



TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

KHOA CÔNG NGHỆ THÔNG TIN

BÁO CÁO THỰC TẬP DỰ ÁN TỐT NGHIỆP

ĐỀ TÀI: PHÂN LOẠI ẢNH CÓ KHỐI U VÀ KHÔNG CÓ KHỐI U Ở PHỔI

CHUYÊN NGÀNH: KHOA HỌC DỮ LIỆU

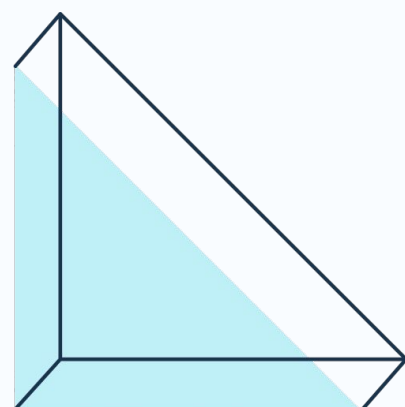
GIẢNG VIÊN HƯỚNG DẪN: PGS. TS LÊ HOÀNG THÁI

SINH VIÊN THỰC HIỆN: HỨA PHÚ THÀNH - 18120563

HÀ MINH TOÀN - 18120597

TRẦN NGỌC TỊNH - 18120599

TRẦN LUẬT VY - 18120656

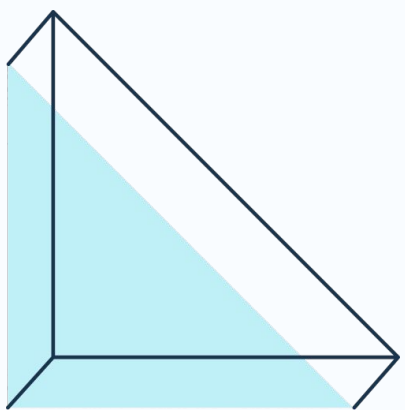


Nội dung chính

- 1 Vấn đề và bài toán nghiên cứu
- 2 Dữ liệu cho bài toán
- 3 Phương pháp đề xuất
- 4 Kết quả thực nghiệm
- 5 Kết luận và hướng phát triển

2

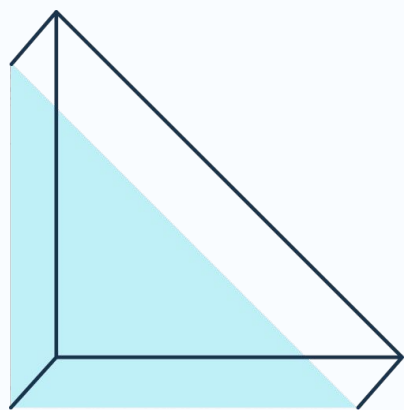
Vấn đề và bài toán nghiên cứu



Lý do nghiên cứu:

- Ung thư phổi là một trong những loại ung thư phổ biến nhất trên thế giới hiện nay.
- Căn bệnh nguy hiểm nhưng có ít hoặc không có bất kỳ triệu chứng lâm sàng.
- Khi có các triệu chứng của bệnh thì ung thư đã vào giai đoạn cuối.
- Phát hiện ung thư ở phổi bằng biện pháp chụp cắt lớp phổi (chụp CT phổi).

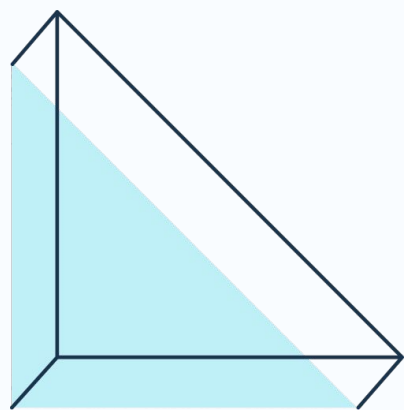
Vấn đề và bài toán nghiên cứu



Mục tiêu nghiên cứu:

- Áp dụng trí tuệ nhân tạo để giải quyết bài toán “Phân loại ảnh có khối u hoặc không có khối u ở phổi”.
- Giảm thời gian chẩn đoán thủ công cho các bác sĩ.
- Xây dựng mô hình có độ chính xác cao trong việc phân loại.
- Giảm tỉ lệ dương tính giả và âm tính giả.

Vấn đề và bài toán nghiên cứu



Hướng giải quyết đã có:

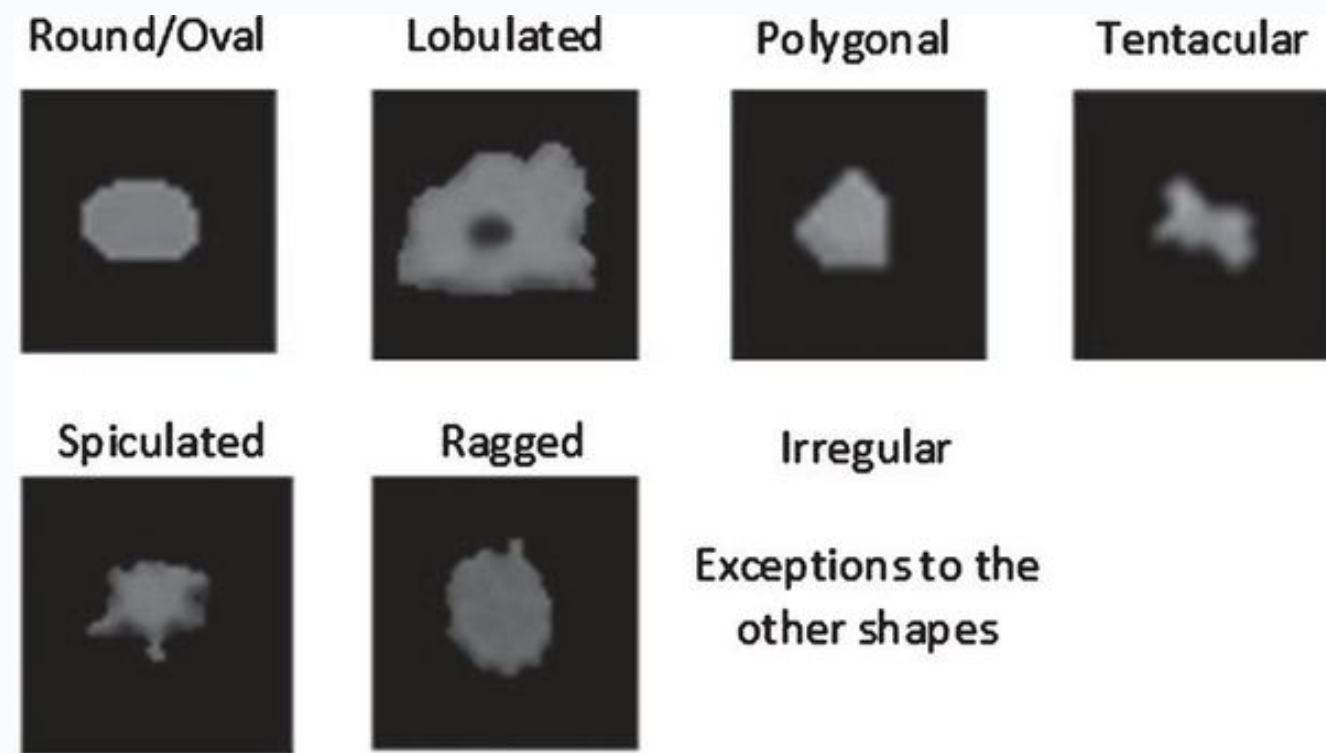
- Sử dụng mạng CNN kết hợp với Attention Mechanism do Lai Đình Khải và các cộng sự¹.
- Các tác giả Wook Jin Choi và Tae Sun Choi tiếp cận theo hướng sử dụng Hierarchical Block kết hợp với SVM để phân loại⁵.
- Trần Giang Sơn cùng những cộng sự đã sử dụng phương pháp học sâu CNN kết hợp với mất mát tiêu điểm (Focal Loss).

