

TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI  
VIỆN ĐIỆN TỬ - VIỄN THÔNG



ĐỒ ÁN  
**TỐT NGHIỆP ĐẠI HỌC**

**Đề tài:**

**PHÂN LOẠI ẢNH UNG THƯ DẠ DÀY  
DỰA VÀO CÁC ĐẶC TRƯNG**

Sinh viên thực hiện: **TRƯƠNG THANH HIỀN**

Lớp ĐT5 – K59

**PHÓ THỊ HÀ ANH**

Lớp ĐT6 – K59

Giảng viên hướng dẫn: **TS. NGUYỄN VIỆT DŨNG**

Hà Nội, 6-2019



TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI  
VIỆN ĐIỆN TỬ - VIỄN THÔNG



ĐỒ ÁN  
**TỐT NGHIỆP ĐẠI HỌC**

**Đề tài:**

**PHÂN LOẠI ẢNH UNG THƯ DẠ DÀY  
DỰA VÀO CÁC ĐẶC TRƯNG**

Sinh viên thực hiện: **TRƯƠNG THANH HIỀN**

Lớp DT5 – K59

**PHÓ THỊ HÀ ANH**

Lớp DT6 – K59

Giảng viên hướng dẫn: **TS. NGUYỄN VIỆT DŨNG**

Cán bộ phản biện:

Hà Nội, 6-2019



# ĐÁNH GIÁ QUYỀN ĐỒ ÁN TỐT NGHIỆP

(Dùng cho giảng viên hướng dẫn)

Tên giảng viên đánh giá: TS. Nguyễn Việt Dũng

Họ và tên Sinh viên: Trương Thanh Hiền. MSSV: 20141612

Tên đồ án: Phân Loại Ảnh Ung Thư Dạng Dày Dựa Vào Các Đặc Trưng

**Chọn các mức điểm phù hợp cho sinh viên trình bày theo các tiêu chí dưới đây:**

Rất kém (1); Kém (2); Đạt (3); Giỏi (4); Xuất sắc (5)

Có sự kết hợp giữa lý thuyết và thực hành (20)					
1	Nêu rõ tính cấp thiết và quan trọng của đề tài, các vấn đề và các giả thuyết (bao gồm mục đích và tính phù hợp) cũng như phạm vi ứng dụng của đồ án	1	2	3	4 5
2	Cập nhật kết quả nghiên cứu gần đây nhất (trong nước/quốc tế)	1	2	3	4 5
3	Nêu rõ và chi tiết phương pháp nghiên cứu/giải quyết vấn đề	1	2	3	4 5
4	Có kết quả mô phỏng/thực nghiệm và trình bày rõ ràng kết quả đạt được	1	2	3	4 5
Có khả năng phân tích và đánh giá kết quả (15)					
5	Kế hoạch làm việc rõ ràng bao gồm mục tiêu và phương pháp thực hiện dựa trên kết quả nghiên cứu lý thuyết một cách có hệ thống	1	2	3	4 5
6	Kết quả được trình bày một cách logic và dễ hiểu, tất cả kết quả đều được phân tích và đánh giá thỏa đáng.	1	2	3	4 5
7	Trong phần kết luận, tác giả chỉ rõ sự khác biệt (nếu có) giữa kết quả đạt được và mục tiêu ban đầu đề ra đồng thời cung cấp lập luận để đề xuất hướng giải quyết có thể thực hiện trong tương lai.	1	2	3	4 5
Kỹ năng viết quyền đồ án (10)					
8	Đồ án trình bày đúng mẫu quy định với cấu trúc các chương logic và đẹp mắt (bảng biểu, hình ảnh rõ ràng, có tiêu đề, được đánh số thứ tự và được giải thích hay đề cập đến trong đồ án, có căn lề, dấu cách sau dấu chấm, dấu phẩy v.v), có mở đầu chương và kết luận chương, có liệt kê tài liệu tham khảo và có trích dẫn đúng quy định	1	2	3	4 5
9	Kỹ năng viết xuất sắc (cấu trúc câu chuẩn, văn phong khoa học, lập luận logic và có cơ sở, từ vựng sử dụng phù hợp v.v.)	1	2	3	4 5
Thành tựu nghiên cứu khoa học (5) (chọn 1 trong 3 trường hợp)					
10a	Có bài báo khoa học được đăng hoặc chấp nhận đăng/đạt giải SVNC khoa học giải 3 cấp Viện trở lên/các giải thưởng khoa học (quốc tế/trong nước) từ giải 3 trở lên/ Có đăng ký bằng phát minh sáng chế	5			
10b	Được báo cáo tại hội đồng cấp Viện trong hội nghị sinh viên nghiên cứu khoa học nhưng không đạt giải từ giải 3 trở lên/Đạt giải khuyến khích trong các kỳ thi quốc gia và quốc tế khác về chuyên ngành như TI contest.	2			
10c	Không có thành tích về nghiên cứu khoa học	0			
Điểm tổng		/50			
Điểm tổng quy đổi về thang 10					

*Nhận xét khác (về thái độ và tinh thần làm việc của sinh viên)*

.....

