

TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP TP HỒ CHÍ MINH
KHOA CÔNG NGHỆ THÔNG TIN



ĐỒ ÁN CUỐI KÌ
DEEP LEARNING

Người thực hiện: **HỒ VÕ HOÀNG DUY – 20087541**

VÕ QUỐC HUY – 20081001

Lớp : **420300411601**

Khoá : **16**

Người hướng dẫn: **TS BÙI THANH HÙNG**

THÀNH PHỐ HỒ CHÍ MINH, NĂM 2023

TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP TP HỒ CHÍ MINH
KHOA CÔNG NGHỆ THÔNG TIN



ĐỒ ÁN CUỐI KÌ
DEEP LEARNING

Người thực hiện: **HỒ VÕ HOÀNG DUY- 20087541**

VÕ QUỐC HUY- 20081001

Lớp : **420300411601**

Khoá : **16**

Người hướng dẫn: **TS. BÙI THANH HÙNG**

THÀNH PHỐ HỒ CHÍ MINH, NĂM 2023

LỜI CẢM ƠN

Chúng tôi xin gửi lời cảm ơn chân thành nhất đến quý thầy TS.Bùi Thanh Hùng – Giảng viên đã tận tình hướng dẫn cho chúng tôi trong suốt quá trình học tập môn Học Sâu và tạo điều kiện cho chúng tôi làm Đồ án này. Dưới sự hướng dẫn của Thầy chúng tôi được tiếp cận với những kiến thức chuyên môn sâu, rộng trong lĩnh vực học sâu. Qua đó có thể hoàn thành Đồ án này một cách thuận lợi.

Một lần nữa chúng tôi xin gửi lời cảm ơn sâu sắc đến Thầy. Chúc Thầy thật nhiều sức khỏe, thành công trong công việc và trong cuộc sống. Xin chân thành cảm ơn Thầy.

ĐỒ ÁN ĐƯỢC HOÀN THÀNH

TẠI TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP TP HỒ CHÍ MINH

Tôi xin cam đoan đây là sản phẩm đồ án của riêng tôi / chúng tôi và được sự hướng dẫn của TS. Bùi Thanh Hùng. Các nội dung nghiên cứu, kết quả trong đề tài này là trung thực và chưa công bố dưới bất kỳ hình thức nào trước đây. Những số liệu trong các bảng biểu phục vụ cho việc phân tích, nhận xét, đánh giá được chính tác giả thu thập từ các nguồn khác nhau có ghi rõ trong phần tài liệu tham khảo.

Ngoài ra, trong đồ án còn sử dụng một số nhận xét, đánh giá cũng như số liệu của các tác giả khác, cơ quan tổ chức khác đều có trích dẫn và chú thích nguồn gốc.

Nếu phát hiện có bất kỳ sự gian lận nào tôi xin hoàn toàn chịu trách nhiệm về nội dung đồ án của mình. Trường đại học Công nghiệp TP Hồ Chí Minh không liên quan đến những vi phạm tác quyền, bản quyền do tôi gây ra trong quá trình thực hiện (nếu có).

TP. Hồ Chí Minh, ngày tháng năm

Tác giả

(ký tên và ghi rõ họ tên)

Hồ Võ Hoàng Duy

Võ Quốc Huy

PHẦN ĐÁNH GIÁ CỦA GIẢNG VIÊN

Tp. Hồ Chí Minh, ngày tháng năm

(kí và ghi họ tên)

TÓM TẮT

Bài toán dự đoán mức độ nghiêm trọng của ung thư tuyến tiền liệt là một bài toán quan trọng trong lĩnh vực y học, giúp các bác sĩ đưa ra quyết định điều trị phù hợp cho bệnh nhân. Mức độ nghiêm trọng của ung thư tuyến tiền liệt được đánh giá dựa trên hệ thống phân loại ISUP grade. Sử dụng các mô hình học sâu như CNN, EfficientNet,... được đào tạo trên bộ dữ liệu PANDA Challenge gồm các hình ảnh mô tuyến tiền liệt để giải quyết bài toán này một cách hiệu quả. Bộ dữ liệu chia thành hai phần là train để huấn luyện mô hình và test để đánh giá chất lượng mô hình. Mô hình CNN, EfficientNet đạt được độ chính xác khá cao trên tập test điều này có nghĩa là mô hình có thể dự đoán mức độ nghiêm trọng của ung thư tuyến tiền liệt trong rất nhiều các trường hợp khác nhau. Kết quả cho thấy các mô hình học sâu có thể được sử dụng hiệu quả trong việc đánh giá mức độ nghiêm trọng của ung thư tuyến tiền liệt một cách hiệu quả. Việc dự đoán mức độ nghiêm trọng của ung thư tuyến tiền liệt có ý nghĩa quan trọng trong y học. Sử dụng các mô hình học sâu để giải quyết vấn đề này có thể giúp cải thiện hiệu quả chuẩn đoán và điều trị ung thư tuyến tiền liệt.

MỤC LỤC

LỜI CẢM ƠN	i
PHẦN ĐÁNH GIÁ CỦA GIẢNG VIÊN	iii
TÓM TẮT	iv
MỤC LỤC	1
DANH MỤC CHỮ VIẾT TẮT.....	3
DANH MỤC CÁC HÌNH VẼ.....	4
DANH MỤC CÁC BẢNG.....	5
1.1. Giới thiệu bài toán.....	6
1.2. Phân tích yêu cầu bài toán.....	7
1.2.1. Yêu cầu bài toán	7
1.2.2. Các phương pháp giải quyết bài toán	8
1.2.2.1. Phương pháp học máy	8
1.2.2.2. Phương pháp học sâu.....	9
1.2.3. Phương pháp đề xuất giải quyết bài toán.....	10
1.3. Phương pháp giải quyết bài toán.....	11
1.3.1. Mô hình tổng quát.....	11
1.3.2. Đặc trưng của mô hình đề xuất.....	12
1.3.2.1. Deep Learning Model.....	12
1.3.2.2. Lưu lại các tham số mô hình	16
1.3.2.3. So sánh và đánh giá kết quả trên tập Test	16

1.4.	Thực nghiệm	17
1.4.1.	Dữ liệu	17
1.4.2.	Xử lý dữ liệu	18
1.4.3.	Công nghệ sử dụng	19
1.4.4.	Cách đánh giá.....	19
1.4.4.1.	Quadratic Weighted Kappa	19
1.4.4.2.	Binary Cross Entropy	20
1.5.	Kết quả đạt được	21
1.5.1.	Hyperparameter	21
1.5.2.	Kết quả	21
1.6.	Kết luận	25
1.6.1.	Kết quả đạt được	25
1.6.2.	Hạn chế	26
1.6.3.	Hướng phát triển	26
	LÀM VIỆC NHÓM	27
	TÀI LIỆU THAM KHẢO.....	28
	PHỤ LỤC	31
	TỰ ĐÁNH GIÁ.....	32

DANH MỤC CHỮ VIẾT TẮT

AI: Artificial Intelligence

CNN: Convolution Neural Network

PANDA: Prostate cANcer graDe Assessment

DANH MỤC CÁC HÌNH VẼ

Hình 1. Giới thiệu ISUP grade	6
Hình 2. Kiến trúc VGG19	12
Hình 3. Mô hình InceptionV3	13
Hình 4 Mô hình ResNet50.....	15
Hình 5. Mô hình DenseNet.....	16
Hình 6. Mô hình tổng quát	11
Hình 7. Accuracy giữa các mô hình	22
Hình 8. Kappa score giữa các mô hình.....	22
Hình 9. Loss giữa các mô hình	23