Dữ liệu cho bài toán

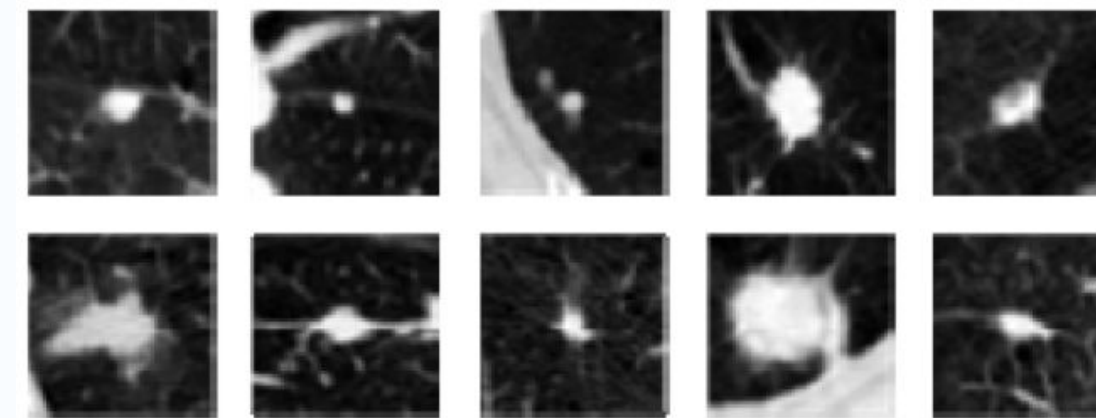
Tập dữ liệu sử dụng: nguồn gốc từ tập dữ liệu LIDC-IDRI² được tác giả Swetha Subramanian điều chỉnh lại.

Tập dữ liệu gồm 2 lớp là **Positive** và **Negative**.

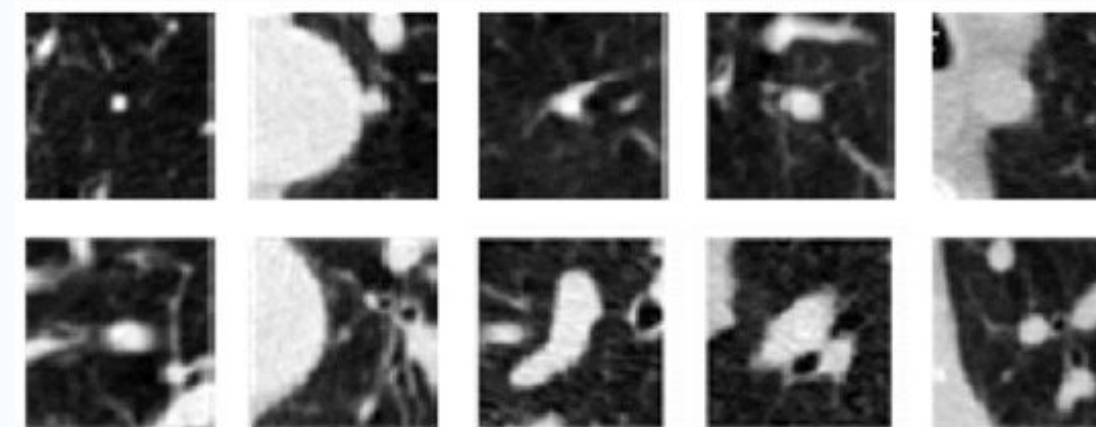
- **Positive:** các nốt sần ở phổi có đường kính lớn hơn hoặc bằng 3mm.
- **Negative:** các nốt sần ở phổi có đường kính bé hơn 3mm và các tổn thương khác ở phổi không phải là nốt sần có đường kính lớn hơn 3mm.



Hình 1. Hình dạng của các nốt sần thuộc lớp Positive ⁴



Hình 2a. Hình dạng của các nốt sần thuộc lớp Positive ⁴



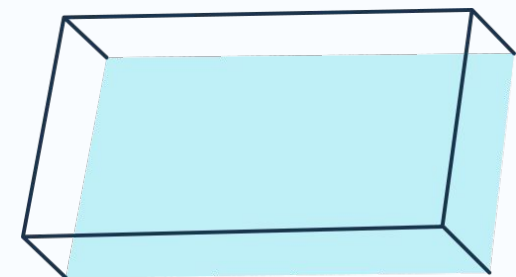
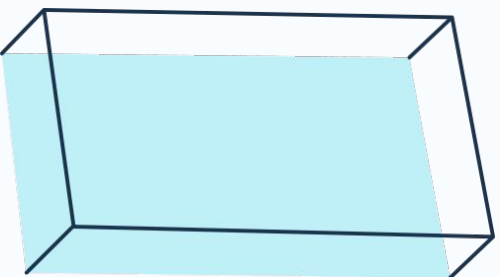
Hình 2b. Hình dạng của các nốt sần thuộc lớp Negative ⁴

Dữ liệu cho bài toán

Tập dữ liệu gồm **8106** ảnh có kích thước **50x50** pixel được chia thành 3 thư mục:

	Positive	Negative	Sum
Train	845	4342	5187
Validation	224	1073	1297
Test	282	1340	1622

Bảng 1. Số lượng các lớp thuộc các tập train, validation và test





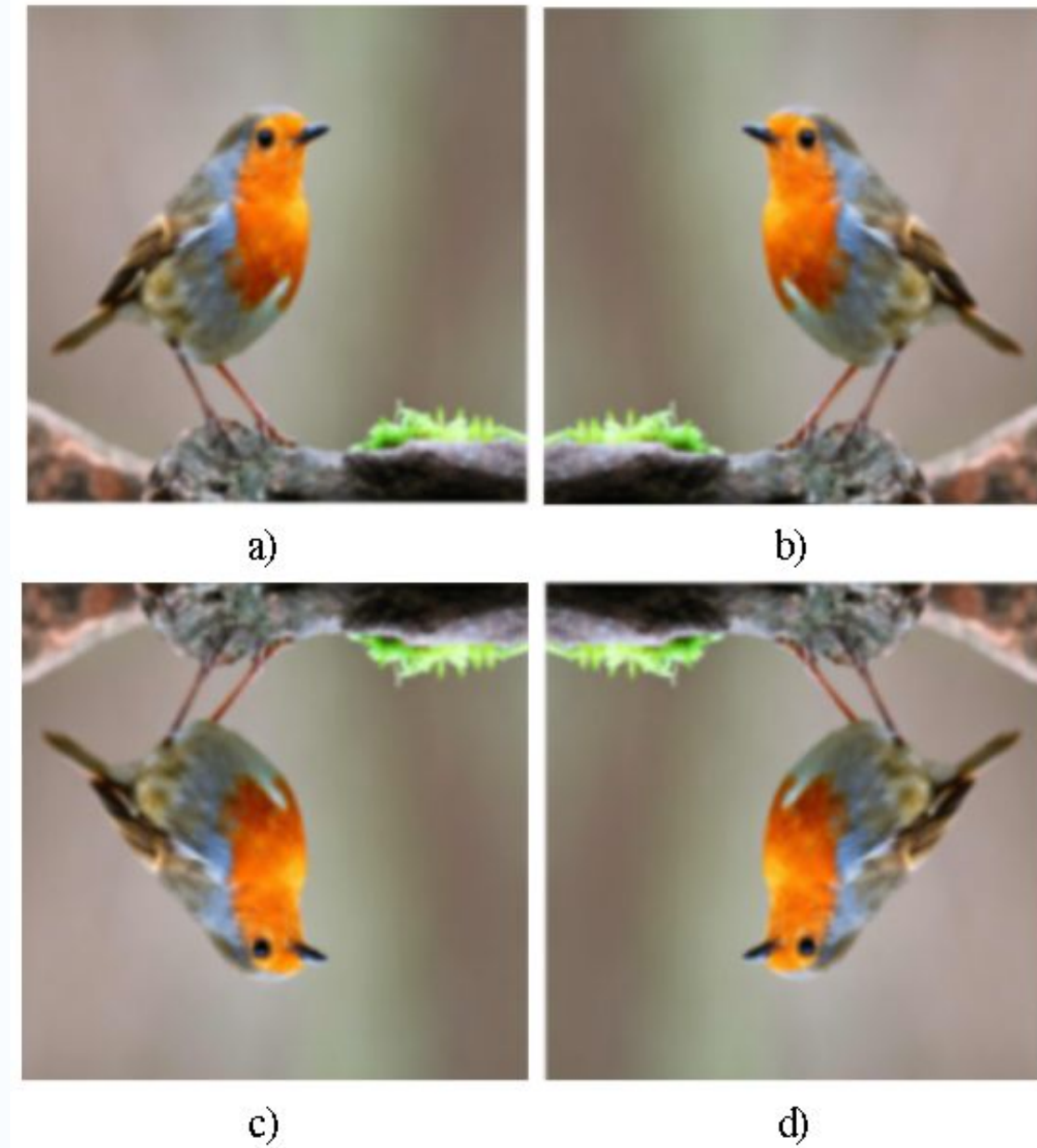
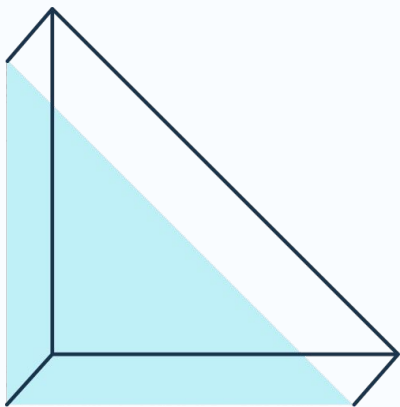
Phương pháp đề xuất

