

PHÂN LOẠI ẢNH CÓ KHỐI U VÀ KHÔNG CÓ KHỐI U Ở PHỔI

TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN, ĐHQG-HCM - KHOA CÔNG NGHỆ THÔNG TIN
GIÁO VIÊN HƯỚNG DẪN: PGS. TS LÊ HOÀNG THÁI
NHÓM SINH VIÊN THỰC HIỆN: HỨA PHÚ THÀNH - 18120563, HÀ MINH TOÀN - 18120597, TRẦN NGỌC TỊNH- 18120599, TRẦN LUẬT VY - 18120656



GIỚI THIỆU

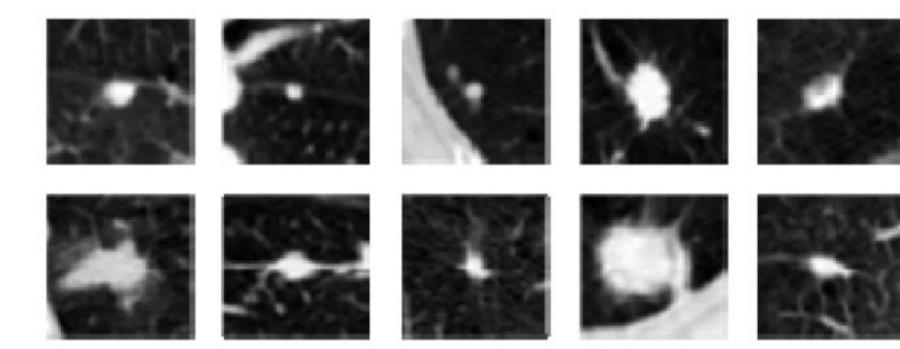
- Lý do nghiên cứu:**
Ung thư phổi là một trong những loại ung thư phổ biến nhất trên thế giới hiện nay. Dù là căn bệnh nguy hiểm nhưng có rất nhiều trường hợp bệnh nhân không có bất kỳ triệu chứng lâm sàng nào và khi có các triệu chứng của bệnh thì ung thư đã vào giai đoạn cuối. Hiện nay, chúng ta có thể phát hiện các bệnh liên quan tới phổi, đặc biệt là ung thư bằng biện pháp chụp cắt lớp phổi (chụp CT phổi).
- 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”**. Qua đó, hướng đến việc giảm thời gian chẩn đoán thủ công cho các bác sĩ và đồng thời mang lại độ chính xác cao trong việc phân loại.
- Hướng giải quyết đã có:**
 - Sử dụng mạng CNN kết hợp với Attention Mechanism do Lai Đình Khải và các cộng sự.
 - Các tác giả Wook Jin Choi và Tea Sun Choi tiếp cận theo hướng sử dụng Hierarchical Block kết hợp với SVM để phân loại.

DỮ LIỆU

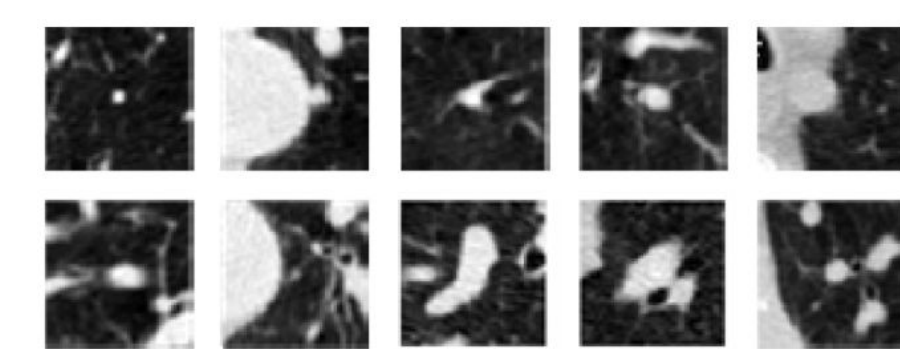
- Tập dữ liệu sử dụng:** nguồn gốc từ tập dữ liệu LIDC-IDRI được tác giả Swetha Subramanian điều chỉnh lại. Tập dữ liệu gồm 2 lớp là Positive và Negative.
 - Positive: các nốt sần ở phổi có đường kính lớn hơn hoặc bằng 3mm.
 - Negative: Các nốt sần ở phổi có đường kính bé hơn 3mm và các tổn thương khác ở phổi không phải là nốt sần có đường kính lớn hơn 3mm.