.....

.....

.....

.....

.....

Ngày: ... / ... / 20...

**Người nhận xét**  
(Ký và ghi rõ họ tên)

# ĐÁNH GIÁ QUYỀN ĐỒ ÁN TỐT NGHIỆP

(Dùng cho giảng viên hướng dẫn)

Tên giảng viên đánh giá: TS. Nguyễn Việt Dũng

Họ và tên Sinh viên: Phó Thị Hà Anh. MSSV: 20140217

Tên đồ án: Phân Loại Ảnh Ung Thư Dạng Dày Dựa Vào Các Đặc Trưng

**Chọn các mức điểm phù hợp cho sinh viên trình bày theo các tiêu chí dưới đây:**

Rất kém (1); Kém (2); Đạt (3); Giỏi (4); Xuất sắc (5)

Có sự kết hợp giữa lý thuyết và thực hành (20)					
1	Nêu rõ tính cấp thiết và quan trọng của đề tài, các vấn đề và các giả thuyết (bao gồm mục đích và tính phù hợp) cũng như phạm vi ứng dụng của đồ án	1	2	3	4 5
2	Cập nhật kết quả nghiên cứu gần đây nhất (trong nước/quốc tế)	1	2	3	4 5
3	Nêu rõ và chi tiết phương pháp nghiên cứu/giải quyết vấn đề	1	2	3	4 5
4	Có kết quả mô phỏng/thực nghiệm và trình bày rõ ràng kết quả đạt được	1	2	3	4 5
Có khả năng phân tích và đánh giá kết quả (15)					
5	Kế hoạch làm việc rõ ràng bao gồm mục tiêu và phương pháp thực hiện dựa trên kết quả nghiên cứu lý thuyết một cách có hệ thống	1	2	3	4 5
6	Kết quả được trình bày một cách logic và dễ hiểu, tất cả kết quả đều được phân tích và đánh giá thỏa đáng.	1	2	3	4 5
7	Trong phần kết luận, tác giả chỉ rõ sự khác biệt (nếu có) giữa kết quả đạt được và mục tiêu ban đầu đề ra đồng thời cung cấp lập luận để đề xuất hướng giải quyết có thể thực hiện trong tương lai.	1	2	3	4 5
Kỹ năng viết quyền đồ án (10)					
8	Đồ án trình bày đúng mẫu quy định với cấu trúc các chương logic và đẹp mắt (bảng biểu, hình ảnh rõ ràng, có tiêu đề, được đánh số thứ tự và được giải thích hay đề cập đến trong đồ án, có căn lề, dấu cách sau dấu chấm, dấu phẩy v.v), có mở đầu chương và kết luận chương, có liệt kê tài liệu tham khảo và có trích dẫn đúng quy định	1	2	3	4 5
9	Kỹ năng viết xuất sắc (cấu trúc câu chuẩn, văn phong khoa học, lập luận logic và có cơ sở, từ vựng sử dụng phù hợp v.v.)	1	2	3	4 5
Thành tựu nghiên cứu khoa học (5) (chọn 1 trong 3 trường hợp)					
10a	Có bài báo khoa học được đăng hoặc chấp nhận đăng/đạt giải SVNC khoa học giải 3 cấp Viện trở lên/các giải thưởng khoa học (quốc tế/trong nước) từ giải 3 trở lên/ Có đăng ký bằng phát minh sáng chế	5			
10b	Được báo cáo tại hội đồng cấp Viện trong hội nghị sinh viên nghiên cứu khoa học nhưng không đạt giải từ giải 3 trở lên/Đạt giải khuyến khích trong các kỳ thi quốc gia và quốc tế khác về chuyên ngành như TI contest.	2			
10c	Không có thành tích về nghiên cứu khoa học	0			
Điểm tổng		/50			
Điểm tổng quy đổi về thang 10					

*Nhận xét khác (về thái độ và tinh thần làm việc của sinh viên)*

.....

.....

.....

.....

.....

.....

Ngày: ... / ... / 20...

