình 1. Chọn láng giềng $k = 2$ và $k = 5$

Trong mô hình này đã sử dụng các tham số `knn_model(knn_n_neighbors=5, knn_algorithm='auto', knn_leaf_size=30, knn_p=2)`. Tức là chọn 5 láng giềng gần nhất. Số láng giềng gần nhất là con số lẻ để tránh trường hợp bầu chọn 50% và 50%.

Mô hình SVM

SVM là một giải thuật máy học dựa trên lý thuyết học thống kê do Vapnik and Chervonenkis (1974), Vapnik (1999) xây dựng. Bài toán cơ bản của SVM là bài toán phân loại hai lớp: Cho trước r điểm, trong không gian n chiều (mỗi điểm thuộc

vào một lớp kí hiệu là $+1$ hoặc -1). Mục đích của giải thuật SVM là tìm một siêu phẳng. Phân hoạch tối ưu cho phép chia các điểm này thành hai phần, sao cho các điểm cùng một lớp nằm về một phía với siêu phẳng này. Cho một minh họa phân lớp với SVM trong mặt phẳng.



Hình 2. Siêu phẳng với lề cực đại trong không gian R²

Xét tập r mẫu huấn luyện $\{(x_1, y_1), (x_2, y_2), \dots, (x_r, y_r)\}$. Trong đó x_i là một vector đầu vào được biểu diễn trong không gian n $X \subseteq \mathbb{R}^n$, y_i là một nhãn lớp; $y_i \in \{1, -1\}$. Siêu phẳng tối ưu phân tập dữ liệu này thành hai lớp là siêu phẳng có thể tách rời dữ liệu thành hai lớp riêng biệt với lề (margin) lớn nhất. Tức là, cần tìm siêu phẳng $H_0: y = w \cdot x + b = 0$ và hai siêu phẳng H_+, H_- hỗ trợ song song với H_0 và có cùng khoảng cách đến H_0 . Với điều kiện không có phần tử nào của tập mẫu nằm giữa H_+ và H_- , khi đó:

$$H_+: w \cdot x + b \geq +1 \text{ với } y = +1$$

$$H_-: w \cdot x + b \leq -1 \text{ với } y = -1$$

Kết hợp hai điều kiện trên, có $y(w \cdot x + b) \geq 1$. Khoảng cách của siêu phẳng H_+ và H_- đến H_0 là: $\frac{1}{\|w\|}$. Cần tìm siêu phẳng H_0 với lề lớn nhất là: $\frac{2}{\|w\|}$. giải bài toán tối ưu tìm w và b sao cho: $\frac{w \cdot w}{2}$ đạt cực đại với ràng buộc $y_i (w \cdot x_i + b) \geq 1$. Tương đương với bài toán cực tiểu hóa: với điều kiện: $y_i (w \cdot x_i + b) \geq 1$, mọi $i=1 \dots r$. Lời giải cho bài toán tối ưu này là cực tiểu hóa hàm Lagrange:

$$L(w, b, \alpha) = \frac{1}{2} \|w \cdot w\| - \sum_{i=1}^r \alpha_i [y_i (\|w \cdot x_i\| + b) - 1]$$

Trong đó, α là các hệ số Lagrange, $\alpha \geq 0$. Lời giải tìm siêu phẳng tối ưu trên có thể mở rộng trong trường hợp dữ liệu không thể tách rời tuyến tính bằng cách ánh xạ dữ liệu vào một không gian có số chiều lớn hơn bằng cách sử dụng một hàm nhân K (kernel). Một số hàm nhân

thường dùng như: Hàm tuyến tính có dạng $K(x,y)=x \cdot y$. Hàm đa thức có dạng $K(x,y)=(x \cdot y + 1)^d$. Hàm RBF (Radial Basis Function) có dạng:

$$K(x, y) = e^{-\frac{\|x-y\|^2}{2\alpha^2}}$$

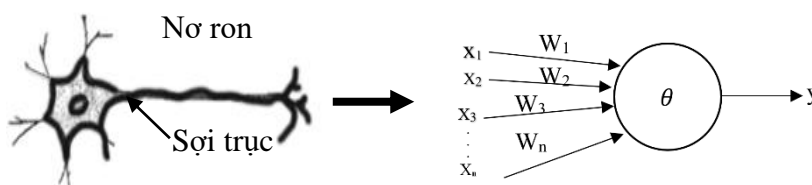
Với khả năng vượt trội của SVM về tính hiệu quả, độ chính xác, khả năng xử lý các bộ dữ liệu một cách linh hoạt, việc sử dụng máy vectơ hỗ trợ SVM đã và đang là sự lựa chọn tối ưu quá trình nhận dạng sau này. Quá trình huấn luyện dữ liệu nhanh hay chậm phụ thuộc vào số lượng mẫu dữ liệu tham gia huấn luyện, thuật toán chọn để huấn luyện dữ liệu.

Các tham số trong mô hình này là: svm_model

(svm_C=1,svm_kernel='linear', num_output = 1). Cùng với tham số 'gamma': [0.0001, 0.0005, 0.001, 0.005, 0.01, 0.1]. Với hàm nhân kernel=RBF.