Bảng 1: Số lượng các lớp trên các tập dữ liệu.

Tập dữ liệu	Postive	Negative
Train	845	4342
Val	224	1073
Test	282	1340



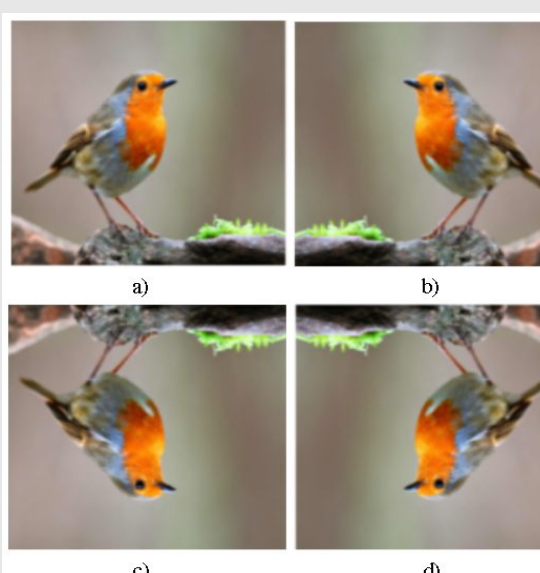
Hình 1: Một số ảnh thuộc lớp Positive.



Hình 2: Một số ảnh thuộc lớp Negative.

TĂNG CƯỜNG DỮ LIỆU

- Tăng cường dữ liệu là một trong những phương pháp phổ biến để làm dồi dào nguồn dữ liệu đầu vào cho mô hình trong việc huấn luyện.
- Chúng tôi sử dụng phương pháp lật ảnh là một trong những phương pháp tăng cường dữ liệu đơn giản để có thể cải thiện hiệu suất của mô hình
- 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.)