**Người nhận xét**  
(Ký và ghi rõ họ tên)



# ĐÁNH GIÁ QUYỀN ĐỒ ÁN TỐT NGHIỆP

(Dùng cho cán bộ phản biện)

Giảng viên đánh giá:.....

Họ và tên sinh viên: Trương Thanh Hiền. MSSV: 20141612

Phó Thị Hà Anh. MSSV: 20140217

Tên đồ án: Phân Loại Ảnh Ung Thư Dạng Dày Dựa Vào Các Đặc Trưng

**Chọn các mức điểm phù hợp cho sinh viên trình bày theo các tiêu chí dưới đây:**

Rất kém (1); Kém (2); Đạt (3); Giỏi (4); Xuất sắc (5)

Có sự kết hợp giữa lý thuyết và thực hành (20)					
1	Nêu rõ tính cấp thiết và quan trọng của đề tài, các vấn đề và các giả thuyết (bao gồm mục đích và tính phù hợp) cũng như phạm vi ứng dụng của đồ án	1	2	3	4 5
2	Cập nhật kết quả nghiên cứu gần đây nhất (trong nước/quốc tế)	1	2	3	4 5
3	Nêu rõ và chi tiết phương pháp nghiên cứu/giải quyết vấn đề	1	2	3	4 5
4	Có kết quả mô phỏng/thực nghiệm và trình bày rõ ràng kết quả đạt được	1	2	3	4 5
Có khả năng phân tích và đánh giá kết quả (15)					
5	Kế hoạch làm việc rõ ràng bao gồm mục tiêu và phương pháp thực hiện dựa trên kết quả nghiên cứu lý thuyết một cách có hệ thống	1	2	3	4 5
6	Kết quả được trình bày một cách logic và dễ hiểu, tất cả kết quả đều được phân tích và đánh giá thỏa đáng.	1	2	3	4 5
7	Trong phần kết luận, tác giả chỉ rõ sự khác biệt (nếu có) giữa kết quả đạt được và mục tiêu ban đầu đề ra đồng thời cung cấp lập luận để đề xuất hướng giải quyết có thể thực hiện trong tương lai.	1	2	3	4 5
Kỹ năng viết quyền đồ án (10)					
8	Đồ án trình bày đúng mẫu quy định với cấu trúc các chương logic và đẹp mắt (bảng biểu, hình ảnh rõ ràng, có tiêu đề, được đánh số thứ tự và được giải thích hay đề cập đến trong đồ án, có căn lề, dấu cách sau dấu chấm, dấu phẩy v.v), có mở đầu chương và kết luận chương, có liệt kê tài liệu tham khảo và có trích dẫn đúng quy định	1	2	3	4 5
9	Kỹ năng viết xuất sắc (cấu trúc câu chuẩn, văn phong khoa học, lập luận logic và có cơ sở, từ vựng sử dụng phù hợp v.v.)	1	2	3	4 5
Thành tựu nghiên cứu khoa học (5) (chọn 1 trong 3 trường hợp)					
10a	Có bài báo khoa học được đăng hoặc chấp nhận đăng/đạt giải SVNC khoa học giải 3 cấp Viện trở lên/các giải thưởng khoa học (quốc tế/trong nước) từ giải 3 trở lên/ Có đăng ký bằng phát minh sáng chế	5			
10b	Được báo cáo tại hội đồng cấp Viện trong hội nghị sinh viên nghiên cứu khoa học nhưng không đạt giải từ giải 3 trở lên/Đạt giải khuyến khích trong các kỳ thi quốc gia và quốc tế khác về chuyên ngành như TI contest.	2			
10c	Không có thành tích về nghiên cứu khoa học	0			
Điểm tổng		/50			
Điểm tổng quy đổi về thang 10					

*Nhận xét khác của cán bộ phản biện*

.....

.....

.....

.....

.....

.....

Ngày: ... / ... / 20...

**Người nhận xét**  
(Ký và ghi rõ họ tên)

## LỜI NÓI ĐẦU

Ngày nay ung thư dạ dày là một trong những căn bệnh nguy hiểm trên toàn thế giới, tuy nhiên loại ung thư này có ít triệu chứng ở giai đoạn đầu. Chính vì vậy, việc cung cấp các công cụ và kỹ thuật tốt để chẩn đoán có thể giúp việc phát hiện sớm và điều trị là vô cùng cần thiết. Các kỹ thuật chụp, phân tích và xử lý ảnh y tế để chẩn đoán ung thư đã và đang được phát triển trong thế kỷ qua, đặc biệt là kỹ thuật nội soi. Nội soi cho phép thu được ảnh của các mô trong dạ dày với độ phân giải tốt. Các kỹ thuật để phân tích những hình ảnh nội soi này đang phát triển và được sử dụng trong hệ thống Chẩn đoán bằng máy tính (Computer Aided Diagnosis - CAD).