Mô hình CNN

Neural là mô hình toán học, mô phỏng neuron trong hệ thống thần kinh con người. Nó biểu hiện cho một số chức năng của neuron (neuron) thần kinh con người.



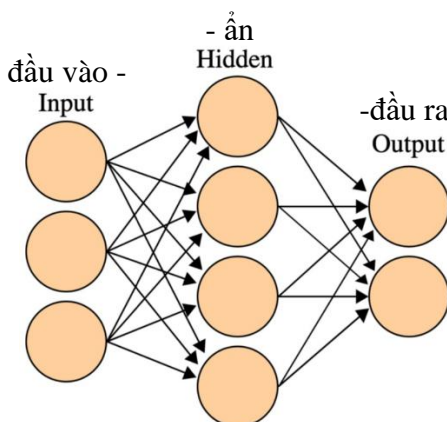
Hình 3. Mạng Neural

Đầu tiên là tính chất truyền đi của thông tin trên neuron. Nó nhận tín hiệu đầu vào, khi tín hiệu vượt qua một ngưỡng (threshold) thì tín hiệu sẽ được truyền đi sang neuron khác (Neurons Fire) theo sợi trục (axon). Neural của mô hình toán học ở đây cũng được mô phỏng tương tự như vậy.

Mạng neural được xây dựng dựa trên mạng neural sinh học. Nó gồm các neural (nút) nối với nhau, và xử lý thông tin bằng cách truyền theo các kết nối sau đó tính giá trị tại các nút. Mạng neuron với mỗi nút sẽ có những dữ liệu đầu vào, biến đổi những dữ liệu đầu vào này bằng cách tính tổng các input, với trọng số tương ứng trên các đầu vào. Sau đó áp

dụng một hàm biến đổi phi tuyến tính cho phép biến đổi này, để tính toán trạng thái trung gian. Sau cùng tạo thành 1 lớp và hàm biến đổi còn được gọi là activation function. Các đầu ra của lớp này sẽ là đầu vào của lớp phía sau.

Thông qua việc lặp lại các bước trên, neural-network học thông qua nhiều lớp và các nút phi tuyến tính rồi. Sau đó kết hợp lại ở lớp cuối cùng để cho ra 1 dự đoán. Neural-network học bằng cách tạo ra các tín hiệu lỗi, đo lường sự khác biệt giữa các dự đoán của mạng và giá trị mong muốn. Sau đó sử dụng tín hiệu lỗi này để cập nhật lại trọng số và bias trong activation function để việc dự đoán sau đó chính xác hơn.



Hình 4. Đầu vào và đầu ra của Neural Network

Các tham số tham gia trong mô hình CNN này:

`model_cnn(input_reshape=(32,32,3),num_classes=2,optimizers_func='Adam',numfilter=20, filtersize=3,numlayercnn_per_maxpool=1,nummaxpool=1,maxpoolsize=2. Và model.add(Conv2D(64, (3, 3), activation="relu", padding='same', input_shape=input_reshape))`

`model.add(MaxPooling2D(pool_size = (2, 2)))`

`model.add(Dropout(0.1))`

`model.add(Conv2D(64, (3, 3), activation="relu"))`

`model.add(MaxPooling2D(pool_size = (2, 2)))`

Chúng tôi đã thử các tham số khác nhau trên 03 mô hình. Cuối cùng những tham số đã nêu trên cho kết quả như đã trình bày ở phần tóm tắt.

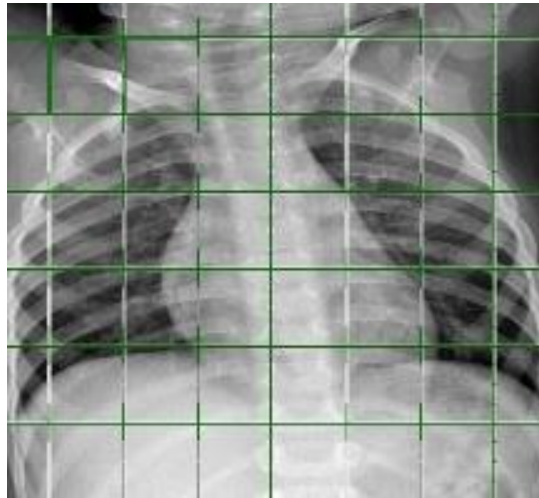
2.2.2. Tiền xử lý ảnh

Mô hình CNN: Nguyên nhân chính để chuẩn hóa hình ảnh là để tránh khả năng bùng nổ các gradient vì dải pixel cao (0-255) và tăng tốc độ hội tụ của mô hình. Trong quá trình lan truyền, thực hiện các phép toán nhân của các giá trị pixel này với ma trận trọng số cho lớp cụ thể. Việc nhân các giá trị pixel lớn này tốn rất nhiều tài nguyên và thời gian tính toán, dẫn đến việc mô hình có thể hội tụ rất chậm.

Người nghiên cứu đã chỉnh lại kích thước cho tất cả ảnh đầu vào là 224x224x3. Trong đó: 224x224 là giá trị của chiều dài và chiều rộng ảnh; 3: là số kênh màu (RGB).