- Tăng cường dữ liệu
- Mô hình 1: DENSENET + CAM
- Mô hình 2: EFFICIENTNET

Tăng cường dữ liệu

- Đây là một trong những phương pháp phổ biến để làm dồi dào nguồn dữ liệu đầu vào cho mô hình trong việc huấn luyện.
- Lật ảnh là một trong những phương pháp tăng cường dữ liệu đơn giản và hiệu quả.

9



Hình 3. Lật ảnh

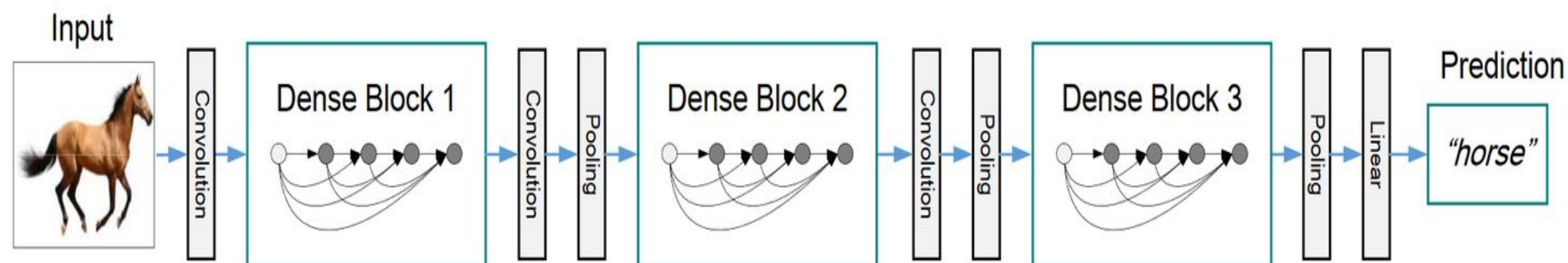
(Hình 3a: ảnh gốc. Hình 3b: lật ảnh ngang. Hình 3c: lật ảnh dọc. Hình 3d: kết hợp lật ảnh ngang và dọc.)

Mô hình 1: DenseNet + CAM

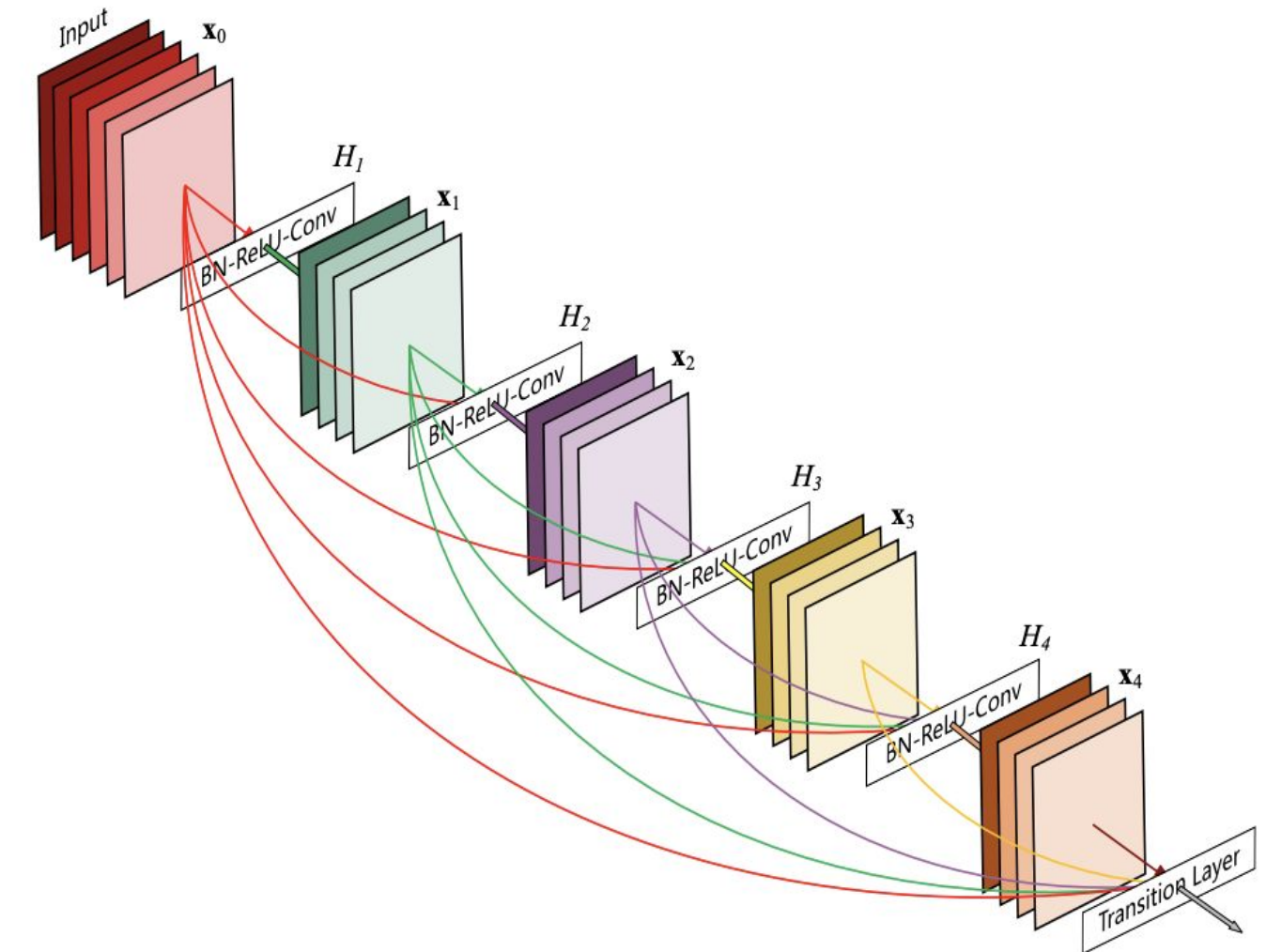
Kiến trúc DenseNet:

- **Dense block:** sử dụng các skip connection (nối tắt), output của layer trước sẽ là input của các layer sau, chúng được **concatenate** lại với nhau.
- **Transition layer** (tầng chuyển tiếp): gồm một layer tích chập giúp giảm độ sâu và một average-pooling layer giúp giảm kích thước dài và rộng.