DANH MỤC CÁC BẢNG

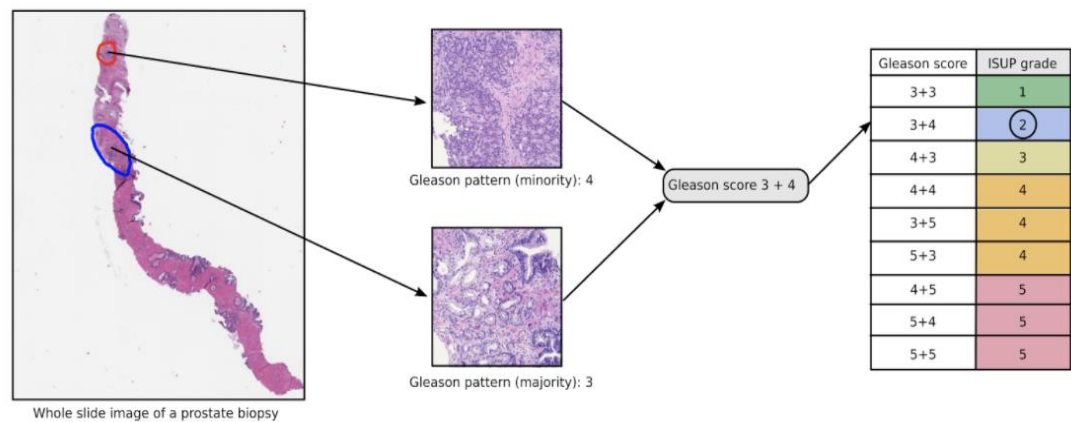
Bảng 1. Mô tả công nghệ sử dụng	19
Bảng 2. Kết quả thực nghiệm	24
Bảng 3. Kết quả so sánh	25

PROSTATE CANCER GRADE ASSESSMENT (PANDA) CHALLENGE

1.1. Giới thiệu bài toán

Bài toán dự đoán mức độ nghiêm trọng của ung thư tuyến tiền liệt là một bài toán quan trọng trong lĩnh vực y học. Ung thư tuyến tiền liệt là một loại ung thư phổ biến ở nam giới, chiếm khoảng 15% tất cả các ca ung thư ở nam giới. Ung thư tuyến tiền liệt có thể được phân loại thành nhiều mức độ nghiêm trọng khác nhau, tùy thuộc vào mức độ xâm lấn của khối u. Mức độ nghiêm trọng của ung thư tuyến tiền liệt được đánh giá dựa trên hệ thống ISUP grade [1], bao gồm 6 mức độ:

- ISUP grade 0: Không bị ung thư tiền liệt.
- ISUP grade 1: Ung thư tiền liệt có tính chất rất thấp, có khả năng di căn thấp.
- ISUP grade 2: Ung thư tiền liệt có tính chất thấp, có khả năng di căn trung bình.
- ISUP grade 3: Ung thư tiền liệt có tính chất trung bình, có khả năng di căn cao.
- ISUP grade 4: Ung thư tiền liệt có tính chất cao, có khả năng di căn rất cao.
- ISUP grade 5: Ung thư tiền liệt có tính chất rất cao và nghiêm trọng nhất, có khả năng di căn gần như chắc chắn.



Hình 1.1. Giới thiệu ISUP grade

Bài toán dự đoán mức độ nghiêm trọng của ung thư tuyến tiền liệt từ bộ dữ liệu PANDA Challenge [2]. Bộ dữ liệu này bao gồm 11.000 mẫu bệnh học tuyến tiền liệt, được đánh giá mức độ nghiêm trọng theo hệ thống ISUP grade.

Bài toán dự đoán mức độ nghiêm trọng của ung thư tuyến tiền liệt từ bộ dữ liệu PANDA Challenge là một bài toán đầy thách thức. Các mẫu bệnh học tuyến tiền liệt thường có kích thước nhỏ và khó phân biệt giữa các tế bào ung thư và tế bào bình thường. Ngoài ra, mức độ nghiêm trọng của ung thư có thể bị ảnh hưởng bởi nhiều yếu tố khác nhau, bao gồm: Tình trạng sức khỏe tổng thể của bệnh nhân, tuổi tác của bệnh nhân và các yếu tố di truyền.

Bài toán dự đoán mức độ nghiêm trọng của ung thư tuyến tiền liệt có ý nghĩa quan trọng trong việc chẩn đoán điều trị bệnh ung thư tuyến tiền liệt. Cải thiện độ chính xác trong chẩn đoán giúp các bác sĩ đưa ra chẩn đoán chính xác hơn, từ đó có phương pháp điều trị phù hợp. Bên cạnh đó việc chẩn đoán chính xác cũng sẽ làm giảm chi phí, thời gian điều trị và tăng khả năng sống sót cho các bệnh nhân.

Tóm lại, bài toán dự đoán mức độ nghiêm trọng của ung thư tuyến tiền liệt là một bài toán đầy thách thức nhưng có ý nghĩa quan trọng trong việc điều trị và chẩn đoán bệnh ung thư tuyến tiền liệt.

1.2. Phân tích yêu cầu bài toán

1.2.1. Yêu cầu bài toán

Nhiệm vụ của bài toán này là phát triển một thuật toán AI có thể phân loại các khối ung thư tuyến tiền liệt thành các cấp độ khác nhau, từ cấp độ thấp nhất đến cấp độ cao nhất. Các cấp độ của ung thư tuyến tiền liệt được xác định dựa trên các đặc điểm mô học của các khối u. Cụ thể, các thuật toán AI phải đáp ứng các yêu cầu như: Phân loại chính xác các khối ung thư tuyến tiền liệt thành các cấp độ khác nhau, có độ chính xác,

độ nhạy và độ đặc thù cao, có thể áp dụng cho các hình ảnh mô bệnh học tuyến tiền liệt từ các bệnh viện và phòng khám khác nhau. Các thuật toán AI sẽ được đánh giá dựa trên một số tiêu chí, bao gồm độ chính xác, độ nhạy và độ đặc thù. Độ chính xác là tỷ lệ các khối ung thư tuyến tiền liệt được phân loại chính xác. Độ nhạy là tỷ lệ các khối ung thư tuyến tiền liệt được phát hiện chính xác. Độ đặc thù là tỷ lệ các khối u lành tính được phân loại chính xác. Bài toán PANDA Challenge là một bài toán khó, nhưng có tiềm năng mang lại lợi ích to lớn cho lĩnh vực y học. Nếu thành công, các thuật toán AI được phát triển từ thử thách này có thể giúp cải thiện đáng kể độ chính xác và hiệu quả của việc chẩn đoán và phân loại ung thư tuyến tiền liệt.

1.2.2. Các phương pháp giải quyết bài toán

1.2.2.1. Phương pháp học máy

Để giải quyết bài toán PANDA Challenge, một số nghiên cứu trước đó đã sử dụng các kỹ thuật học máy để học cách phân loại các hình ảnh mô bệnh học tuyến tiền liệt thành các cấp độ khác nhau. Các kỹ thuật này có thể được sử dụng để tính toán các đặc trưng của các hình ảnh, chẳng hạn như kích thước, hình dạng và màu sắc của các tế bào ung thư. Sau đó, các đặc trưng này có thể được sử dụng để phân loại các hình ảnh. Bài báo "Prediction of prostate cancer using machine learning algorithms" của tác giả Muktevi Srivenkatesh được công bố trên tạp chí International Journal of Recent Technology and Engineering (IJRTE) vào năm 2020 [3]. Bài báo này trình bày một nghiên cứu sử dụng các thuật toán học máy để dự đoán ung thư tuyến tiền liệt. Nghiên cứu này đã sử dụng tập dữ liệu gồm 62 bệnh nhân ung thư tuyến tiền liệt và 38 bệnh nhân không mắc ung thư tuyến tiền liệt. Các thuật toán học máy được sử dụng trong nghiên cứu này bao gồm: Logistic regression, Support vector machine (SVM), Random forest, K-Nearest Neighbour, Naive Bayes. Các thuật toán học máy này được sử dụng để xây dựng các mô hình dự đoán ung thư tuyến tiền liệt. Các mô hình này được đánh

giá trên tập dữ liệu kiểm tra và đạt được độ chính xác dự đoán khá cao. Kết quả của nghiên cứu này cho thấy rằng các thuật toán học máy có thể được sử dụng để dự đoán ung thư tuyến tiền liệt một cách hiệu quả. Điều này có thể giúp các bác sĩ chẩn đoán sớm ung thư tuyến tiền liệt và đưa ra phương pháp điều trị phù hợp cho bệnh nhân. Hạn chế của nghiên cứu này là tập dữ liệu được sử dụng tương đối nhỏ và chỉ bao gồm bệnh nhân đến từ một bệnh viện. Cần có thêm nghiên cứu với tập dữ liệu lớn hơn và đa dạng hơn để xác nhận kết quả của nghiên cứu này.