Trong hệ tiêu hóa, các tiến bộ về mặt kỹ thuật của hệ thống CAD đã có thể giúp các bác sĩ chụp và phân tích được các ảnh nội soi. Tuy nhiên, việc nhận biết ung thư trong đường tiêu hóa là một vấn đề hết sức phức tạp, và thường chỉ có các bác sĩ được đào tạo bài bản và có kinh nghiệm thì mới có tỷ lệ chẩn đoán thành công cao. Trong quá khứ, một số giải pháp cho bài toán phân loại ảnh nội soi dạ dày để trợ giúp các bác sĩ chẩn đoán đã được ứng dụng. Vì vậy trong luận văn này, chúng em đã đề xuất một số phương pháp trích xuất đặc trưng cho ảnh nội soi dạ dày để giải quyết cho bài toán phân loại trên. Luận văn sẽ tập trung trình bày chính vào đề tài: “Phân loại ảnh ung thư dạ dày dựa vào các đặc trưng”.

Để có thể hoàn thành đồ án này, chúng tôi xin chân thành cảm ơn TS. Nguyễn Việt Dũng đã tạo điều kiện giúp đỡ chúng tôi. Cũng xin cảm ơn các bạn hiện đang làm việc tại phòng lab “Phân tích, Xử lý tín hiệu và ảnh y sinh” phòng 308A-C9 đã hỗ trợ chúng tôi trong quá trình thực hiện đồ án. Cuối cùng, chúng tôi xin cảm ơn bạn bè và người thân đã đứng sau ủng hộ vô điều kiện về mọi mặt để có thể hoàn thành đồ án tốt nghiệp này.

Chúng tôi xin chân thành cảm ơn!



## **LỜI CAM ĐOAN**

Chúng tôi là Trương Thanh Hiền, mã số sinh viên 20141612, sinh viên lớp ĐT 05, khóa 59 và Phó Thị Hà Anh, mã số sinh viên 20140217, sinh viên lớp ĐT 06, khóa 59. Người hướng dẫn là TS. Nguyễn Việt Dũng. Chúng tôi xin cam đoan toàn bộ nội dung được trình bày trong đồ án Phân loại ảnh ung thư dạ dày sử dụng các đặc trưng là kết quả quá trình tìm hiểu và nghiên cứu của chúng tôi. Các dữ liệu được nêu trong đồ án là hoàn toàn trung thực, phản ánh đúng kết quả đo đạc thực tế. Mọi thông tin trích dẫn đều tuân thủ các quy định về sở hữu trí tuệ; các tài liệu tham khảo được liệt kê rõ ràng. Chúng tôi xin chịu hoàn toàn trách nhiệm với những nội dung được viết trong đồ án này.

Hà Nội, ngày 11 tháng 06 năm 2019

**Người cam đoan**

**Trương Thanh Hiền**

**Phó Thị Hà Anh**



# MỤC LỤC

DANH MỤC KÝ HIỆU VÀ CHỮ VIẾT TẮT .....	i
DANH MỤC HÌNH VẼ.....	i
DANH MỤC BẢNG BIỂU .....	iii
TÓM TẮT ĐỒ ÁN.....	v
ABSTRACT .....	vi
MỞ ĐẦU.....	vii
CHƯƠNG 1. TỔNG QUAN VỀ NỘI SOI VÀ UNG THƯ DẠ DÀY .....	1
<i>1.1 Khái niệm về nội soi .....</i>	<i>1</i>
1.1.1 Khái niệm về nội soi.....	1
1.1.2 Các kỹ thuật nội soi .....	2
<i>1.2 Khái niệm về ung thư dạ dày .....</i>	<i>4</i>
<i>1.3 Tầm quan trọng và ý nghĩa của phân loại ảnh nội soi về ung thư dạ dày.....</i>	<i>6</i>
<i>1.4 Tình hình nghiên cứu hiện nay.....</i>	<i>6</i>
<i>1.5 Hướng tiếp cận bài toán phân loại ảnh nội soi dạ dày.....</i>	<i>7</i>
<i>1.6 Mục đích và phạm vi nghiên cứu.....</i>	<i>7</i>
<i>Tổng kết chương.....</i>	<i>8</i>
CHƯƠNG 2. CƠ SỞ LÝ THUYẾT .....	9
<i>2.1 Trích chọn đặc trưng ảnh .....</i>	<i>9</i>
2.1.1 Khái niệm và phân loại đặc trưng ảnh .....	9
2.1.2 Đặc trưng tăng tốc mạnh mẽ SURF.....	11
2.1.3 Túi từ trực quan BoW ( Bag Of Visual Words) .....	15
2.1.4 Đặc trưng sai khác xác suất ngược khối BDIP (Block Difference Of Inverse Probabilities) .....	19
2.1.5 Đặc trưng biến thiên hệ số tương quan cục bộ khối BVLC (Block Variation Of Local Correlation Coefficients) .....	20
<i>2.2 Giới thiệu về học máy (Machine Learning) và thuật toán máy vector hỗ trợ SVM .....</i>	<i>21</i>
2.2.1 Giới thiệu về học máy (Machine Learning) .....	21