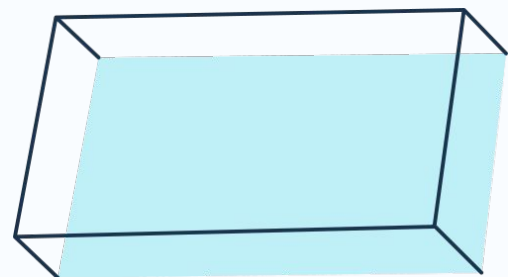
10



Hình 4a: Kiến trúc DenseNet⁵



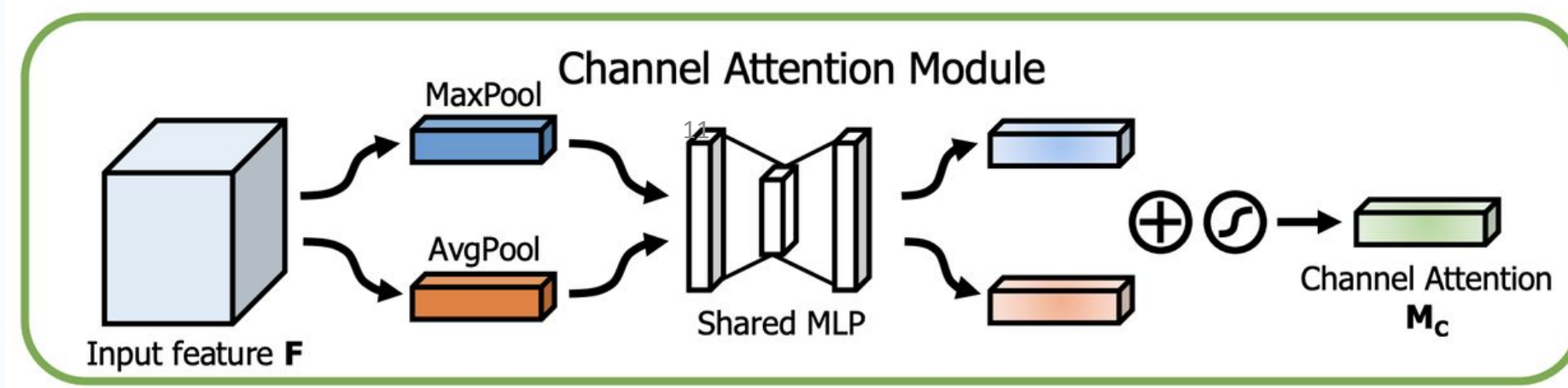
Hình 4b: Kiến trúc DenseBlock⁵



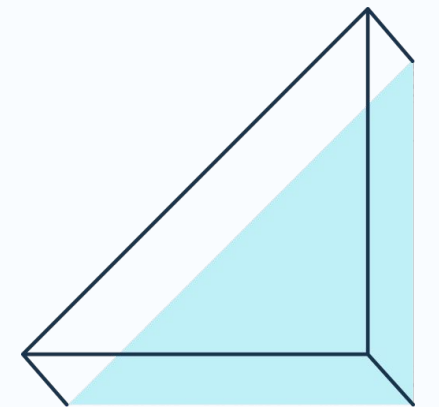
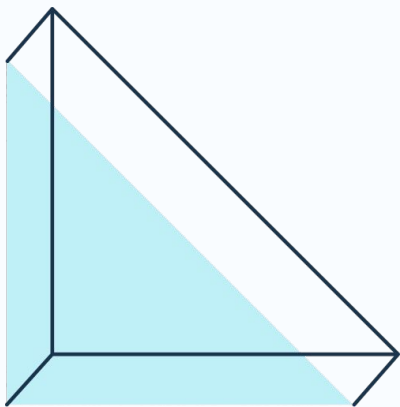
Mô hình 1: DenseNet + CAM

Kiến trúc Channel Attention Module:

- **Channel Attention Module** được tạo ra bằng cách khai thác mối quan hệ giữa các channel của các feature.
- Tổng hợp thông tin không gian của **Feature map** bằng cách sử dụng **average-pooling** và **max-pooling**.
- Sau đó, cả hai biến được đẩy đến một mạng chung (**shared network**) để tạo ra **Channel Attention Map**.
- Hợp nhất các vector feature bằng phép toán **element-wise summation**.



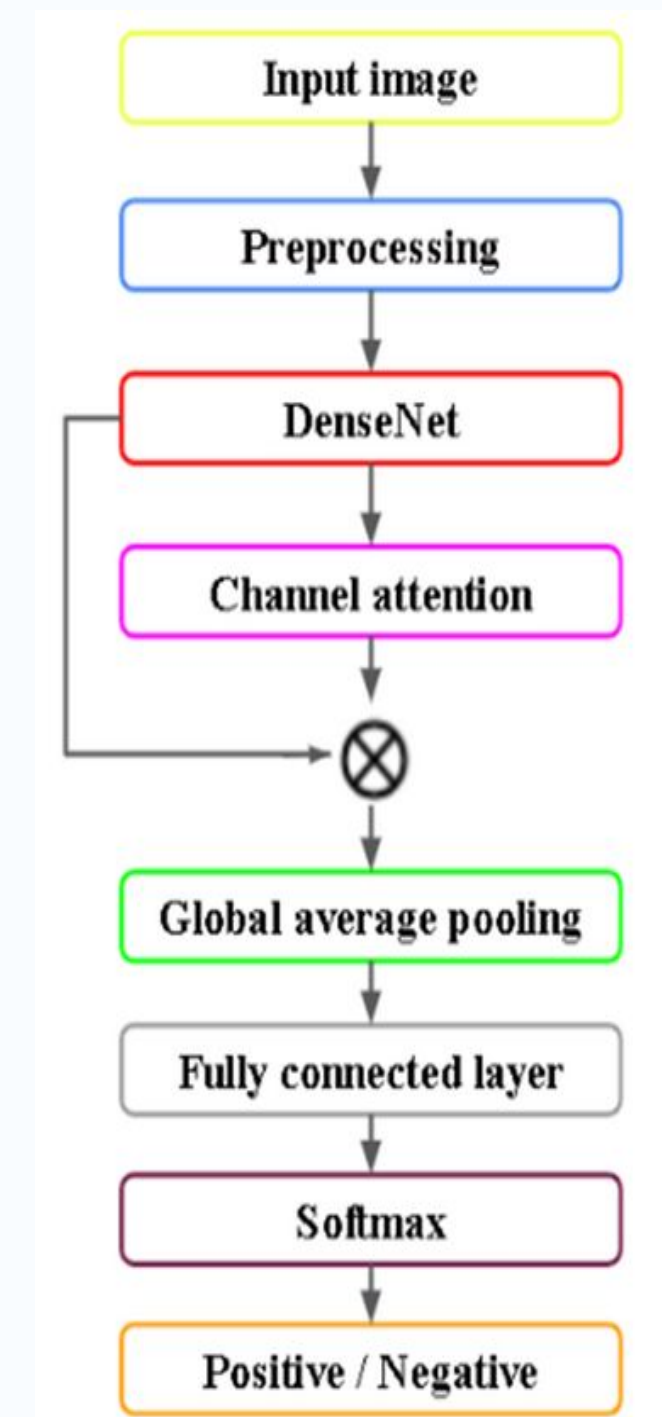
Hình 5: Channel Attention Module⁶



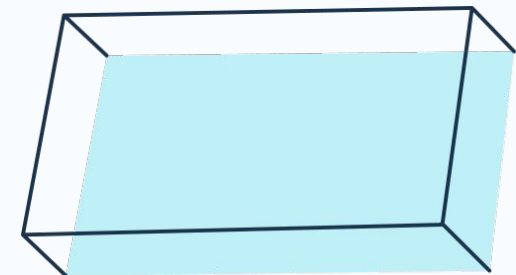
Mô hình 1: DenseNet + CAM

Kiến trúc mô hình 1:

- **Input:** dữ liệu đầu vào.
- **Preprocessing:** thay đổi kích thước và chuẩn hóa ảnh.
- **DenseNet:** rút trích đặc trưng ảnh thành các feature map.
- **Channel attention module:** đánh trọng số trên các feature map.
- **Fully connected:** kết hợp các feature ở các bước trước một cách đa dạng.
- **Softmax:** đưa ra xác suất của ảnh ở 2 lớp **Positive** và **Negative**.