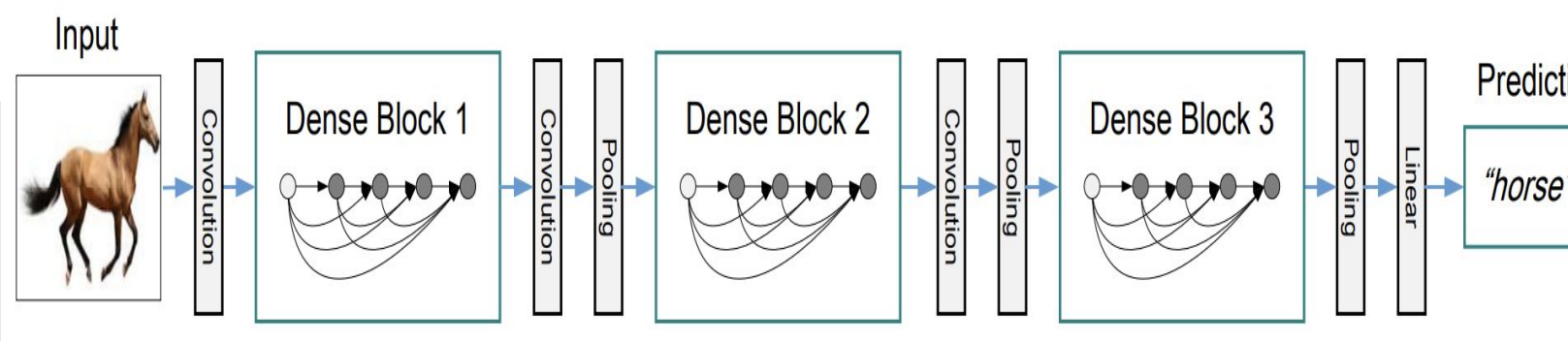
Hình 3: Lật ảnh

MÔ HÌNH 1: DENSENET + CAM

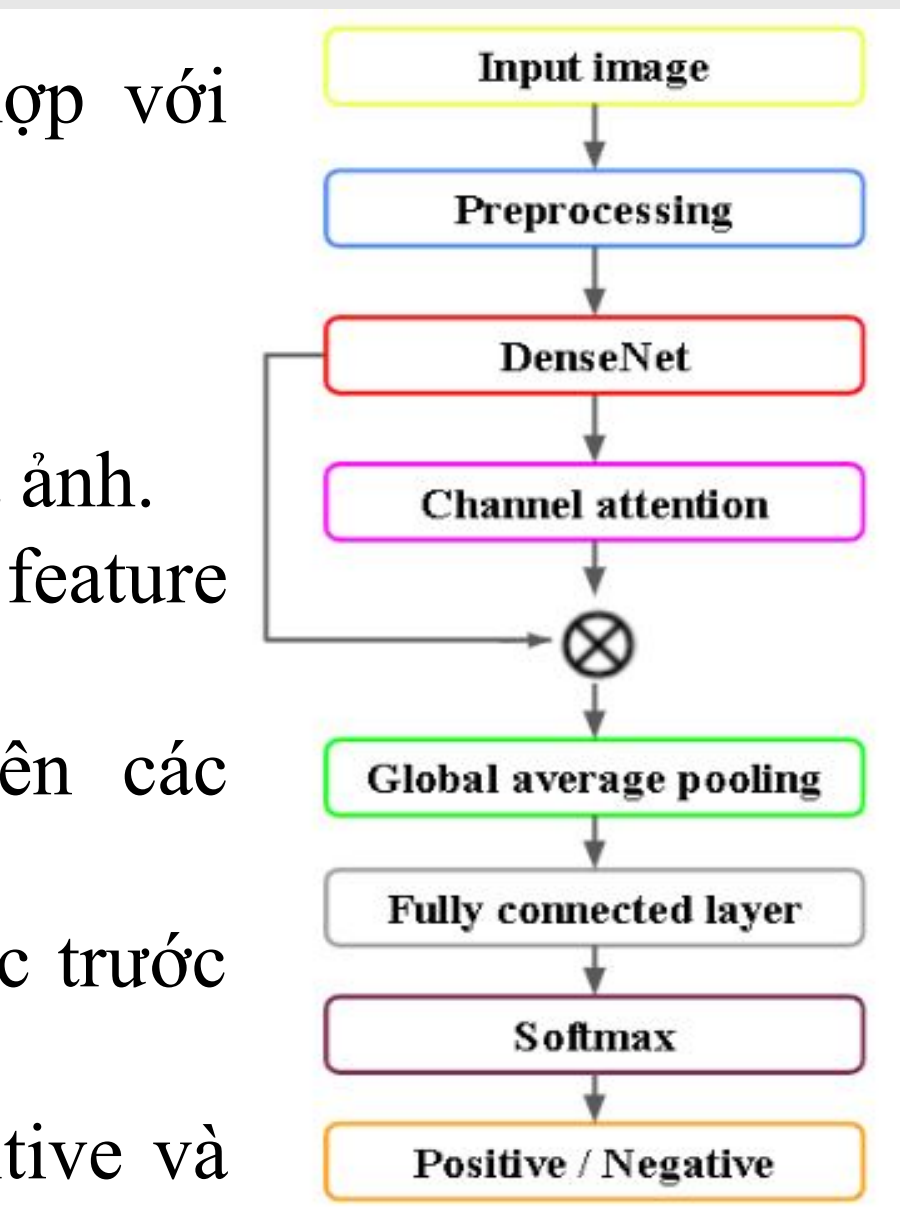
Nhóm chúng tôi sử dụng mạng DenseNet kết hợp với Channel Attention Module.

Hình 4 mô tả kiến trúc của mô hình 1:

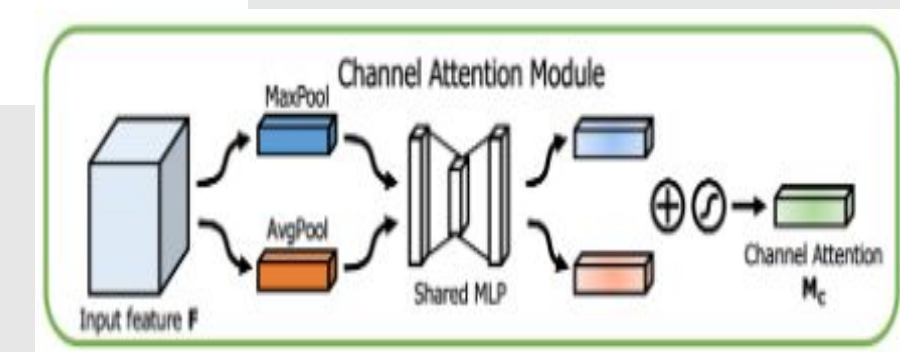
- Input: dữ liệu đầu vào.
- Preprocessing: thay đổi kích thước và chuẩn hóa ảnh.
- DenseNet: rút trích đặc trưng ảnh thành các feature map.
- Channel attention module: đánh trọng số trên các feature map.
- Fully connected: kết hợp các feature ở các bước trước một cách đa dạng.
- Softmax: đưa ra xác suất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



Hình 5: Kiến trúc DenseNet [1].