2.2.2 Một số phương pháp học máy phổ biến.....	22
2.2.3 Thuật toán máy vector hỗ trợ SVM phân lớp dữ liệu .....	23
2.2.4 Một số thuật toán phân loại khác được sử dụng trong thử nghiệm.....	27
<b>Tổng kết chương .....</b>	<b>30</b>
<b>CHƯƠNG 3. ĐỀ XUẤT PHƯƠNG PHÁP PHÂN LOẠI ẢNH NỘI SOI DẠ DÀY.....</b>	<b>31</b>
<b>3.1 Đề xuất các phương pháp trích chọn đặc trưng SURF, BDIP, BVLC .....</b>	<b>31</b>
<b>3.2 Chi tiết các bước chính trong mô hình .....</b>	<b>32</b>
3.2.1 Biến đổi ảnh RGB thành ảnh xám .....	32
3.2.2 Trích chọn đặc trưng SURF, BDIP và BVLC .....	33
3.2.3 Đào tạo và kiểm định bộ phân loại, tính độ chính xác của mô hình phân loại .....	35
<b>3.3 Đề xuất phương pháp túi từ đặc trưng SURF (SURF-based Bag of Words).....</b>	<b>35</b>
<b>Tổng kết chương .....</b>	<b>36</b>
<b>CHƯƠNG 4. KẾT QUẢ THỬ NGHIỆM .....</b>	<b>38</b>
<b>4.1 Cơ sở dữ liệu.....</b>	<b>38</b>
<b>4.2 Phương pháp đánh giá .....</b>	<b>39</b>
<b>4.3 Kết quả thử nghiệm.....</b>	<b>41</b>
4.3.1 Kết quả thử nghiệm đối với phân loại đặc trưng SURF.....	41
4.3.2 Kết quả thử nghiệm đối với phân loại túi từ đặc trưng SURF (SURF-based-BoW) .....	43
4.3.3 Kết quả thử nghiệm đối với phân loại đặc trưng BDIP và BVLC .....	44
4.3.4 So sánh các phương pháp đề xuất với các phương pháp khác .....	56
<b>Tổng kết chương .....</b>	<b>56</b>
<b>KẾT LUẬN .....</b>	<b>58</b>
<b>TÀI LIỆU THAM KHẢO .....</b>	<b>59</b>



## DANH MỤC KÝ HIỆU VÀ CHỮ VIẾT TẮT

CAD	Computer Aided Diagnosis
SURF	Speed Up Robust Features
SIFT	Scale-Invariant Feature Transform
BDIP	Block Difference of Inverse Probabilities
BVLC	Block Variation Of Local Correlation Coefficients
SVM	Suport Vector Machine
RGB	Red-Green-Blue
Acc	Accuracy
AUC	Area Under the Curve
CH	Chromoendoscope
GLDM	Gray Level Difference Matrix
BoW	Bag of Word
KNN	K Nearest Neighbors
ROC	Receiver Operating Characteristic
CT	Classification Tree