1.2.2.2. Phương pháp học sâu

Bên cạnh sự thành công của các phương pháp học máy, các kỹ thuật học sâu cho bài toán này cũng đạt được nhiều thành tựu đáng kể. Bài báo "Deep learning in prostate cancer diagnosis and Gleason grading in histopathology images: An extensive study" của Linkon, Ali Hasan Md, et al [4] được công bố trên tạp chí Informatics in Medicine Unlocked vào năm 2021. Bài báo này trình bày một nghiên cứu tổng quan về việc sử dụng học sâu để chẩn đoán ung thư tuyến tiền liệt và phân loại cấp độ Gleason [5] trên hình ảnh mô bệnh học. Nghiên cứu này đã tổng hợp các nghiên cứu được công bố trước về việc sử dụng học sâu để chẩn đoán ung thư tuyến tiền liệt và phân loại cấp độ Gleason trên hình ảnh mô bệnh học. Các nghiên cứu này sử dụng nhiều mô hình học sâu khác nhau như: CNN, MobileNet, Unet, InceptionV3, GCN,... Các nghiên cứu này cũng đạt được độ chính xác chẩn đoán ung thư tuyến tiền liệt và độ chính xác phân loại cấp độ Gleason khá cao. Kết quả của nghiên cứu này cho thấy rằng học sâu có tiềm năng lớn để cải thiện độ chính xác và hiệu quả của việc chẩn đoán và điều trị ung thư tuyến tiền liệt. Tuy nhiên, nghiên cứu này cũng chỉ ra một số hạn chế của việc sử dụng học sâu để chẩn đoán ung thư tuyến tiền liệt và phân loại cấp độ Gleason. Một trong những hạn chế lớn nhất là các mô hình học sâu thường yêu cầu một tập dữ liệu lớn và đa dạng để đào tạo. Ngoài ra, các mô hình học sâu có thể bị ảnh hưởng bởi nhiễu trong dữ liệu và các vấn đề về độ tổng quát hóa.

1.2.3. Phương pháp đề xuất giải quyết bài toán

Như đã trình bày ở phần trên, trong bài toán này, đề xuất sử dụng các phương pháp học sâu để giải quyết. Các phương pháp học sâu sẽ phù hợp với các mô hình có dữ liệu đầu vào lớn và là các hình ảnh. Một phương pháp học sâu phổ biến thường được sử dụng trong các bài toán xử lý hình ảnh là CNN. Convolutional Neural Network (CNN) [6] là một loại mạng nơ-ron nhân tạo (neural network) được sử dụng rộng rãi trong các bài toán xử lý hình ảnh. CNN có thể học các đặc trưng không gian của hình ảnh, chẳng hạn như hình dạng, kích thước, màu sắc,... và sử dụng các đặc trưng này để phân loại, nhận dạng, hay thực hiện các tác vụ khác trên hình ảnh phù hợp cho bài toán dự đoán mức độ nghiêm trọng của ung thư tuyến tiền liệt trong đề án này. CNN có một số điểm nổi bật sau:

- Khả năng học các đặc trưng không gian của hình ảnh: CNN có thể học các đặc trưng từ các vùng nhỏ của hình ảnh và kết hợp chúng lại để tạo thành các đặc trưng tổng quát hơn. Điều này giúp CNN có thể phát hiện các đặc trưng không gian của hình ảnh, chẳng hạn như hình dạng, kích thước, biến chúng của ung thư tuyến tiền liệt ở các giai đoạn cụ thể.
- Khả năng giảm thiểu số lượng tham số: CNN có thể học các đặc trưng không gian của hình ảnh bằng cách sử dụng các bộ lọc (filter). Các bộ lọc này được chia sẻ giữa các pixel của hình ảnh, giúp giảm thiểu số lượng tham số của CNN.
- Khả năng học các đặc trưng không biến đổi: CNN có thể học các đặc trưng không biến đổi theo vị trí của hình ảnh. Điều này giúp CNN có thể phát hiện các đặc trưng của hình ảnh dù chúng xuất hiện ở bất kỳ vị trí nào trên hình ảnh. .

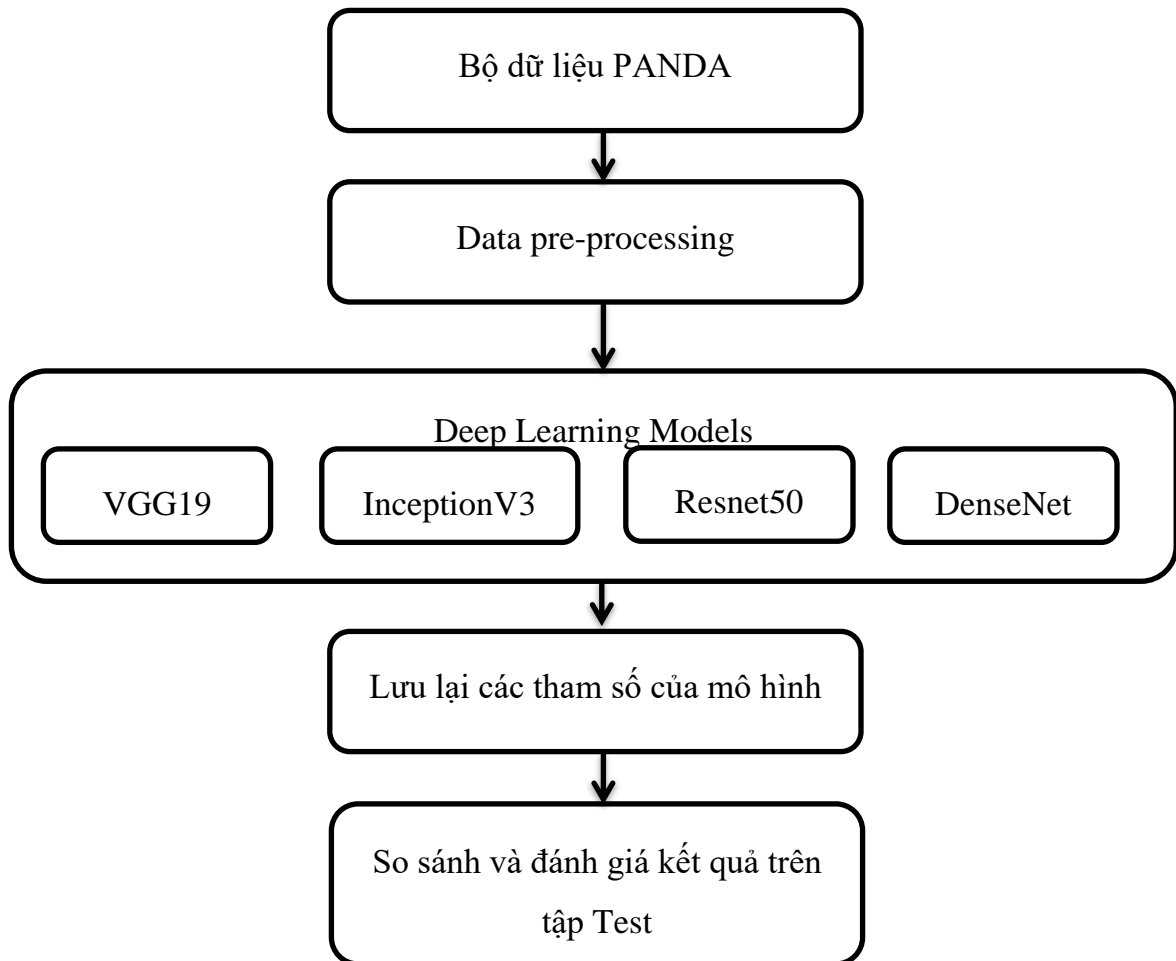
Trong đề án này, chúng tôi đã sử dụng các biến thể của CNN để giải quyết bài toán và tiến hành so sánh kết quả. Một số biến thể của CNN được chúng tôi sử dụng cho bài toán này: VGG19 [7], InceptionV3 [8], ResNet50 [9], Densenet [10]. Kết quả cho thấy,