Các đặc trưng

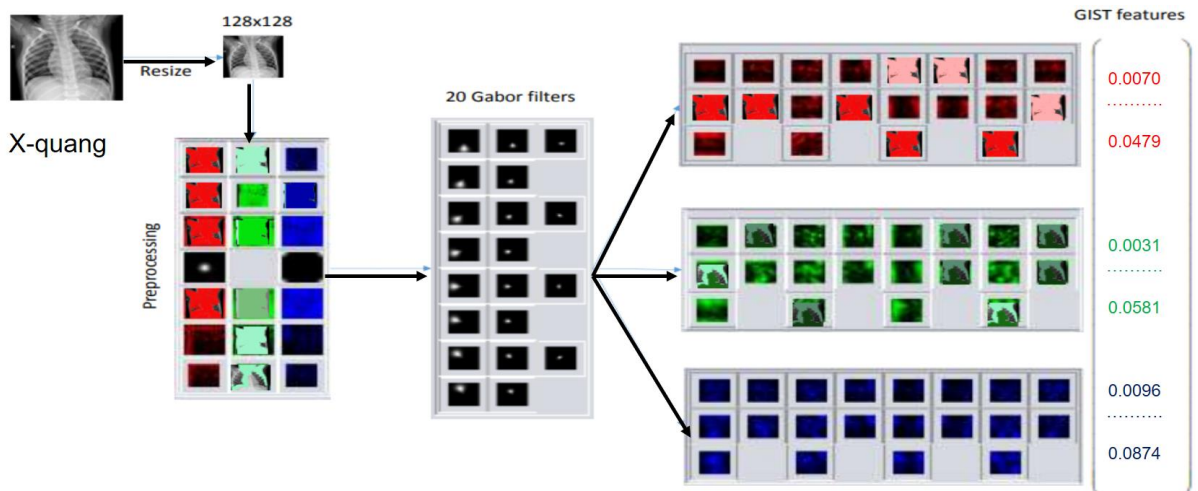
Đặc trưng HOG: Chuẩn hóa: ảnh X-quang thành 80x80, kích thước block 8x8, kích thước cell: 1x1 mỗi cell gồm 64 pixels.



Hình 5. Đặc trưng HOG

Sau đó kết hợp đặc trưng theo unsigned-HoG(p=9 pins). Mỗi ảnh sau khi tách đặc trưng HOG sẽ gồm 900 chiều dữ liệu.

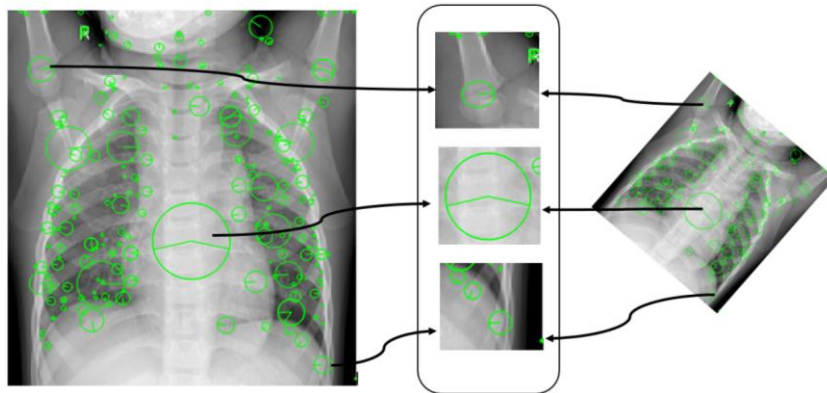
Đặc trưng Gist: Là đặc trưng toàn cục, dựa vào phép biến đổi Gabor theo các hướng và tần số khác nhau.



Hình 6. Đặc trưng Gist

Từ dữ liệu đầu vào chuẩn hóa ảnh với kích thước 128x128. Sau khi trích đặc trưng sẽ cho ra một vectơ 960 chiều dữ liệu.

Đặc trưng Sift: hay còn gọi là đặc trưng cục bộ bất biến SIFT. So sánh đối chiếu các điểm ảnh không cùng tỉ lệ, cũng như góc độ từ các điểm nổi bật.

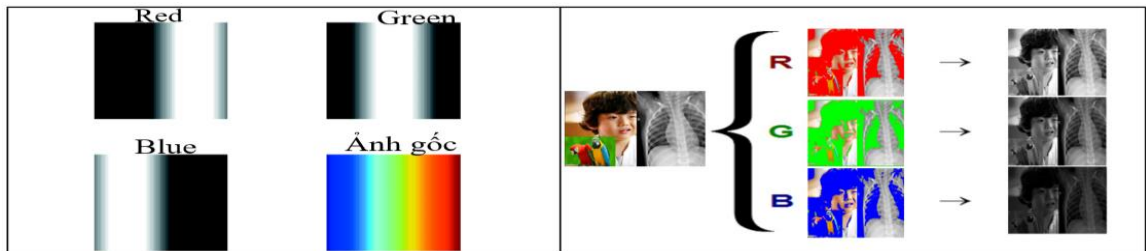


Hình 7. Đặc trưng Sift

Bất biến khi: Thay đổi tỷ lệ ảnh; Quay ảnh; Thay đổi góc nhìn; Thêm nhiều ảnh; Thay đổi cường độ chiếu sáng ảnh.