# DANH MỤC HÌNH VẼ

Hình 1. 1 Minh họa 1 ảnh nội soi dạ dày.....	2
Hình 1. 2 Minh họa 7 công nghệ nội soi [1].....	4
Hình 1. 3 Ảnh ví dụ về ung thư dạ dày [3].....	5
Hình 1. 4 Các giai đoạn ung thư dạ dày [4].....	5
Hình 1. 5 Ví dụ hai ảnh nội soi dạ dày. Ảnh a) là ảnh nội soi dạ dày bị ung thư, ảnh b) là ảnh nội soi của người bình thường .....	6
Hình 2. 1 Mô hình phân loại đặc trưng ảnh.....	9
Hình 2. 2 Ví dụ đặc trưng ngữ nghĩa:(1) quần jean; (2) tóc dài; (3) quần short [17]....	10
Hình 2. 3 Ví dụ về đặc trưng SURF của một ảnh nội soi dạ dày .....	11
Hình 2. 4 Xấp xỉ đạo hàm cấp 2 hàm Gaussian bằng hộp lọc.....	12
Hình 2. 5 Lọc Haar wavelet để tính sự ảnh hưởng trên hai hướng $x$ và $y$ .....	13
Hình 2. 6 Vùng hình tròn xung quanh và hướng đại diện cho điểm đặc trưng.....	14
Hình 2. 7 4x4 hình vuông con xung quanh điểm đặc trưng .....	14
Hình 2. 8 Mô tả tạo túi từ đặc trưng BoW [12] .....	16
Hình 2. 9 Minh họa tạo đồ thị từ mã – vector đặc trưng [12] .....	17
Hình 2. 10 Mô hình BoW chứa mặt người [20] .....	18
Hình 2. 11 Mô hình BoW cho ảnh ô tô [20].....	19
Hình 2. 12 Ảnh xám và ảnh BDIP tương ứng: (a) Ảnh gốc, (b) Ảnh BDIP [13].....	20
Hình 2. 13 Khối kích thước 2 x 2 (nét liền) và các khối sau khi dịch (nét đứt) theo 4 hướng: (a) $\rho(0,1)$ , (b) $\rho(1,0)$ , (c) $\rho(1,1)$ và (d) $\rho(1,-1)$ [13] .....	20
Hình 2. 14 Ảnh BVLC cho các ảnh gốc ở Hình 2.10 [13] .....	21
Hình 2. 15 Các ứng dụng sử dụng Machine Learning [22].....	22
Hình 2. 16 Tập dữ liệu được phân chia tuyến tính [25] .....	25
Hình 2. 17 Tập dữ liệu phân chia tuyến tính nhưng có nhiễu [25].....	26
Hình 2. 18 Tập dữ liệu không thể phân chia tuyến tính [25].....	26
Hình 2. 19 Bản đồ phân loại của 1NN [26].....	28
Hình 2. 20 Ví dụ về một mô hình cây quyết định [29].....	29
Hình 3. 1 Sơ đồ khối thực hiện bài toán phân loại ảnh nội soi dạ dày.....	31
Hình 3. 2 Một ảnh trong bộ dữ liệu được chuyển thành ảnh xám.....	32
Hình 3. 3 Ảnh xám trong Hình 3.2 được lấy 100 vector đặc trưng SURF.....	32
Hình 3. 4 Thang màu xám của ảnh.....	33
Hình 3. 5 Ảnh xám trong Hình 3.2 được lấy 50 đặc trưng SURF mạnh nhất (a) và 1000 đặc trưng SURF mạnh nhất (b) .....	33
Hình 3. 6 Ảnh xám trong Hình 3.2 được lấy đặc trưng BDIP.....	34
Hình 3. 7 Ảnh xám trong Hình 3.2 được lấy đặc trưng BVLC .....	35
Hình 3. 8 Sơ đồ khối mô hình phân loại sử dụng túi từ đặc trưng SURF .....	36

Hình 4. 1 a) Một ảnh nội soi thuộc lớp Normal, b) Một ảnh nội soi thuộc lớp Abnormal .....	38
Hình 4. 2 Minh họa ma trận nhầm lẫn [33] .....	39
Hình 4. 3 Minh họa đường cong ROC [34] .....	40
Hình 4. 4 Minh họa AUC [34].....	41
Hình 4. 5 Đường cong ROC cho 3 bộ phân loại ứng với đặc trưng SURF [31].....	42
Hình 4. 6 Đường cong ROC 3 bộ phân loại khác nhau đối với túi từ đặc trưng SURF .....	44