Hình 6. Kiến trúc mô hình 1



Mô hình 2: EfficientNet

Thu phóng mô hình (compound model scaling)

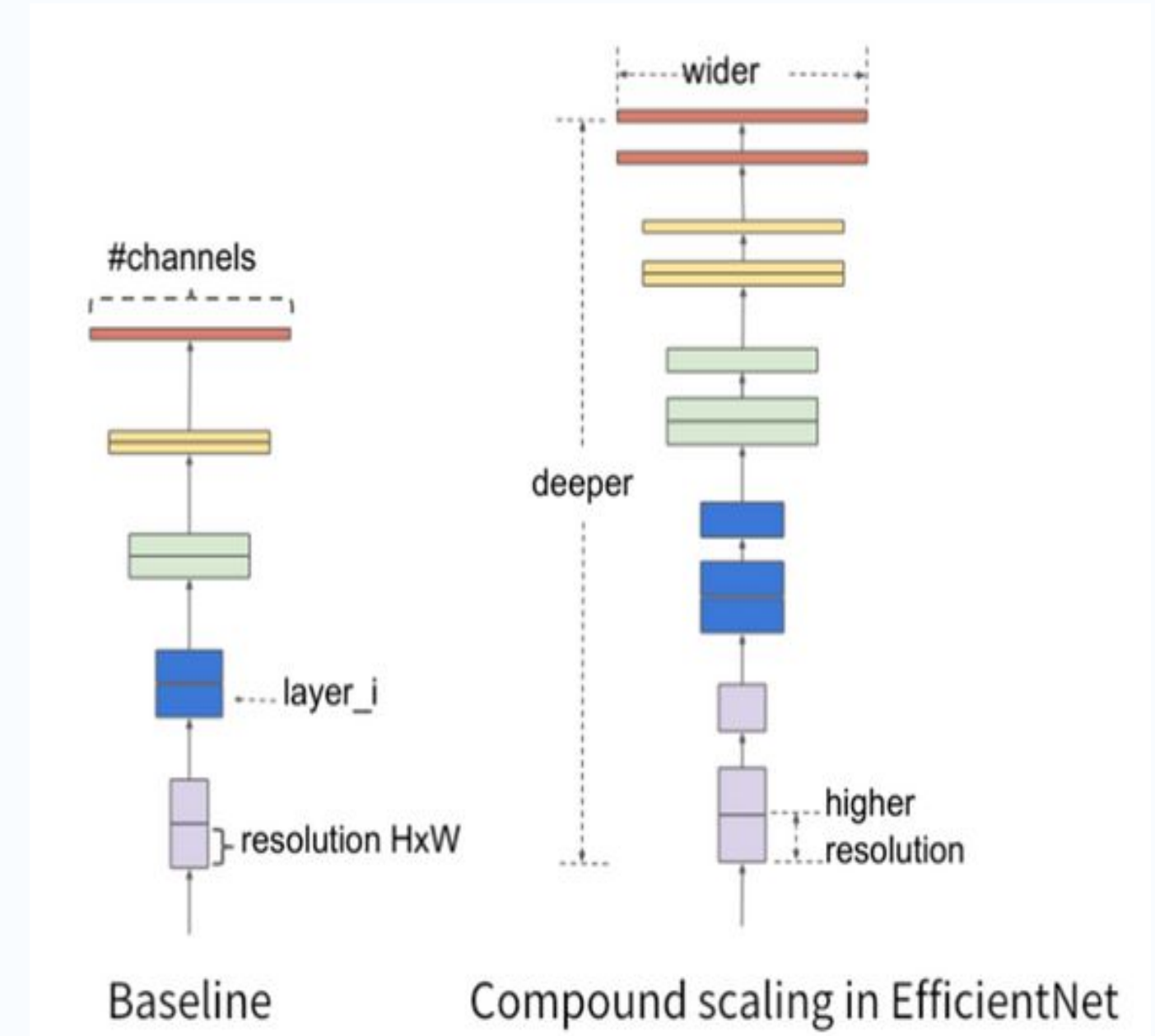
- Thu phóng mô hình đồng thời ở các chiều (**chiều cao, chiều rộng, độ phân giải**) nhằm tăng hiệu suất tổng thể.
- Thu phóng được thực hiện theo nguyên tắc nhất định với công thức:

Trong đó:

α , β , γ : là các hằng số có thể được xác định bằng small grid search.

Φ : là hệ số do người dùng chỉ định để kiểm soát số lượng tài nguyên khác có sẵn để thu phóng mô hình.

$$\begin{aligned} \text{depth} : d &= \alpha^\Phi \\ \text{width} : w &= \beta^\Phi \\ \text{resolution} : r &= \gamma^\Phi \\ s.t. \alpha \cdot \beta^2 \cdot \gamma^2 &\approx 2 \\ \alpha \geq 1, \beta \geq 1, \gamma &\geq 1 \end{aligned}$$



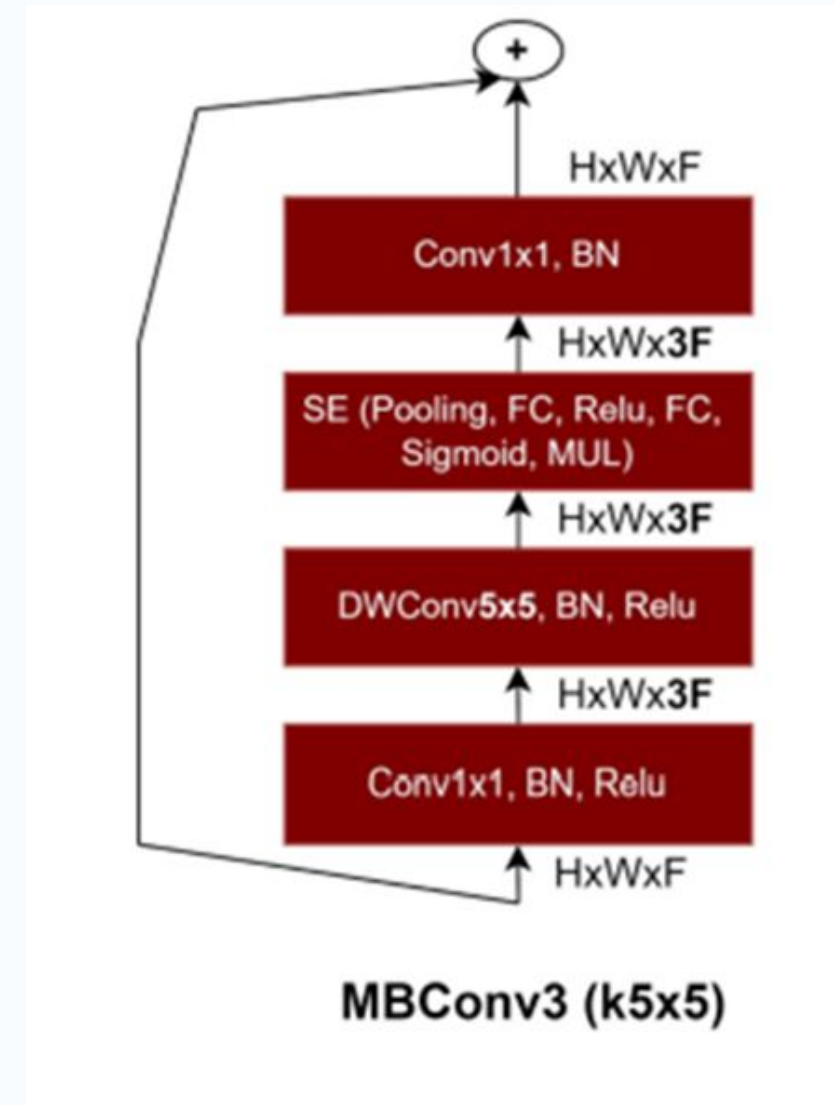
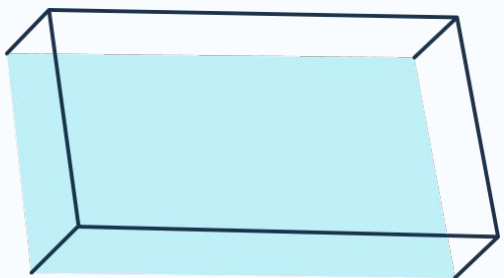
Hình 7. Thu phóng mô hình trong EfficientNet⁷

Mô hình 2: EfficientNet

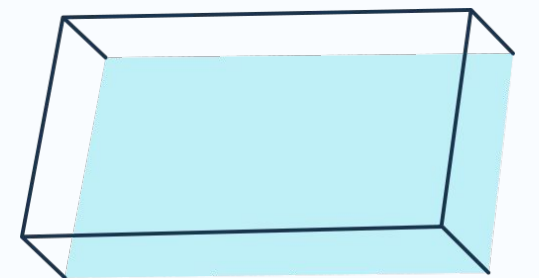
Kiến trúc mạng

- Sử dụng kiến trúc các khối **mobile inverted bottleneck convolutional** (MBConv) để giảm tham số, chú trọng đặc điểm quan trọng, tăng hiệu suất cho mô hình.