Đặc trưng màu sắc: Người nghiên cứu đã thay đổi kích thước cho tất cả các

ảnh đầu vào về cỡ 80x80. Đặc trưng này biểu diễn như một điểm trong không gian màu sắc ba chiều. Với các màu tự nhiên đỏ, xanh và lơ (R, G, B).



Hình 8. Đặc trưng màu sắc

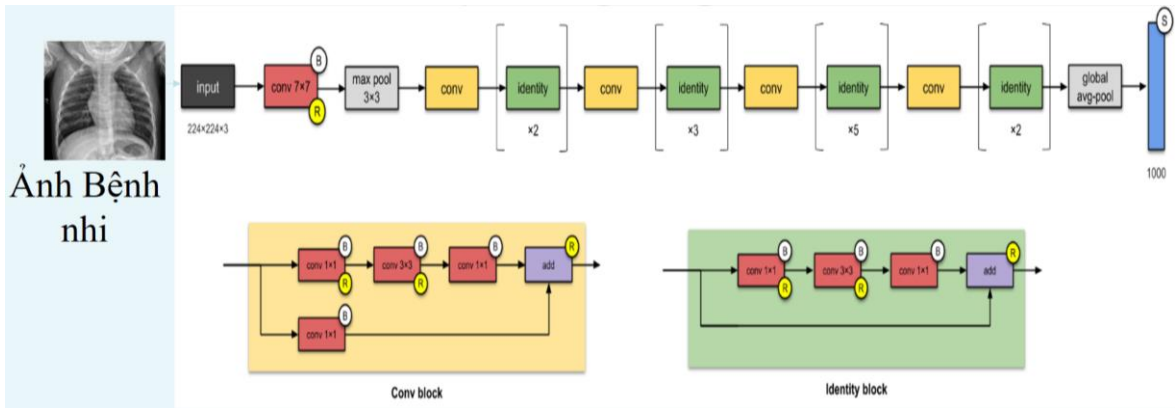
Bước đầu tiên tách kênh màu, kể đến sẽ chuyển ảnh từng kênh màu thành 3 kênh tương ứng là 3 kênh xám. Nhằm xác định tỉ trọng các điểm ảnh của ảnh chứa các giá trị đặc biệt (màu sắc). Cuối cùng phân vùng ảnh có mối quan hệ tương đồng với nhau.

Đặc trưng Deep: Đầu vào là ảnh RGB được chuẩn hóa thành 224x224x3. Đặc

trung Deep được xuất thân từ các mô hình học VGG-16 và ResNet-50.

Dữ liệu đầu vào cho các mô hình SVM, KNN chính là đặc trưng này. Sẽ trích xuất từ feature map cuối cùng.

Tất cả 5 đặc trưng được kể trên lần lượt chạy trên 2 mô hình KNN và SVM được mô tả bên dưới. Mục đích cuối cùng chúng tôi chọn đặc trưng với mô hình tốt nhất để đối sánh với mô hình CNN.



Hình 9. Đặc trưng Deep

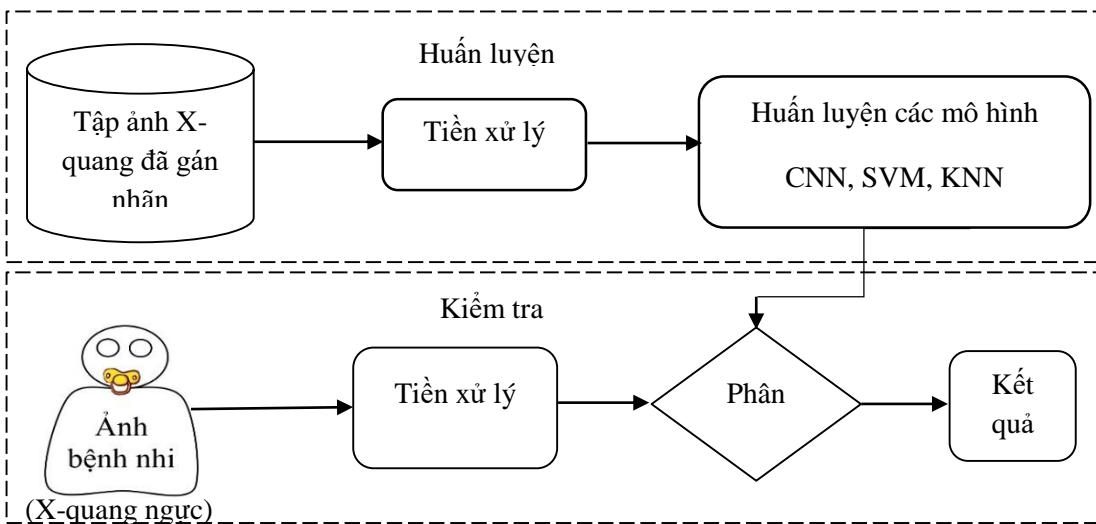
2.2.3. Các bước thực hiện

Dựa vào ảnh X-quang đã được gán nhãn chúng tôi đã đề xuất giải pháp cụ thể là:

- Một là: Trích xuất các đặc trưng (Color, Gist, Hog, Deep, Sift)
- Hai là: Huấn luyện mô hình CNN, SVM, KNN

- Ba là: Đối sánh đặc trưng
- Bốn là: Đối sánh mô hình
- Cuối cùng: Chọn mô hình hiệu quả nhất chẩn đoán bệnh án mới.