## DANH MỤC BẢNG BIỂU

Bảng 4. 1 Ảnh hưởng của số lượng đặc trưng SURF được chọn lên kết quả phân loại [31].....	41
Bảng 4. 2 Kết quả phân loại của các bộ phân loại khác nhau kết hợp với đặc trưng SURF và KAZE [31] .....	42
Bảng 4. 3 Kết quả thay đổi số lượng đặc trưng trên ảnh .....	43
Bảng 4. 4 Kết quả thay đổi tỷ lệ huấn luyện – kiểm định .....	43
Bảng 4. 5 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 6x6 không chồng lấp .....	44
Bảng 4. 6 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 7x7 không chồng lấp .....	45
Bảng 4. 7 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 8x8 không chồng lấp .....	45
Bảng 4. 8 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 10x10 không chồng lấp.....	45
Bảng 4. 9 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 7x7 không chồng lấp với tỷ lệ trainingsets: testingsets=6:4.....	46
Bảng 4. 10 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 7x7 không chồng lấp với tỷ lệ trainingsets: testingsets=5:5.....	46
Bảng 4. 11 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 7x7 không chồng lấp với tỷ lệ trainingsets: testingsets=4:6.....	47
Bảng 4. 12 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 7x7 không chồng lấp với tỷ lệ trainingsets: testingsets=3:7.....	47
Bảng 4. 13 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 7x7 chồng lấp 1 pixel .....	48
Bảng 4. 14 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 7x7 chồng lấp 3 pixel .....	48
Bảng 4. 15 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 7x7 chồng lấp 5 pixel .....	48
Bảng 4. 16 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 7x7 chồng lấp 6 pixel .....	49

Bảng 4. 17 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 10x10 không chồng lấp .....	50
Bảng 4. 18 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 15x15 không chồng lấp .....	50
Bảng 4. 19 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 25x25 không chồng lấp .....	50
Bảng 4. 20 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 35x35 không chồng lấp .....	51
Bảng 4. 21 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 45x45 không chồng lấp .....	51
Bảng 4. 22 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 50x50 không chồng lấp .....	52
Bảng 4. 23 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 45x45 không chồng lấp với tỷ lệ trainingsests: testingsests=6:4.....	52
Bảng 4. 24 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 45x45 không chồng lấp với tỷ lệ trainingsests: testingsests=5:5.....	53
Bảng 4. 25 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 45x45 không chồng lấp với tỷ lệ trainingsests: testingsests=4:6.....	53
Bảng 4. 26 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 45x45 không chồng lấp với tỷ lệ trainingsests: testingsests=3:7.....	53
Bảng 4. 27 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 45x45 chồng lấp 5 pixel .....	54
Bảng 4. 28 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 45x45 chồng lấp 10 pixel .....	54
Bảng 4. 29 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 45x45 chồng lấp 15 pixel .....	55
Bảng 4. 30 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 45x45 chồng lấp 20 pixel .....	55
Bảng 4. 31 So sánh các đặc trưng và bộ phân loại.....	56

## TÓM TẮT ĐỒ ÁN

Trong những năm gần đây, ung thư đường tiêu hóa đang có xu hướng gia tăng và thực sự đáng báo động. Ở Việt Nam, mỗi năm ghi nhận có hơn 8000 ca mắc mới và gần 5000 ca tử vong do ung thư đường tiêu hóa (ung thư đại tràng và trực tràng). Ngoài ra, Việt Nam còn được ghi nhận đứng thứ 18 trong số 20 nước có tỷ lệ ung thư dạ dày cao nhất thế giới. Một trong những nguyên nhân gây tử vong cao là bệnh không có triệu chứng rõ ràng, nhất là ở những giai đoạn đầu, gần giống với các bệnh lý thông thường nên rất khó phát hiện. Tuy nhiên, việc phát hiện và chẩn đoán bệnh vẫn còn gặp nhiều khó khăn do phụ thuộc nhiều vào chất lượng hình ảnh cũng như trình độ chuyên môn của bác sĩ. Do đó, việc cung cấp các công cụ hỗ trợ chẩn đoán sớm là rất cần thiết để hỗ trợ điều trị sớm và giảm tỷ lệ tử vong. Từ thực tế này, nhóm nghiên cứu đề xuất phương pháp xây dựng mô hình phân loại ảnh ung thư dạ dày bằng Machine Learning. Từ bộ dataset ảnh nội soi dạ dày Chromoendoscope (Viện Ung thư Bờ Đào Nha), nhóm nghiên cứu đề xuất biểu diễn lại ảnh dưới 4 dạng: đặc trưng SURF (Speeded-up Robust Features), túi từ đặc trưng SURF (SURF-based BoW), đặc trưng về toán tử sai khác xác suất ngược dạng khối BDIP (Block Difference of Inverse Probabilities) và đặc trưng về hệ số biến thiên tương quan cục bộ dạng khối BVLC (Block Variation Of Local Correlation Coefficients). Các đặc trưng này sau đó được đưa vào bộ phân loại SVM (Support Vector Machine) để phân loại ảnh thành bình thường hay bất thường. Độ chính xác và AUC (Area Under the Curve) trung bình của 4 loại đặc trưng lần lượt là: 81.5% và 0.92, 98.5% và 0.99, 88.1% và 0.93, 78.7% và 0.81, tương đối khả quan so với kết quả tương ứng 90% và 0.93 của phương pháp Gabor kết hợp thuật toán Genetic [10] khi thử nghiệm trên cùng một cơ sở dữ liệu. Phương pháp mới này đã đạt được kết quả như mong đợi là nâng cao độ chính xác và độ tin cậy trong việc phân loại ảnh nội soi dạ dày, cũng như giảm thiểu thời gian thực hiện phân loại.