các biến thể của CNN đều có thể đạt được hiệu quả tốt trong bài toán dự đoán mức độ nghiêm trọng của ung thư tuyến tiền liệt.

1.3. Phương pháp giải quyết bài toán

1.3.1. Mô hình tổng quát

Từ tập dữ liệu PANDA được cung cấp ban đầu chúng tôi tiến hành tiền xử lý dữ liệu như resize ảnh, tăng cường dữ liệu,... Tiếp đến chúng tôi tiến hành quá trình huấn luyện mô hình trên các mô hình học sâu đã được giới thiệu trước đó. Sau đó chúng tôi sẽ lưu lại các tham số của mô hình cuối cùng là tiến hành so sánh và đánh giá giữa các phương pháp đề xuất. Mô hình tổng quát được trình bày ở Hình 3.1



Hình 6. Mô hình tổng quát

1.3.2. Đặc trưng của mô hình đề xuất

1.3.2.1. Deep Learning Models

Sau khi có dữ liệu thích hợp thì thực hiện việc huấn luyện mô hình *đã đề xuất ở trên là VGG19, InceptionV3, Resnet50 và Densenet*.

- **VGG19 [7]** là một mô hình mạng nơ-ron tích chập được phát triển bởi nhóm nghiên cứu của Karen Simonyan và Andrew Zisserman tại Đại học Oxford. Mô hình này được công bố lần đầu tiên trong bài báo "Very Deep Convolutional Networks for Large-Scale Image Recognition" tại Hội nghị Quốc tế về Máy học Thực tế (ICCV) năm 2014. Mô hình gồm 19 lớp tích chập, được chia thành 5 khối. Mỗi khối gồm 2 hoặc 3 lớp tích chập được nối tiếp nhau. Mỗi lớp tích chập đều được theo sau bởi một lớp kích hoạt ReLU và một lớp max pooling. VGG19 đạt được kết quả tốt trong các cuộc thi phân loại hình ảnh, bao gồm ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2014 với độ chính xác 92,9%. Mô hình này được sử dụng trong nhiều ứng dụng phân loại hình ảnh như phân loại đối tượng, dự đoán nhãn, tìm kiếm hình ảnh. VGG19 có ưu điểm là kiến trúc đơn giản, dễ hiểu và dễ triển khai, tuy nhiên nhược điểm của mô hình này là kiến trúc sâu, đòi hỏi nhiều dữ liệu và thời gian để huấn luyện.



Hình 1.3. Kiến trúc VGG19 [7]

- InceptionV3 [8]** là một mô hình mạng nơ-ron tích chập (CNN) được phát triển bởi Google AI. Mô hình này được công bố lần đầu tiên trong bài báo "Inception-v3: Inception Architectures for Large-Scale Image Recognition" tại Hội nghị Quốc tế về Máy học Thực tế (ICCV) năm 2015. Kiến trúc của InceptionV3 bao gồm 22 lớp tích chập, được chia thành 9 khối. Mỗi khối gồm một hoặc nhiều mô hình Inception. Mô hình Inception kết hợp các kích thước kernel khác nhau trong một lớp, cho phép InceptionV3 học các đặc trưng ở nhiều kích thước khác nhau, giúp cải thiện hiệu năng phân loại hình ảnh. InceptionV3 đạt được độ chính xác 93,1% trong cuộc thi ImageNet Large Scale Visual Recognition Challenge (ILSVRC) năm 2015. Mô hình này có thể được sử dụng trong nhiều ứng dụng phân loại hình ảnh, chẳng hạn như phân loại đối tượng, dự đoán nhãn và tìm kiếm hình ảnh. Ưu điểm của InceptionV3 là kiến trúc hiệu quả, hiệu năng tốt và tính phổ biến tuy nhiên mô hình này cũng có một số nhược điểm là kiến trúc phức tạp và kích thước lớn, đòi hỏi nhiều tài nguyên lưu trữ và sử dụng.

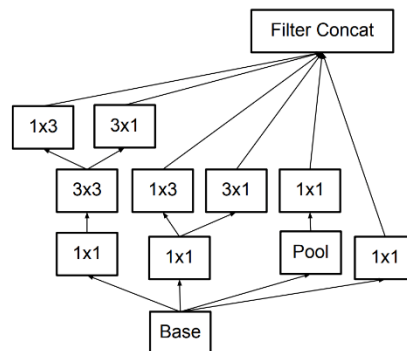
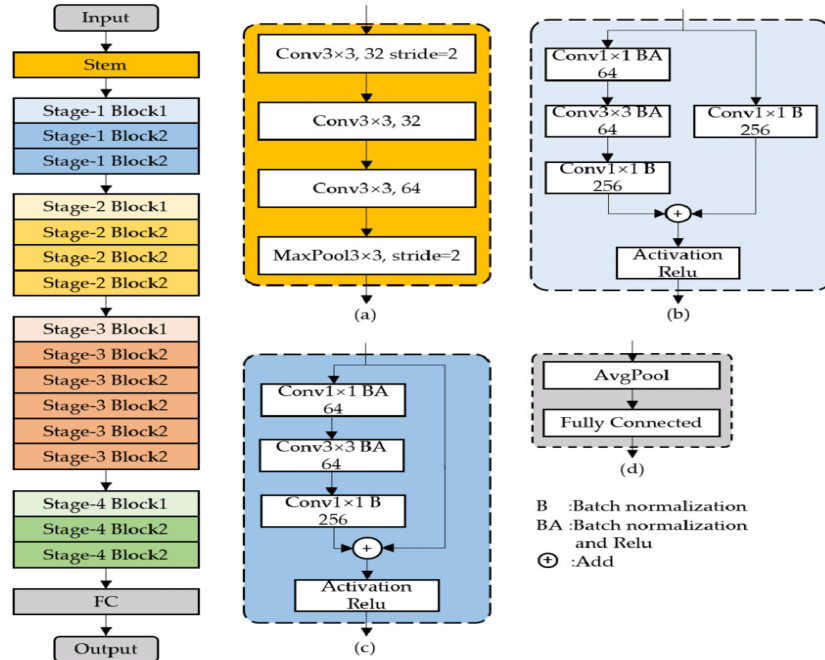


Figure 7. Inception modules with expanded the filter bank outputs. This architecture is used on the coarsest (8×8) grids to promote high dimensional representations, as suggested by principle 2 of Section 2. We are using this solution only on the coarsest grid, since that is the place where producing high dimensional sparse representation is the most critical as the ratio of local processing (by 1×1 convolutions) is increased compared to the spatial aggregation.