Để chẩn đoán cho trẻ em từ 1 đến 5 tuổi. Chúng tôi đề xuất mô hình, thông qua bảng vẽ hệ thống chẩn đoán bệnh viêm phổi ở bên dưới.

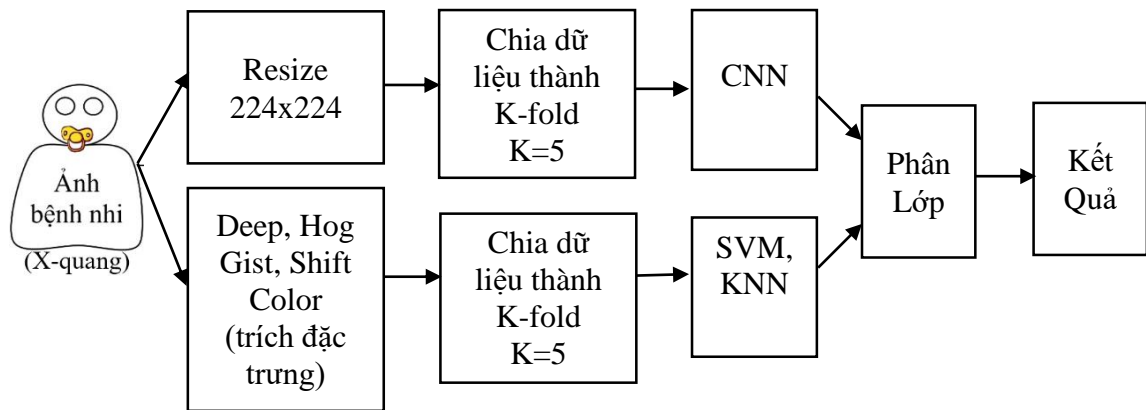


Hình 10. Mô hình huấn luyện và đối sánh tổng quát

Tiếp theo là 3 nội dung chính: Tập dữ liệu huấn luyện, tiêu chí đánh giá, hệ thống chẩn đoán.

Trước tiên người nghiên cứu đã thực hiện cho hai nhóm mô hình: Thứ nhất là

mô hình CNN, thứ hai là nhóm các mô hình: SVM, KNN. Đối với mô hình thứ 2 này sử dụng các đặc trưng: Gist, Shift, Hog, Deep, Color. Quá trình này được mô tả chi tiết như Hình 11.



Hình 11. Mô hình huấn luyện và đối sánh chi tiết

2.3. Phân chia dữ liệu

Chúng tôi chia thành 2 thư mục: Normal chiếm 1.583 ảnh, Pneumonia

chiếm 4.273 ảnh. Được trình bày theo bảng bên dưới:

Bảng 1. Phân chi dữ liệu huấn luyện

Tên kết quả chẩn đoán	Số lượng
Bình thường	1.583
Bệnh phổi	4.273
Tổng cộng	5.856

Từ bảng dữ liệu này chúng tôi chia thành 2 phần. Phần huấn luyện: 4.685 ảnh X-quang chiếm 80%; phần kiểm tra mô hình 1.171 ảnh X-quang chiếm 20% trong tập dữ liệu.

2.3.1. Dữ liệu độc lập cho demo

Ngoài ra, chúng tôi còn thu thập ngẫu nhiên 20 ảnh X-quang của 20 bệnh nhi trên địa bàn TP. Cần Thơ trong khoảng thời gian từ ngày 01/06/2021 đến 07/06/2021, được trình bày ở Bảng 2.

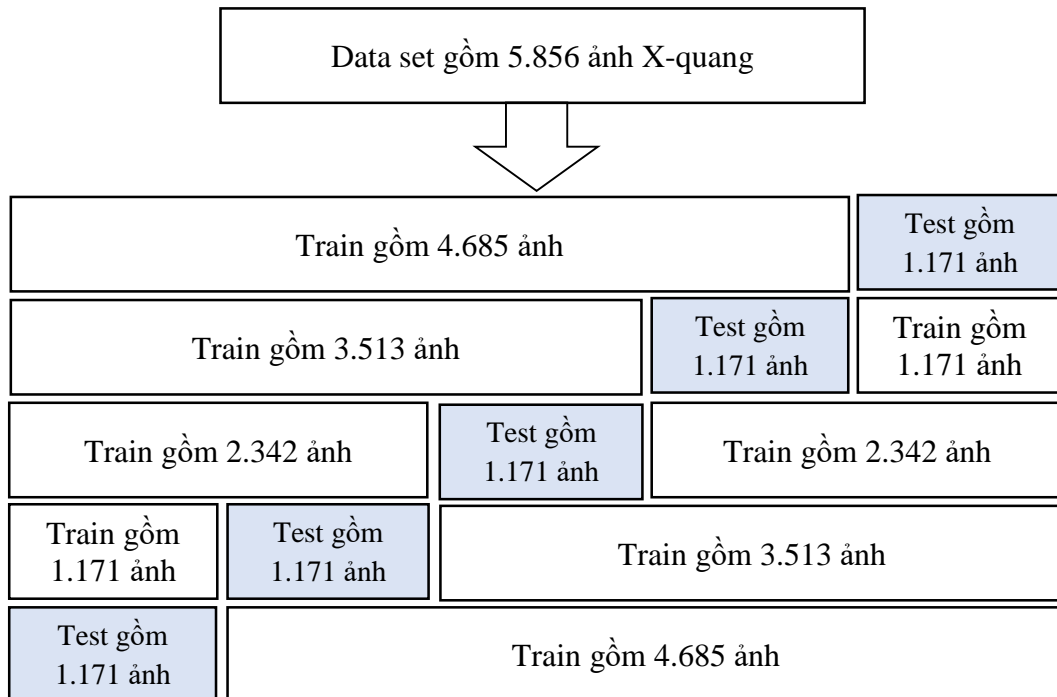
Bảng 2. Ảnh X-quang tại Bệnh viện Nhi Đồng Cần Thơ