Hình 4: Mô hình 1



Hình 6: Channel attention module [2].

MÔ HÌNH 2: EFFICIENTNET

Nhóm chúng tôi sử dụng mạng EfficientNet cho mô hình 2.

Những điểm nổi bật của EfficientNet nằm ở:

- Cơ chế thu phóng mô hình.
 - Kiến trúc mạng EfficientNet.
- Hình 2 mô tả kiến trúc của 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.

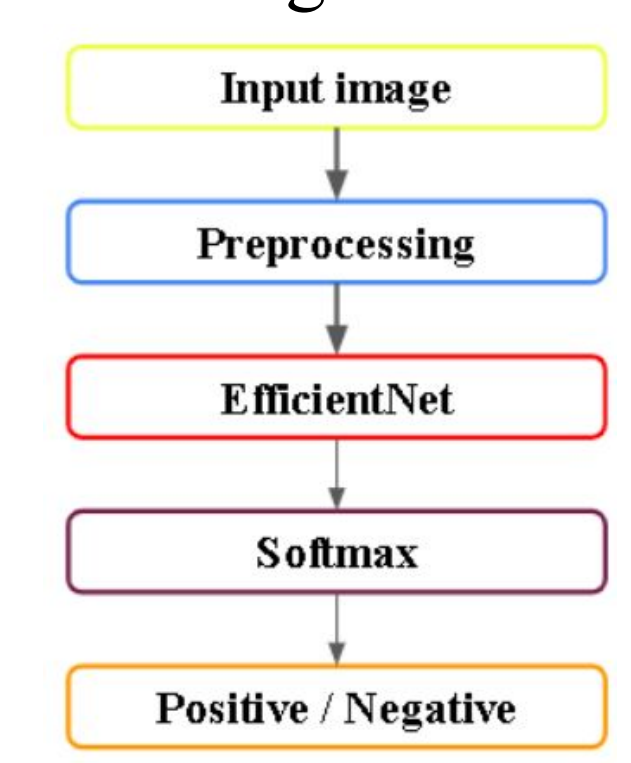
Nguyên tắc thu phóng:

Depth: $d = \alpha^\Phi$
Width: $w = \beta^\Phi$
Resolution: $r = \gamma^\Phi$
s.t. $\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2$
 $\alpha \geq 1, \beta \geq 1, \gamma \geq 1$

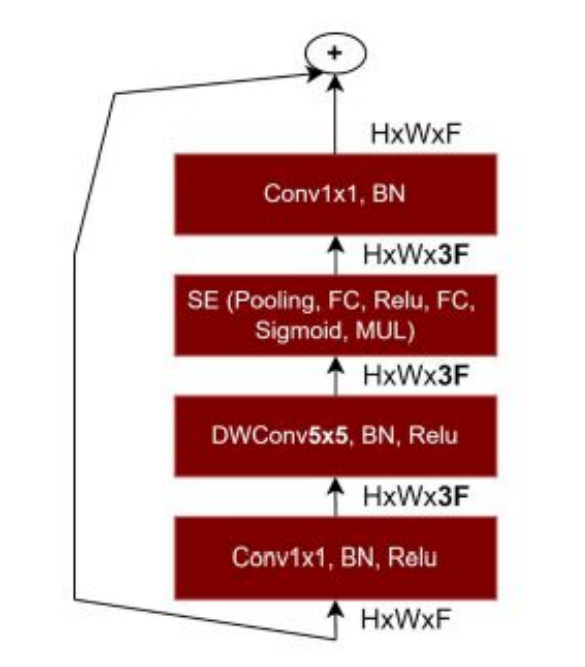
Trong đó:

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.



Hình 7: Mô hình 2.