14



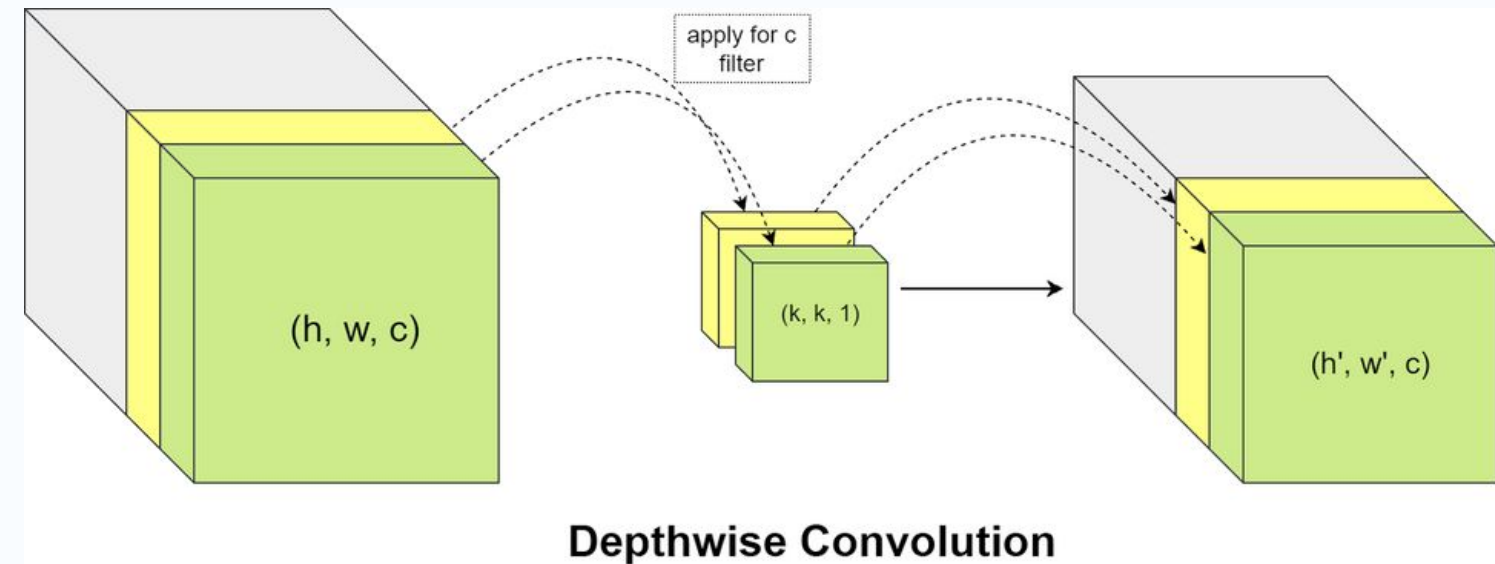
Hình 8. Khối MBConv⁸



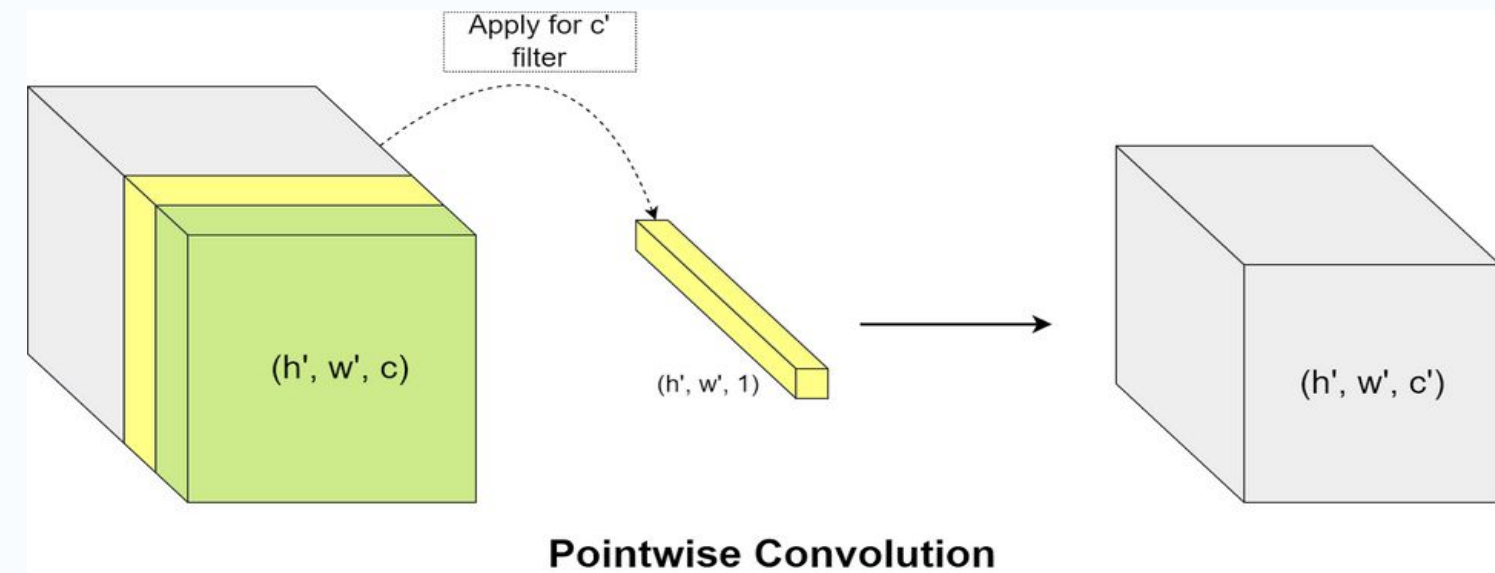
Mô hình 2: EfficientNet

Kiến trúc mạng

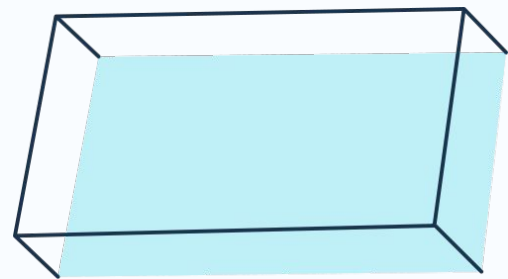
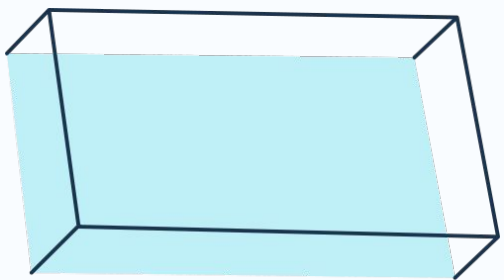
- Khối MB Conv sử dụng **tích chập tách biệt theo chiều sâu** (depthwise separable convolution)
- Giúp giảm lượng tham số đầu vào.
- Giảm chi phí tính toán cho mô hình.



Hình 9. Cách hoạt động của tích chập theo chiều sâu⁹



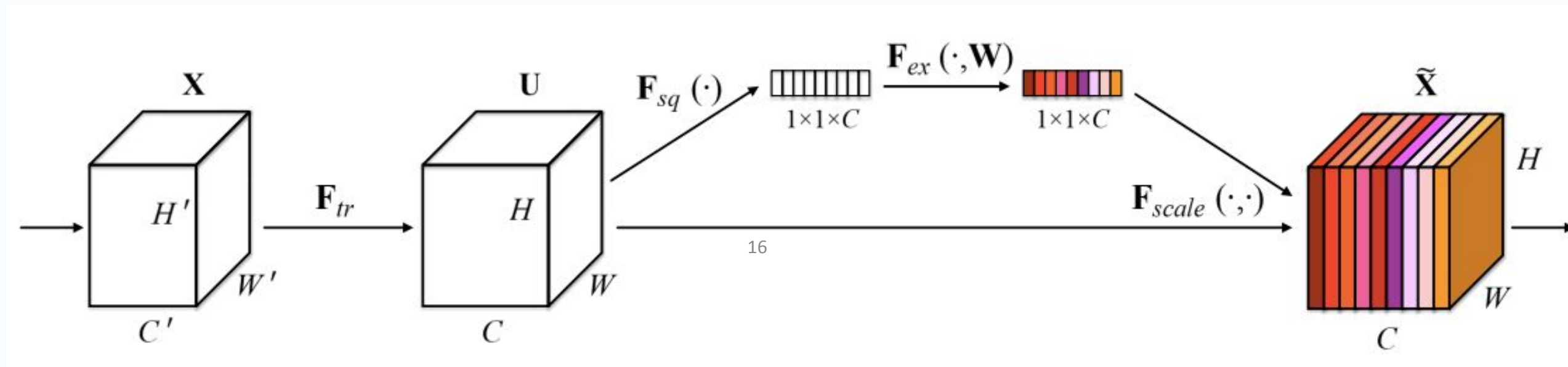
Hình 10. Cách hoạt động của tích chập điểm⁹