Tên kết quả chẩn đoán	Số lượng
Bình thường	7
Bệnh phổi	13
Tổng cộng	20

2.3.2. Chia dữ liệu huấn luyện thành K-Fold

Để hoàn thành quá trình huấn luyện và kiểm thử. Người nghiên cứu chia dữ liệu

cho 4 mô hình với K=5. Dữ liệu ảnh ngẫu nhiên dùng huấn luyện chiếm 80%, còn lại 20% dùng cho kiểm thử. Được miêu tả chi tiết như Hình 12.



Hình 12. Dữ liệu K-Fold với K=5

2.3.3. Tiêu chí đánh giá

Trong nghiên cứu này, sau khi kiểm thử cho ra 2 trạng thái. Trạng thái thứ nhất

là bệnh, trạng thái thứ 2 là không bệnh. Chúng tôi đánh giá 4 mô hình dựa và độ chính xác (accuracy). Được liệt kê theo Hình 13.

		Predicted class	
		Positive	Negative
Actual class	Positive	Có bệnh đoán ra bệnh (TP)	Có bệnh đoán ra không bệnh (FN)
	Negative	Không bệnh đoán ra bệnh (FP)	Không bệnh đoán ra không bệnh (TN)

Hình 13. Các trường hợp xảy ra trong dự đoán

Cuối cùng sẽ dùng công thức tính độ chính xác trung bình này cho bài toán nghiên cứu. Từ đây người nghiên cứu sẽ đối sánh 4 mô hình. Kết quả đối sánh sẽ

được trình bày ở bảng 10 (bảng đối sánh 4 mô hình). Bên dưới là công thức tính độ chính xác cho nghiên cứu.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{FP} + \text{TN}}$$

3. KẾT QUẢ THỰC NGHIỆM VÀ ĐÁNH GIÁ GIẢI PHÁP

3.1. Mô hình kết hợp các đặc trưng

3.1.1. Mô hình kết hợp đặc trưng màu sắc

Sau khi chạy 2 mô hình SVM, KNN với đặc trưng **màu sắc**. Chỉ số K-Fold(K=5), chúng tôi thu được kết quả như Bảng 3.

Bảng 3. Mô hình kết hợp đặc trưng màu

model	val_accuracy
svm	0.91
knn	0.90

Đổi sánh mô hình, **SVM** với đặc trưng **màu sắc** cho kết quả cao nhất **0.91**.

3.1.2. Mô hình kết hợp đặc trưng Gist

Sau khi chạy 2 mô hình SVM, KNN với đặc trưng **Gist**. Chỉ số K-Fold(K=5), chúng tôi thu được kết quả như Bảng 4.

Bảng 4. Mô hình kết hợp đặc trưng Gist

model	val_accuracy
knn	0.80
svm	0.82

Đổi sánh mô hình, **SVM** với đặc trưng **Gist** cho kết quả cao nhất **0.82**.

3.1.3. Mô hình kết hợp đặc trưng Hog

Sau khi chạy 2 mô hình SVM, KNN với đặc trưng **Hog**. Với số K-Fold(K=5), chúng tôi thu được kết quả như Bảng 5.

Bảng 5. Mô hình kết hợp đặc trưng Hog

model	val_accuracy
knn	0.93
svm	0.95

Đổi sánh mô hình, **SVM** với đặc trưng **Hog** cho kết quả cao nhất **0.95**.

3.1.4. Mô hình kết hợp đặc trưng Sift

Sau khi chạy 2 mô hình SVM, KNN với đặc trưng **Sift**. Chỉ số K-Fold(K=5), chúng tôi thu được kết quả như Bảng 6.

Bảng 6. Mô hình kết hợp đặc trưng Sift

model	val_accuracy
svm	0.91
knn	0.77

Đối sánh mô hình, SVM với đặc trưng Sift cho kết quả cao nhất **0.91**.

3.1.5. Mô hình kết hợp đặc trưng Deep

Sau khi chạy 2 mô hình SVM, KNN với đặc trưng Deep. Chỉ số K-Fold(K=5), chúng tôi thu được kết quả như Bảng 7.