Hình 3. Mô hình InceptionV3 [8]

- **ResNet50 [9]** là một mô hình mạng nơ-ron tích chập (CNN) được phát triển bởi Kaiming He và cộng sự tại Viện Công nghệ California, Berkeley. Mô hình này được công bố lần đầu tiên trong bài báo "Deep Residual Learning for Image Recognition" tại Hội nghị Quốc tế về Máy học Thực tế (ICCV) năm 2015. Kiến trúc của ResNet50 bao gồm 53 lớp tích chập, được chia thành 4 khối. Mỗi khối gồm 3 hoặc 4 mô hình ResNet. Các mô hình ResNet sử dụng kết nối thẳng và kết nối nhảy cóc để giảm hiện tượng suy giảm gradient và cải thiện hiệu năng của mô hình. ResNet50 cũng đạt được kết quả tốt trong các cuộc thi phân loại hình ảnh, bao gồm cuộc thi ImageNet Large Scale Visual Recognition Challenge (ILSVRC) năm 2015. Mô hình này đạt được độ chính xác 92,7% trong cuộc thi này. ResNet50 được sử dụng trong nhiều ứng dụng phân loại hình ảnh, chẳng hạn như: phân loại đối tượng, dự đoán nhãn, tìm kiếm hình ảnh. Một số ưu điểm của ResNet50 như có thể xử lý các bài toán phức tạp, tận dụng được tính linh hoạt của các khối residual để kết nối các lớp khác nhau. Tuy nhiên cũng có một số nhược điểm như cần nhiều tài nguyên tính toán và bộ nhớ để huấn luyện mô hình, còn gặp phải vấn đề vanishing gradient [11] khi số lượng layer quá cao hoặc khi kích thước ảnh quá nhỏ.



Hình 4. Mô hình ResNet50 [9]

- DenseNet [10]** là một mô hình mạng nơ-ron tích chập (CNN) được phát triển bởi Gao Huang và cộng sự tại Đại học Stanford. Mô hình này được công bố lần đầu tiên trong bài báo "Densely Connected Convolutional Networks" tại Hội nghị Quốc tế về Máy học Thực tế (ICCV) năm 2016. Kiến trúc của DenseNet: DenseNet có kiến trúc gồm các khối tích chập được kết nối với nhau theo cách "dense". Điều này có nghĩa là đầu ra của mỗi khối tích chập được kết nối với tất cả các khối tích chập tiếp theo. DenseNet đạt được kết quả tốt trong các cuộc thi phân loại hình ảnh, Mô hình này đạt được độ chính xác 93,5% trong cuộc thi ImageNet Large Scale Visual Recognition Challenge (ILSVRC) năm 2016. DenseNet có thể được sử dụng trong nhiều ứng dụng phân loại hình ảnh, chẳng hạn như: phân loại đối tượng, dự đoán nhãn, tìm kiếm hình ảnh. Ưu điểm của mô hình này chính là có thể giảm thiểu số lượng tham số của mạng, do đó tiết kiệm thời gian huấn luyện và tăng hiệu suất, có khả năng học được các đặc trưng phức tạp từ dữ liệu, do đó có thể áp dụng cho nhiều bài toán khác nhau.

Mô hình cũng có một số nhược điểm đó là cần nhiều tài nguyên tính toán và bộ nhớ để huấn luyện mô hình

Layers	Output Size	DenseNet-121	DenseNet-169	DenseNet-201	DenseNet-264
Convolution	112×112	7×7 conv, stride 2			
Pooling	56×56	3×3 max pool, stride 2			
Dense Block (1)	56×56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition Layer (1)	56×56	1×1 conv			
	28×28	2×2 average pool, stride 2			
Dense Block (2)	28×28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition Layer (2)	28×28	1×1 conv			
	14×14	2×2 average pool, stride 2			
Dense Block (3)	14×14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 64$
Transition Layer (3)	14×14	1×1 conv			
	7×7	2×2 average pool, stride 2			
Dense Block (4)	7×7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$
Classification Layer	1×1	7×7 global average pool			
		1000D fully-connected, softmax			

Hình 5. Mô hình DenseNet [10]

1.3.2.2. Lưu lại các tham số mô hình

Lưu lại các tham số mô hình có hai mục đích chính: Đánh giá và kiểm tra: Các tham số mô hình được lưu lại để có thể sử dụng để đánh giá và kiểm tra hiệu năng của mô hình trên tập dữ liệu khác. Tái sử dụng mô hình: Các tham số mô hình được lưu lại để có thể tái sử dụng mô hình trên các tập dữ liệu khác.

1.3.2.3. So sánh và đánh giá kết quả trên tập Test

Sau khi lưu lại các tham số của mô hình, tiến hành so sánh và đánh giá kết quả của mô hình trên tập dữ liệu Test. Tập dữ liệu Test là tập dữ liệu không được sử dụng để huấn luyện mô hình, do đó kết quả của mô hình trên tập dữ liệu này phản ánh chính xác hiệu năng của mô hình. Sử dụng độ đo Accuracy [12] và Kappa [13] để đánh giá

mô hình. Chúng tôi sẽ trình bày chi tiết hơn ở phần sau về các bảng so sánh và đánh giá hiệu suất của từng mô hình.

1.4. Thực nghiệm

1.4.1. Dữ liệu

Bộ dữ liệu PANDA (Prostate cANcer graDe Assessment) [2] bao gồm 11.000 hình ảnh. Tập dữ liệu được phát triển bởi Computational Pathology Group (CPG) của Trường Đại học Radboud Medical Center và Karolinska Institute's Department of Medical Epidemiology and Biostatistics (MEB). Hiện nay, đây là bộ dữ liệu công cộng có sẵn lớn nhất, khoảng 8 lần kích thước so với CAMELYON17 Challenge [14]. Quá trình đánh giá bao gồm việc tìm và phân loại mô u ác thành các mẫu Gleason dựa trên các mẫu tăng trưởng kiến trúc của khối u. Do đó, nó được chuyển đổi thành một ISUP grade trên một thang điểm từ 1 đến 5, dựa trên sự xuất hiện của hai mẫu Gleason rõ rệt. Trung bình, mỗi hình ảnh cá nhân có kích thước khoảng $25.000\text{px} \times 25.000\text{px}$ với các kênh màu RGB. Tập huấn luyện chứa một số nhiều nhãn. Nhiều nhãn này được giới thiệu vì một số lý do, bao gồm các báo cáo của bác sĩ dự phòng không rõ ràng, lỗi chú thích, lỗi trong chẩn đoán ban đầu và sự không đồng tình giữa các bác sĩ dự phòng. Đối với dữ liệu của Radboudumc và dữ liệu của Karolinska, từng tập kiểm tra đã được đánh giá độc lập bởi ba bác sĩ dự phòng chuyên ngành bệnh lý niệu đạo.

Bộ dữ liệu bao gồm file chứa hình ảnh [train/test]_images là thư mục chứa các hình ảnh về ung thư tuyến tiền liệt và train_label_masks chứa các mặt nạ phân đoạn hiển thị phần nào của hình ảnh dẫn đến ISUP grade. Ngoài các file chứa hình ảnh và mặt nạ thì còn 2 file train.csv và test.csv. Chi tiết các features trong file csv như sau:

- `image_id`: Mã ID cho hình ảnh.
- `data_provider`: Tên của tổ chức cung cấp dữ liệu. Cả Viện Karolinska và Trung tâm Y tế Đại học Radboud đều đóng góp dữ liệu. Họ sử dụng các máy quét khác nhau với độ phân giải kính hiển vi tối đa hơi khác nhau và hợp tác với các nhà bệnh lý khác nhau để dán nhãn hình ảnh của họ.
- `isup_grade`: Chỉ dành cho tập train. Biến mục tiêu. Mức độ nghiêm trọng của ung thư trên thang điểm từ 0 đến 5.
- `gleason_score`: Chỉ dành cho tập train. Một hệ thống đánh giá mức độ nghiêm trọng của ung thư thay thế với nhiều cấp hơn so với thang điểm ISUP. Để biết chi tiết về cách so sánh các hệ thống Gleason và ISUP.