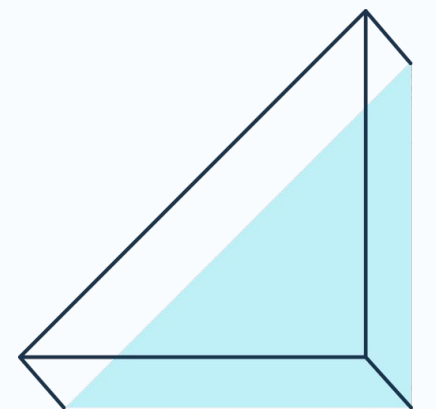
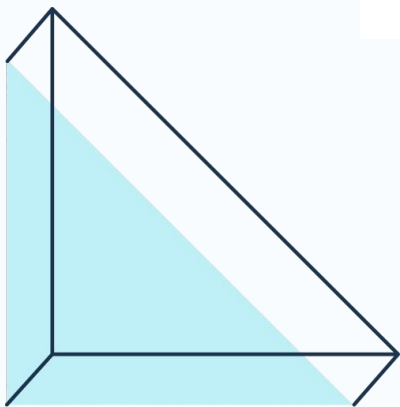
Mô hình 2: EfficientNet

Kiến trúc mạng

- Khối MB Conv còn có một module gọi là **SE** (Squeeze-and-Excitation).
- Khối SE nhằm giúp mô hình đánh trọng số trên các feature map.
- Từ đó phân loại ảnh tốt hơn, cải thiện hiệu suất của mô hình.



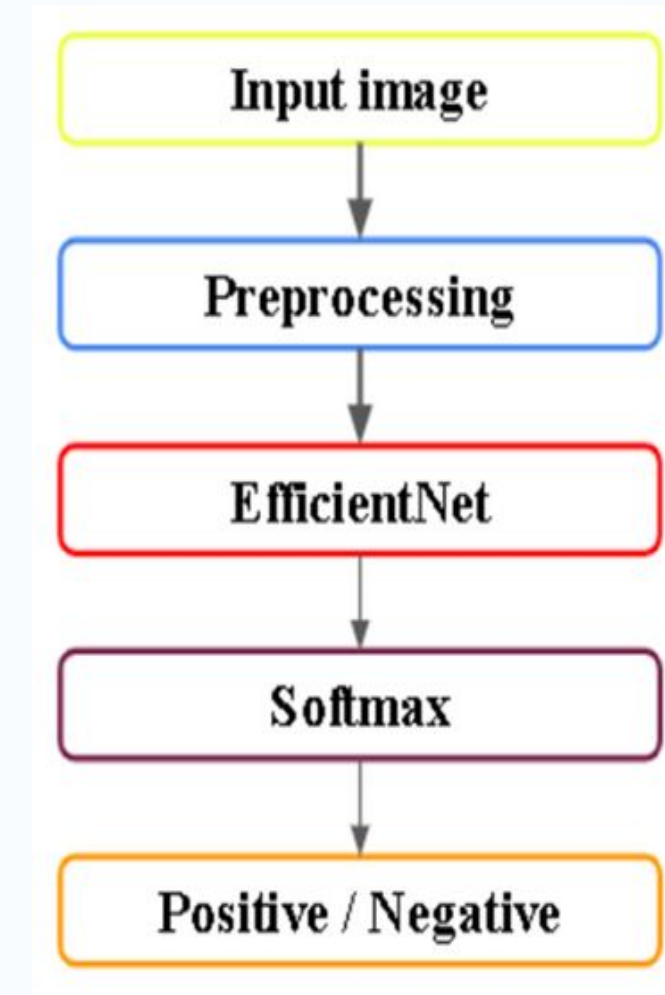
Hình 11: Cách hoạt động của khối SE¹⁰



Mô hình 2: EfficientNet

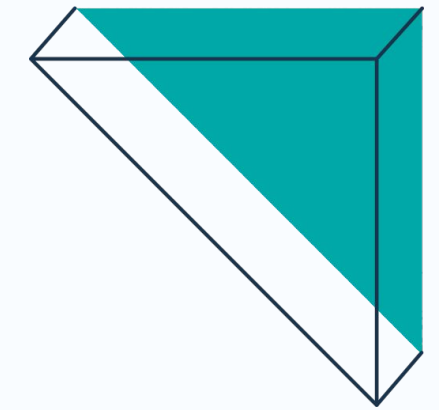
Kiến trúc mô hình 2:

- **Input:** dữ liệu đầu vào.
- **Preprocessing:** thay đổi kích thước và chuẩn hóa ảnh.
- **EfficientNet:** rút trích các đặc trưng của ảnh
- **Softmax:** đưa ra xác suất của ảnh ở 2 lớp Positive và Negative.

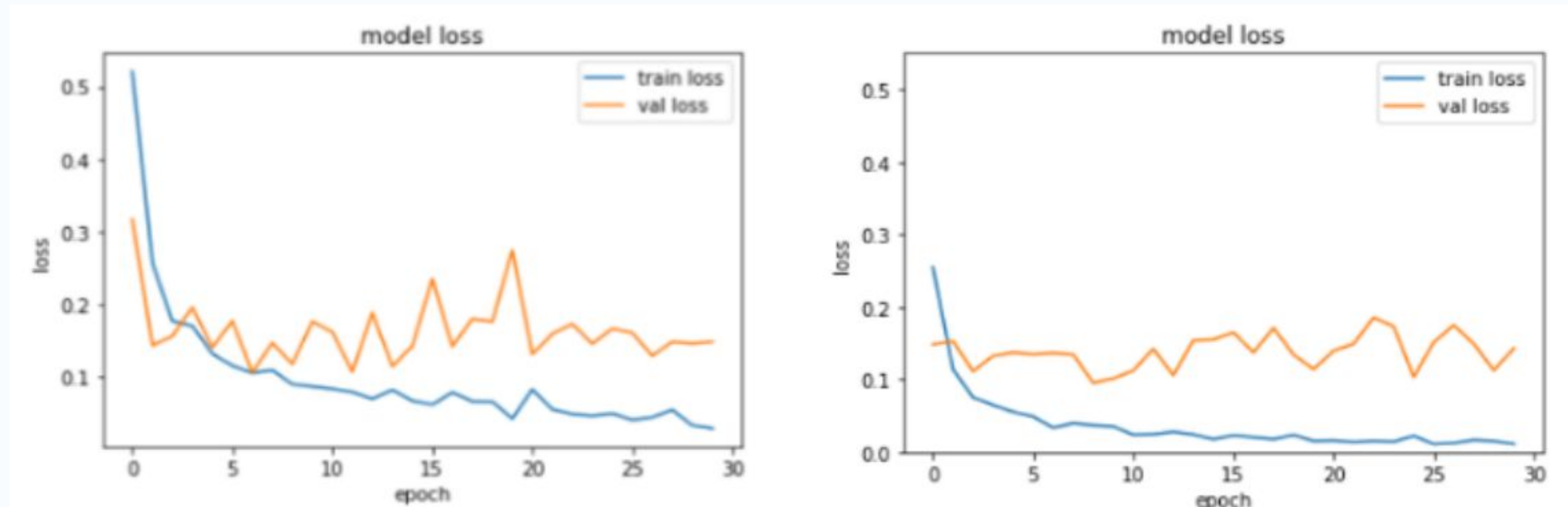


Hình 12. Kiến trúc mô hình 2

Kết quả thực nghiệm

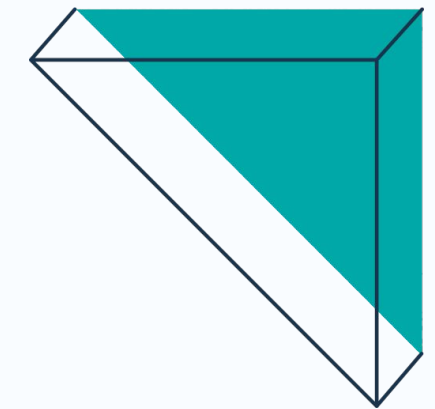


Biểu đồ loss thu được khi huấn luyện



Hình 13. Biểu đồ thể hiện độ lỗi của các mô hình trong quá trình huấn luyện¹⁸
(Hình bên trái thể hiện độ lỗi của mô hình DenseNet201 kết hợp với Channel attention module.
Hình bên phải thể hiện độ lỗi của mô hình EfficientNet B7)