Bảng 7. Mô hình kết hợp đặc trưng Deep

model	val_accuracy
svm	0.970
knn	0.953

Tổng hợp của 5 đặc trưng trên 2 mô hình theo Bảng 8.

Bảng 8. Tổng hợp của 5 đặc trưng trên 2 mô hình

	SVM	KNN
Color	0.910	0.900
Gist	0.820	0.800
Hog	0.950	0.930
Sift	0.910	0.770
Deep	0.970	0.953

Đối sánh mô hình, SVM với đặc trưng Deep cho kết quả cao nhất 0.97. Trong 2 mô hình, thì mô hình này đạt độ chính xác cao nhất trong 5 đặc trưng. Người nghiên cứu sẽ chọn mô này đối sánh với mô hình CNN ở phần sau.

3.2. Đối sánh các mô hình với CNN

3.2.1. Mô hình CNN

Theo bảng dữ liệu tại mục 2.3.2. Chia dữ liệu huấn luyện thành K-Fold. Quá trình huấn luyện trên: 4.685 dữ liệu train và 1.171 ảnh test. Người nghiên cứu đã kiểm thử cho kết quả ở Bảng 9.

Bảng 9. Bảng mô tả kết quả phân lớp theo từng fold

K-Fold	v_acc	tn	fp	fn	tp	time
fold 1/5	0.961	293	24	22	833	167.38
fold 2/5	0.961	290	27	19	835	397.33
fold 3/5	0.945	274	43	21	833	499.41
fold 4/5	0.935	270	46	30	825	589.37
fold 5/5	0.955	290	26	27	828	755.5
Tổng		1417	166	119	4154	2409
Trung bình	0.951	283.4	33.2	23.8	830.8	

Cột đầu tiên K=5: Mô tả từ fold 1/5 đến fold 5/5.

- v_acc: Mô tả độ chính xác của mỗi fold.

- TN: Có bệnh mà đoán ra không bệnh.
- FP: Không bệnh mà đoán ra bệnh.
- FN: Có bệnh mà đoán ra không bệnh.
- TP: Có bệnh đoán ra bệnh thật.
- Time: Thời gian thực hiện

Kết luận độ chính xác của mô hình CNN là **0.951**.

3.2.2. Đối sánh mô hình

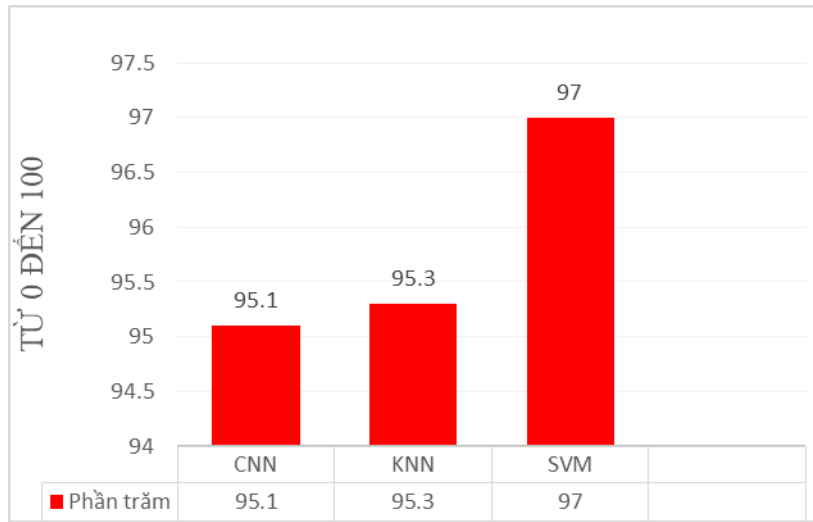
Công trình nghiên cứu của chúng tôi đã đối sánh 2 mô hình SVM, KNN cho thấy kết quả cao nhất là mô hình SVM với đặc trưng Deep. Việc này cho chúng tôi độ chính xác là 0.97. Sau khi thực nghiệm trên mô hình CNN cho kết quả khả quan là: 0.95 được mô tả bằng Bảng 10.

Bảng 10. Bảng đối sánh 4 mô hình

Model	Độ chính xác %
RF	94,7
CNN	95,1
KNN	95,3
SVM	97,0

Kết quả nghiên cứu đã chọn được mô hình tối ưu nhất là Mô hình SVM với đặc trưng Deep đã cho độ chính xác 97%. Kết

quả của công trình nghiên cứu này đã thể hiện ở Hình 14.



Hình 14. Đối sánh kết quả 3 mô hình

4. KẾT LUẬN

Mô hình tối ưu cho nghiên cứu là SVM. Mô hình đã kết hợp với đặc trưng Deep qua quá trình đối sánh với 4 đặc trưng: Color, Gist, Shift, Hog.

Độ tin cậy đạt 97% góp phần hỗ trợ cho bác sĩ chẩn đoán bệnh viêm phổi cho bệnh nhi. Hơn nữa mô hình này đã rút ngắn thời gian chẩn đoán.