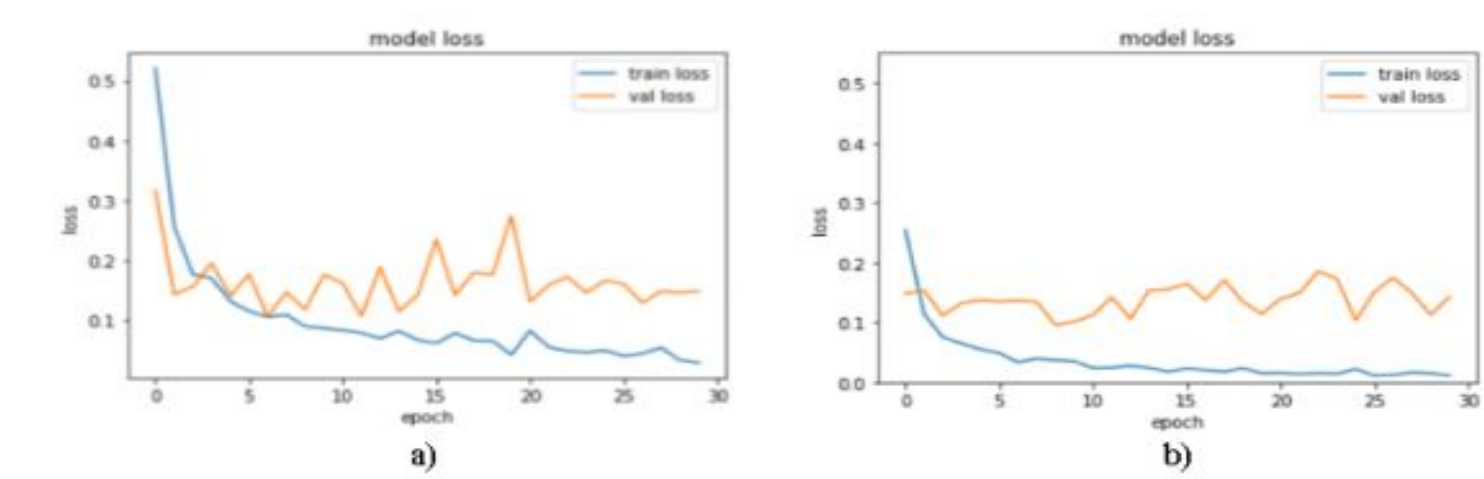


Hình 8: Kiến trúc khối MBConv3.

KẾT QUẢ THỰC NGHIỆM

Nhóm chúng tôi tiến hành huấn luyện và kiểm thử mô hình.

Các kết quả thu được ở hình 9, bảng 2 và bảng 3.



Hình 9: Biểu đồ loss của 2 mô hình train tập train và validation. (Hình a: biểu đồ loss của mô hình 1. Hình b: biểu đồ loss của mô hình 2.)

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

Mô hình	TP	FP	TN	FN
DenseNet201 + CAM	232	27	1313	50
EfficientNet B7	255	16	1324	27

Bảng 3: So sánh kết quả của 2 mô hình trên tập test theo các độ đo Precision, Recall, Specificity, Auc.

Mô hình	Precision	Recall	Specificity	Auc
DenseNet201 + CAM	0.8958	0.8227	0.9799	0.9773
EfficientNet B7	0.9410	0.9043	0.9881	0.9874

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” [3])

Bảng 4: So sánh kết quả của 2 mô hình trên tập test theo các độ đo Precision, Recall, Specifict, Auc.

Mô hình	Precision	Recall	Specificity	Auc
Paper [3]	0.950	0.864	0.982	0.992
EfficientNet B7	0.950	0.8759	0.9903	0.993

KẾT LUẬN

- Mô hình nhóm xây dựng cho kết quả với độ chính xác cao, tỉ lệ False-Positive thấp.
- Kết quả của thực nghiệm còn cho thấy rằng việc áp dụng mô hình EfficientNet vào bài toán phân loại khối u trong phổi đạt hiệu quả cao hơn kết quả của mô hình Densenet kết hợp Channel attention module của nhóm và một số nghiên cứu trước đó

THAM KHẢO

[1] Phạm Đình Khanh. Bài 38 - Các kiến trúc CNN hiện đại. url: <https://phamdingkhanh.github.io/2020/05/31/CNNHistory.html> (visited on 06/23/2022).
[2] Sanghyun Woo Jongchan Park, Joon-Young Lee In So Kweon. “CBAM: Convolutional Block Attention Module”. In: ECCV 2018 (2018).
[3] 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”. In: Published by International Association of Educators and Researchers (IAER) (2021).