1.4.2. Xử lý dữ liệu

Dữ liệu PANDA là một tập dữ liệu hình ảnh về ung thư tuyến tiền liệt. Dữ liệu được cung cấp từ cuộc thi Prostate cANcer graDe Assessment (PANDA) Challenge, nhằm mục đích phát triển các mô hình máy học để phân loại ung thư tuyến tiền liệt. Dữ liệu PANDA bao gồm các hình ảnh ung thư tuyến tiền liệt, được chụp bằng máy cộng hưởng từ hạt nhân (MRI). Các hình ảnh được phân loại thành 4 cấp độ. Dữ liệu đã được xử lý và chia ra thành tập train và test. Tập train, chiếm 90% dữ liệu. Tập test, chiếm 10% dữ liệu. Các bước tiền xử lý dữ liệu PANDA được thực hiện nhằm mục đích chuẩn hóa dữ liệu và cải thiện hiệu quả của mô hình phân loại. Các bước tiền xử lý bao gồm:

- **Normalizing the images:** Chuẩn hóa dữ liệu hình ảnh bằng cách chuyển đổi giá trị pixel thành giá trị từ 0 đến 1.
- **Chuyển đổi kích thước của hình ảnh:** Chuyển đổi kích thước của hình ảnh cho phù hợp với từng mô hình. Kích thước của hình ảnh ảnh hưởng đến hiệu quả của mô hình phân loại. Các mô hình phân loại thường hoạt động tốt hơn với các

hình ảnh có kích thước nhỏ. Trong đồ án này, chúng tôi đã chuyển đổi kích thước của hình ảnh thành 224 x 224 pixel, phù hợp với các mô hình phân loại phổ biến.

- **Tăng cường ảnh:** Thực hiện các bước tăng cường ảnh như xoay, lật để làm cho bộ dữ liệu đa dạng hơn và cho ra kết quả tốt hơn khi huấn luyện mô hình. Tăng cường ảnh là một kỹ thuật giúp làm cho bộ dữ liệu đa dạng hơn. Điều này sẽ giúp cho mô hình phân loại có thể học được các đặc trưng của đối tượng trong nhiều điều kiện khác nhau.

Các bước tiền xử lý dữ liệu PANDA là một bước quan trọng trong quá trình phát triển các mô hình phân loại ung thư tuyến tiền liệt. Các bước tiền xử lý giúp chuẩn hóa dữ liệu và cải thiện hiệu quả của mô hình phân loại đã được đề xuất ở trên.

1.4.3. Công nghệ sử dụng

Ngôn ngữ	Python [15]
Thư viện	Keras [16], TensorFlow [17]
Môi trường	Kaggle GPU P100

Bảng 1. Mô tả công nghệ sử dụng

1.4.4. Cách đánh giá

1.4.4.1. Quadratic Weighted Kappa

Quadratic Weighted Kappa (Kappa score) [13] được dùng để đo lường mức độ nhất quán giữa hai kết quả. Chỉ số này thường dao động từ 0 (nhất quán ngẫu nhiên) đến 1 (nhất quán hoàn toàn). Trong trường hợp mức độ nhất quán thấp hơn dự kiến một cách ngẫu nhiên, chỉ số này có thể giảm xuống dưới 0.

Kappa score được tính như sau. Đầu tiên, xây dựng một ma trận biểu đồ O có kích thước NxN , sao cho $O_{i,j}$ tương ứng với số lượng *isup_grades* i (thực tế) nhận được giá trị dự đoán j. Một ma trận trọng số w kích thước N x N, được tính toán dựa trên sự khác biệt giữa các giá trị thực tế và dự đoán:

$$w_{i,j} = \frac{(i - j)^2}{(N - 1)^2} \quad (1)$$

Một ma trận biểu đồ E có kích thước N x N của các kết quả dự kiến, được tính toán giả định rằng không có mối tương quan giữa các giá trị. Điều này được tính toán dưới dạng tích ngoài giữa vector biểu đồ thực tế của kết quả và vector biểu đồ dự đoán, được chuẩn hóa sao cho E và O có cùng tổng.

Từ ba ma trận này, Kappa score được định nghĩa như sau:

$$\kappa = 1 - \frac{\sum_{i,j} w_{i,j} O_{i,j}}{\sum_{i,j} w_{i,j} E_{i,j}} \quad (2)$$

1.4.4.2. Binary Cross Entropy

Binary Cross-Entropy (BCE) [18] là một hàm mất mát phổ biến trong các bài toán phân loại nhị phân. Hàm BCE được sử dụng để đo lường sai khác giữa các dự đoán nhị phân và nhãn thực tế trong bài toán phân loại nhị phân. Thông thường, đầu ra của một mô hình phân loại nhị phân được ánh xạ vào khoảng giá trị [0, 1]. BCE tính toán sai số chéo (cross-entropy) giữa dự đoán nhị phân và nhãn thực tế dưới dạng một giá trị số.

$$BCE = -\frac{1}{N} \sum_{i=0}^N y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i)$$