Tuy đã đối sánh được 5 đặc trưng và đối sánh 3 mô hình, cải thiện thời gian chẩn đoán bệnh viêm phổi, nhưng hệ thống chẩn đoán cần được kết nối với cơ sở dữ liệu đa phương tiện trong các cơ sở y tế. Nghiên cứu tiếp theo sẽ hướng tới việc chẩn đoán ra kết quả chi tiết hơn.

TÀI LIỆU THAM KHẢO

1. ACR, 2020. Recommendations for the use of Chest Radiography and Computed Tomography (CT) for Suspected COVID-19 Infection.

(American College of Radiology), Accessed 05 01, 2021, from <https://www.acr.org>

2. Esteva, Kuprel, Novoa, R, A, & B, 2017. Dermatologist level classification of skin cancer with deep neural networks. *Nature*, 582, pp 115–118.

3. Jiamin Liu, Song Yang, Hong Huang, & Zhengying Li & Guangyao Shi, 2020. A deep feature manifold embedding method for hyperspectral image classification. *Remote Sensing Letters*, 11:7, 620-629. doi:10.1080/2150704X.2020.1746855

4. Kermany, 2018. Chest X-Ray Images (Pneumonia). Accessed 08 01, 2020, from <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>

5. Kuhlmann L, Lehnertz K, Richardson MP, Schelter B, & Zaveri

HP, 2018. Seizure prediction - ready for a new era. *Nat Rev Neurol*, 14(10):618-630. doi:10.1038/s41582018-

6. Kwon JM, Lee Y, Lee Y, Lee S, & Park J, 2018. An Algorithm Based on Deep Learning for Predicting In-Hospital Cardiac Arrest. *J Am Heart Assoc*, 7(13):e008678. doi:10.1161/JAHA.118.008678

7. Nguyễn Thanh Hải, 2020. Explanation of the Convolutional Neural Network Classifying Chest X-Ray Images Supporting Pneumonia Diagnosis. *EAI Endorsed Transactions*. Accessed 09 04, 2021, from <https://eprints.eudl.eu/id/eprint/265/1/eai.13-7-2018.165349.pdf?fbclid=iwar19qy1x14sxnyakvtfkufkfqipzc4hhl1dh0xqp-yh117qw9pflsqgdme>

8. Nguyễn Văn Hùng, 2020. Xây dựng hỗ trợ chẩn đoán và đề xuất các hướng điều trị bệnh phổi tắc nghẽn mạn tính. Đại học Cần Thơ.

9. NHS, & NHS England, 2021. Diagnostic imaging dataset statistical. Accessed 05 01, 2021, from <https://www.england.nhs.uk/>

10. Ma, S. Xu, X. Yi, & L. Li and C. Yu, 2020. Research on Image Classification Method Based on DCNN. *International Conference on Computer Engineering and Application*, (p. 873876). Guangzhou, China. doi:10.1109/ICCEA50009.2020.00192

11. Shickel, P. J. Tighe, & A. Bihorac and P. Rashidi. 2018. Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis. *IEEE Journal of Biomedical and Health Informatics*, 22, pp. 1589-1604. doi:doi:10.1109/JBHI.2017.2767063

12. Stokes JM, Yang K, & Swanson K, 2020. A Deep Learning Approach to Antibiotic Discovery appears in *Cell*. *Cell*, 181(2):475-483. doi:doi:10.1016/j.cell.2020.01.021.

13. Tseng, Y. Lin, & W. Hsu and C. Huang, 2017. Joint Sequence Learning and Cross-Modality Convolution for 3D Biomedical Segmentation. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3739-3746. doi:10.1109/CVPR.2017.398

BUILDING A SYSTEM SUPPORTING THE CLASSIFICATION OF PNEUMONIA IN CHEST X-RAY IMAGES BASING ON DEEP LEARNING METHOD

Bui Xuan Tung^{1*}, Huynh Thanh Danh¹, Trinh Quang Minh¹,
Ngo Thi Lan¹ and Hoang Van Tu²

¹Tay Do University, ²Hoan My Cuu Long Hospital
(*Email: bxtung@tdu.edu.vn)

ABSTRACT

The objective of this study was to assist doctors in diagnosing pneumonia in children aged 1 to 5 years, through chest X-ray images. The chest X-ray image source was selected from Guangzhou Women's and Children's Medical Center. The data is organized into 2 folders (train, test) and contains subfolders for each image type. There is a total of 5856 X-Ray images, of which there are 2 types (pneumonia and normal). In addition, the researcher also recorded the diagnostic results of 20 random X-ray images at Can Tho Children's Hospital. Deep learning in image-based disease diagnosis was studied to obtain many good results. However, to increase the accuracy of the classification, we used feature matching and model matching. In this study, we propose feature matching which includes: Color, Gist, Hog, Deep, and Sift. Finally, we compared 3 models: CNN, SVM, KNN. The experimental results of the SVM model combined with Deep features are 96.72% superior to the other two models, CNN 95.1%, KNN 95.2%. From this result, we can apply chest X-ray image classification to support the diagnosis of pneumonia.

Keywords: *Convolutional Neutron Network-CNN, K. Nearest Neighbor-KNN, Support Vector Machine –SVM*