Kết quả thực nghiệm



So sánh kết quả của các mô hình

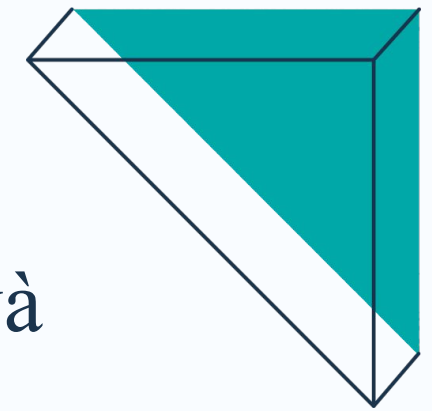
Mô hình	TP	FP	TN	FN
DenseNet201 + CAM	232	27	1313	50
EfficientNetB7	255	16	1324	27

Bảng 2. So sánh kết quả của 2 mô hình trên tập test theo TP, FN, FP, TN.

Mô hình	¹⁹ Precision	Recall	Specificity	AUC
DenseNet201 + CAM	0.8958	0.8227	0.9799	0.9773
EfficientNetB7	0.9410	0.9043	0.9881	0.9874

Bảng 3. So sánh kết quả của 2 mô hình trên tập test theo các độ đo.

Kết quả thực nghiệm

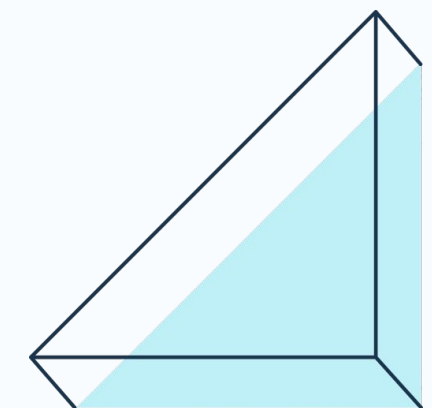
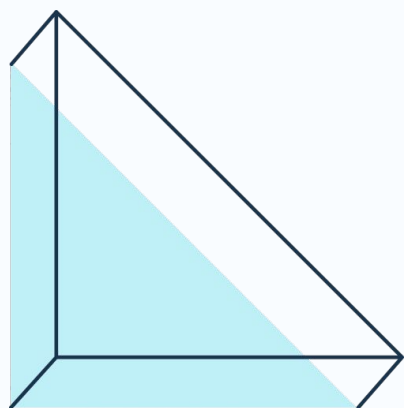


So sánh kết quả với các nghiên cứu trước đó

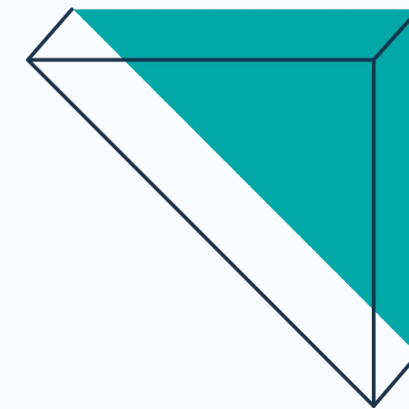
- Chúng tôi tiến hành huấn luyện lại mô hình nhưng lần này là với toàn bộ dữ liệu và so sánh kết quả với các nghiên cứu trước đó trên cùng bộ dữ liệu.
- Với bài báo “Detection of Lung Nodules on CT Images based on the Convolutional Neural Network with Attention Mechanism” [1]

Mô hình	Precision	Recall	Specificity	AUC
Paper [1]	0.950 ²⁰	0.864	0.982	0.992
EfficientNet B7	0.950	0.8759	0.9903	0.993

Bảng 4. So sánh kết quả của 2 mô hình trên tập test theo các độ đo.

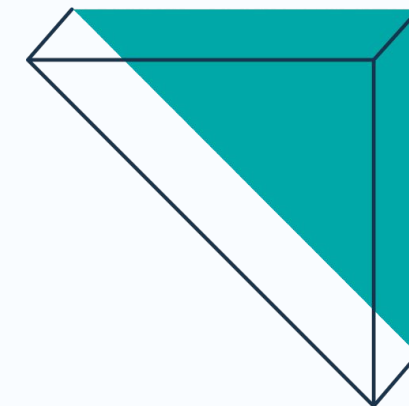


Kết luận

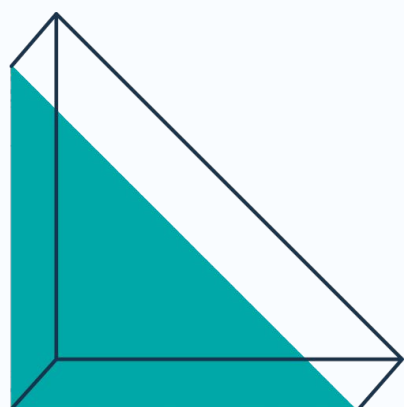


- Để giải quyết bài toán về phân loại khối u trong phổi, một mô hình giải quyết bài toán phức tạp, với độ chính xác cao, tỉ lệ False-Positive thấp, đã được đưa ra. Mô hình dựa vào cách thu phóng theo chiều rộng, chiều sâu và độ phân giải.
- Kết quả của thực nghiệm còn cho thấy rằng việc áp dụng mô hình EfficientNet vào bài toán phân loại khối u trong phổi₂₁ đạt hiệu quả cao hơn kết quả của mô hình Densenet kết hợp CAM module của nhóm và một số nghiên cứu trước đó.

Hướng phát triển



- Triển khai mô hình đã xây dựng lên Web-app.
- Xây dựng mô hình phân loại với độ chính xác cao hơn và thời gian xử lý ngắn hơn.
- Có thể áp dụng một số thuật toán,²² mô hình mới nhất để cải thiện kết quả.



Tài liệu tham khảo

- 
- [1] Lai, Khai Dinh, Nguyen, Thuy Thanh, and Le, Thai Hoang. “Detection of Lung Nodules on CT Images based on the Convolutional Neural Network with Attention Mechanism”
 - [2] Armato, Samuel G, McLennan, Geoffrey, and Bidaut, Luc. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI).
 - [3] Choi, Wook-Jin, Choi, Tae-Sun. "Automated Pulmonary Nodule Detection System in Computed Tomography Images: A Hierarchical Block Classification Approach"
 - [4] Huafenga, Wang et al. “A hybrid CNN feature model for pulmonary nodule malignancy risk differentiation”. In: Journal of X-Ray Science and Technology (2018).
 - [5] Gao Huang, Zhuang Liu, Laurens van der Maaten, Kilian Q. Weinberger, Densely Connected Convolutional Networks.
 - [6] Sanghyun Woo Jongchan Park, Joon-Young Lee In So Kweon. “CBAM: Convolutional Block Attention Module”. In: ECCV 2018 (2018).
 - [7] Mingxing Tan, Quoc V. Le. “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks”. In: International Conference on Machine Learning, 2019 (2019).
 - [8] Mingxing Tan Bo Chen, Ruoming Pang Vijay Vasudevan Mark Sandler Andrew Howard Quoc V. Le. “MnasNet: Platform-Aware Neural Architecture Search for Mobile”. In: CVPR 2019 (2019).
 - [9] Pham Dinh Khanh. Mobilenet model. url: <https://phamdinhkhanh.github.io/2020/09/19/MobileNet.html>
 - [10] Bui Quang Manh. Những mô hình trợ thủ đặc lực trong các mô hình Deep learning [Phần 1]