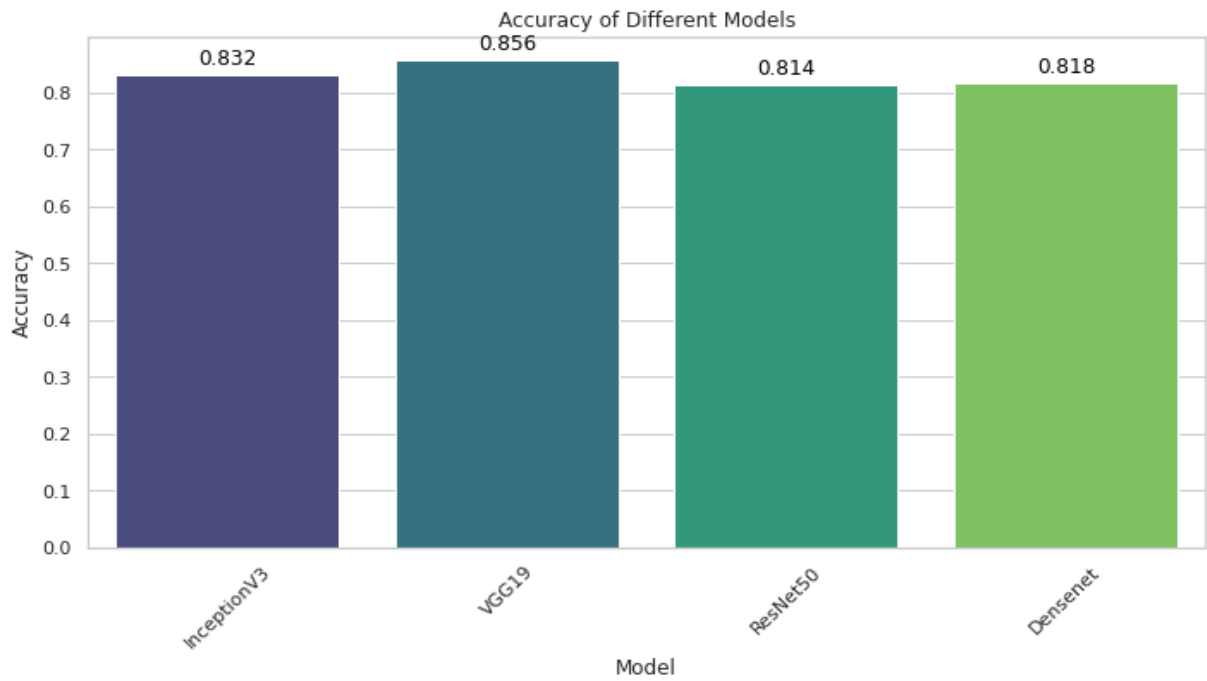
1.5. Kết quả đạt được

1.5.1. Hyperparameter

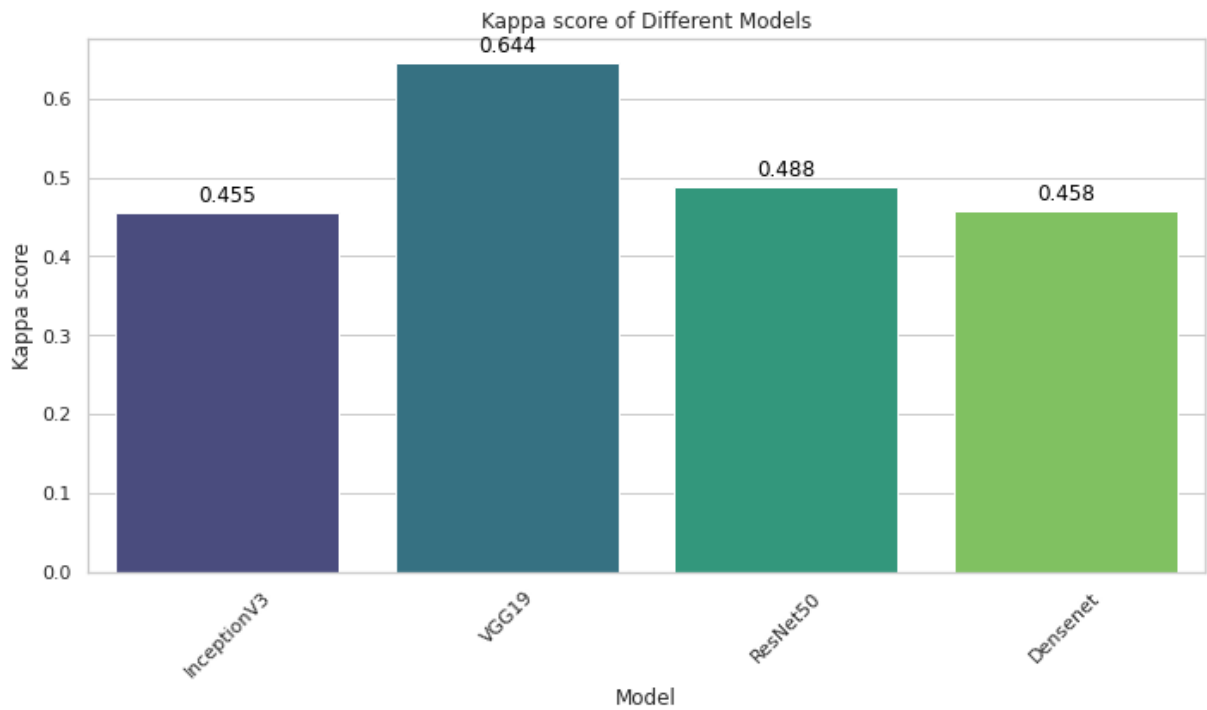
Đồ án này sử dụng bộ tối ưu hóa SGD Momentum [19] với learning rate 0.0005, momentum = 0.9. Đồng thời áp dụng chuẩn hóa batch normalization [20] và đặt batch size là 16. Chúng tôi khởi tạo mô hình đề xuất ở phần trên là các mô hình có sẵn trong thư viện keras cho tác vụ phân loại hình ảnh. Hàm này nhận tham số num_classes là số lớp đầu ra của mô hình. Chúng tôi sử dụng các trọng số đã được huấn luyện trên tập dữ liệu ImageNet [21] cho các mô hình. Tiếp theo, thêm một lớp Flatten() và một lớp Dense với 32 neuron và hàm kích hoạt ReLU, và một lớp Dropout với tỉ lệ 20% để tránh hiện tượng quá khớp (overfitting). Cuối cùng, thêm một lớp Dense với số neuron bằng num_classes và hàm kích hoạt Softmax để phân loại hình ảnh. Chúng tôi huấn luyện các mô hình với 20 epoch, sử dụng các callback để lưu lại mô hình với các điểm checkpoint tốt nhất.

1.5.2. Kết quả

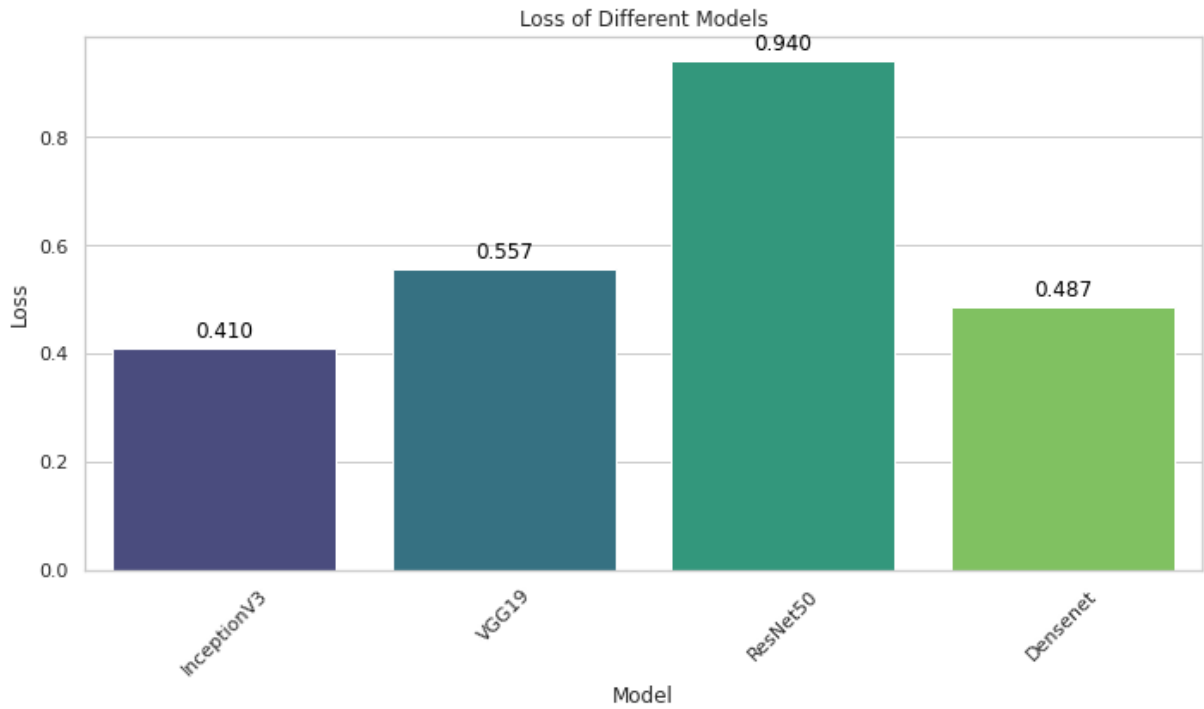
Dựa trên các phương pháp, tham số đã được nêu ra ở trên, chúng tôi tiến hành thực nghiệm để dự đoán mức độ nghiêm trọng của ung thư tuyến tiền liệt.



Hình 7. Accuracy giữa các mô hình






Hình 8. Kappa score giữa các mô hình



Hình 9. Loss giữa các mô hình

Thông qua các biểu đồ trực quan hóa ở trên, ta thấy mô hình VGG19 mang lại kết quả tốt nhất. Do đó, chúng tôi tiến hành dự đoán một số mẫu dựa trên mô hình này.

Image	Predict ISUP grade
	3

	1
	5

Bảng 2. Kết quả thực nghiệm

Chúng tôi cũng tiến hành so sánh phương pháp của chúng tôi với một số phương pháp học máy như: K-Nearest Neighbour, Support Vector Machines, Naive Bayes. [3]

Đối với Accuracy của 3 mô hình trên chúng tôi tính toán dựa trên số liệu từ Confusion matrix trong bài báo. Đối với Kappa score chúng tôi sử dụng các kết quả có sẵn.

	Accuracy	Kappa score
K-Nearest Neighbour	0.8	0.61
Support Vector Machines	0.7	0.44
Naive Bayes	0.8	0.61
VGG19 (Ours)	0.856	0.644

Bảng 3. Kết quả so sánh

Từ các kết quả so sánh trên có thể thấy mô hình của chúng tôi cho ra các kết quả tốt hơn so với các mô hình còn lại. Bởi vì, mô hình VGG19 là một mô hình Deep Learning được sử dụng để nhận dạng hình ảnh. Mô hình này có thể học hỏi các đặc trưng phức tạp từ dữ liệu hình ảnh, chẳng hạn như các đặc trưng về màu sắc, hình dạng và kết cấu. Điều này cho phép mô hình VGG19 đạt được độ chính xác cao hơn trong các nhiệm vụ nhận dạng và phân loại hình ảnh.

1.6. Kết luận

1.6.1. Kết quả đạt được

Về mặt lý thuyết, Đồ án đã tìm hiểu về các phương pháp giải quyết bài toán dự đoán mức độ nghiêm trọng của ung thư tuyến tiền liệt, đồng thời Đồ án cũng đề xuất phương pháp học sâu để giải quyết bài toán này.

Kết quả thực nghiệm của nghiên cứu đã chứng minh tính hiệu quả và tiềm năng của phương pháp đề xuất trong bài toán dự đoán mức độ nghiêm trọng của ung thư tuyến tiền liệt. Thực nghiệm trên các tập dữ liệu đa dạng đã chứng minh tính chính xác trong phương pháp của chúng tôi và các phương pháp khác.

1.6.2. Hạn chế

Mặc dù mô hình đã đạt được các kết quả thực nghiệm khá tốt, tuy nhiên vẫn còn một số hạn chế bởi dữ liệu bị nhiễu hay các đặc điểm bệnh lý khác trong tuyến tiền liệt, chẳng hạn như viêm tuyến tiền liệt. Điều này có thể dẫn đến sai sót trong quá trình huấn luyện và kiểm tra của mô hình.

1.6.3. Hướng phát triển

Để cải thiện và phát triển phương pháp trong tương lai, chúng tôi có đề xuất một số hướng tiếp cận để có thể tiếp tục nghiên cứu chuyên sâu hơn. Đầu tiên, mở rộng kiến trúc mô hình để xử lý các trường hợp đặc biệt và khó khăn hơn, bao gồm chất lượng hình ảnh thấp, các trường hợp đặc biệt trong bệnh lý tuyến tiền liệt. Thứ hai, nghiên cứu các phương pháp để kết hợp các mô hình học sâu khác nhau để cải thiện độ chính xác của kết quả đánh giá.

Với sự phát triển tiếp tục và nghiên cứu sâu hơn, phương pháp đề xuất có thể trở thành một công cụ quan trọng và mạnh mẽ trong việc dự đoán mức độ nghiêm trọng của ung thư tuyến tiền liệt trong lĩnh vực y học.

LÀM VIỆC NHÓM

Trình bày tóm tắt cách thức làm việc nhóm

Phân chia công việc của các thành viên trong nhóm

Tổng số lần gặp nhau (tính theo buổi)

Tổng thời gian gặp nhau (tính theo giờ)

TÀI LIỆU THAM KHẢO

1. Egevad, L., Delahunt, B., Srigley, J. R., & Samaratunga, H. (2016). International Society of Urological Pathology (ISUP) grading of prostate cancer—An ISUP consensus on contemporary grading. *Apmis*, 124(6), 433-435.
2. Datasets PANDA (Prostate cANcer graDe Assessment): [Prostate cANcer graDe Assessment \(PANDA\) Challenge | Kaggle](#)
3. Srivenkatesh, M. (2020). Prediction of prostate cancer using machine learning algorithms. *Int. J. Recent Technol. Eng*, 8(5), 5353-5362.
4. Linkon, A. H. M., Labib, M. M., Hasan, T., & Hossain, M. (2021). Deep learning in prostate cancer diagnosis and Gleason grading in histopathology images: An extensive study. *Informatics in Medicine Unlocked*, 24, 100582.
5. Silva-Rodríguez, J., Colomer, A., Sales, M. A., Molina, R., & Naranjo, V. (2020). Going deeper through the Gleason scoring scale: An automatic end-to-end system for histology prostate grading and cribriform pattern detection. *Computer methods and programs in biomedicine*, 195, 105637.
6. O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*.
7. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
8. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826).

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. Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).
11. Hochreiter, S. (1998). The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 6(02), 107-116.
12. Šimundić, A. M. (2009). Measures of diagnostic accuracy: basic definitions. *ejifcc*, 19(4), 203.
13. Vaughn, D., & Justice, D. (2015). On the direct maximization of quadratic weighted kappa. *arXiv preprint arXiv:1509.07107*.
14. CAMELYON17 Challenge: [Home - Grand Challenge \(grand-challenge.org\)](http://grand-challenge.org)
15. Python: <https://www.python.org>
16. Keras: <https://keras.io>
17. Tensorflow: <https://www.tensorflow.org>
18. Ruby, U., & Yendapalli, V. (2020). Binary cross entropy with deep learning technique for image classification. *Int. J. Adv. Trends Comput. Sci. Eng*, 9(10).
19. Shi, B. (2021). On the hyperparameters in stochastic gradient descent with momentum. *arXiv preprint arXiv:2108.03947*.
20. Bjorck, N., Gomes, C. P., Selman, B., & Weinberger, K. Q. (2018). Understanding batch normalization. *Advances in neural information processing systems*, 31.

21. Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition* (pp. 248-255). Ieee.

PHỤ LỤC

Phần này bao gồm những nội dung cần thiết nhằm minh họa hoặc hỗ trợ cho nội dung đề án như số liệu, biểu mẫu, tranh ảnh. . . . nếu sử dụng những câu trả lời cho một *bảng câu hỏi thì bảng câu hỏi mẫu này phải được đưa vào phần Phụ lục ở dạng nguyên bản* đã dùng để điều tra, thăm dò ý kiến; **không được tóm tắt hoặc sửa đổi**. Các tính toán mẫu trình bày tóm tắt trong các biểu mẫu cũng cần nêu trong Phụ lục của luận văn. Phụ lục không được dày hơn phần chính của đề án

TỰ ĐÁNH GIÁ

STT	Nội dung	Điểm chuẩn	Tự chấm	Ghi chú
1 (8.5)	1.1 Giới thiệu về bài toán	0.5		
	1.2 Phân tích yêu cầu của bài toán	1		
	1.3 Phương pháp giải quyết bài toán	1.5		
	1.4 Thực nghiệm	4		
	1.5 Kết quả đạt được	1		
	1.6 Kết luận	0.5		
2 (1)	Báo cáo (chú ý các chú ý 2,3,4,6 ở trang trước, nếu sai sẽ bị trừ điểm nặng)	1đ		
3 (0.5)	Điểm nhóm (chú ý trả lời các câu hỏi trong mục làm việc nhóm)	0.5đ		
Tổng điểm				