## ABSTRACT

In recent years, gastrointestinal cancer has been increasingly found in people of all ages. In Vietnam, more than 8000 new and approximately 5000 death cases are recorded every year due to this disease (including large-intestine and rectum cancer). Furthermore, Vietnam also is at the 18 over 20 countries which have the highest gastrointestinal cancer ratio in the world. One of the reasons leading to the high mortality is that gastrointestinal cancer either has no clear symptom in the early stage or appears to be similar to the common pathologies, making it even harder to be detected. Additionally, there are some difficulties existing in both diagnosis and detection procedures due to their dependencies on image quality and qualification of the physicians. From the fact that it is important to create a tool helping the early diagnosis in order reduce mortality ratio, we propose a Machine Learning method which builds the gastrointestinal cancer image classifier. Based on the gastic chromoendoscopy image dataset (from Portuguese Institute of Oncology (IPO) Porto, Portugal), we represent the images using 4 types of features: Speeded-up Robust Features feature (SURF), SURF based Bag of Word feature, Block Difference of Inverse Probabilities (BDIP) feature and Block Variation Of Local Correlation Coefficients (BVLC) feature. Then, the feature sets are classified into 'normal' or 'abnormal' by a Support Vector Machine model. The average accuracy and Area-Under-Curve (AUC) of the proposed method for each type of feature is 81.5% and 0.92, 98.5% and 0.99, 88.1% and 0.93, and 78.7% and 0.81. On the same dataset, this result is comparable to the one of the Gabor method combined with Genetic algorithm [10], which is 90% and 0.93. In general, the proposed method has successfully improved the accuracy and the reliability in gastic chromoendoscopy image classification, as well as reduced the classification time.



## MỞ ĐẦU

Ngày nay, với sự trợ giúp của các hệ thống chẩn đoán hỗ trợ máy tính, ung thư đường tiêu hóa có thể được theo dõi phát hiện và phân loại thành các giai đoạn khác nhau. Tuy nhiên, vẫn còn tồn tại một số vấn đề, chẳng hạn như sự phụ thuộc của các đặc điểm của ảnh và kinh nghiệm của các bác sĩ có thể cho kết quả chẩn đoán không hoàn toàn chính xác. Để tránh những vấn đề này, cũng như hỗ trợ các bác sĩ trong chẩn đoán bệnh, đề án này đề xuất sử dụng các đặc trưng của ảnh để phân loại ảnh nội soi dạ dày bị ung thư và dạ dày bình thường. Vì vậy mục tiêu của đề án này là: Tìm hiểu cơ sở lý thuyết về ảnh nội soi và nghiên cứu phát triển về việc phân loại ảnh nội soi dạ dày từ các loại đặc trưng khác nhau. Đề án có nội dung gồm 4 chương như sau:

Chương 1: Giới thiệu về nội soi và ung thư dạ dày, những thuật ngữ cơ bản, tầm quan trọng của phân loại ảnh nội soi về ung thư dạ dày và ý nghĩa trong thực tế. Chương này chỉ ra những khó khăn, thách thức, hướng tiếp cận bài toán và nêu ra tình hình nghiên cứu trong, ngoài nước về phân loại ảnh nội soi dạ dày.

Chương 2: Trình bày về các phương pháp trích chọn đặc trưng ảnh, lý thuyết về các thuật toán.

Chương 3: Nội dung chương này nêu ra đề xuất giải quyết bài toán phân loại ảnh nội soi, chỉ ra sơ đồ khối và chức năng từng khối của bài toán.

Chương 4: Trình bày toàn bộ những kết quả thử nghiệm qua các bước được nêu ra ở chương 3. So sánh các kết quả đạt được qua thử nghiệm và đưa ra đánh giá về tính hiệu quả của phương pháp thử nghiệm.

Trong quá trình tìm hiểu, phân tích, nhóm nghiên cứu không thể tránh được những khiếm khuyết, sai sót. Mong thầy cô và các bạn góp ý để các tác giả ngày càng hoàn thiện hơn sản phẩm của mình.

**Sinh viên thực hiện**



# CHƯƠNG 1. TỔNG QUAN VỀ NỘI SOI VÀ UNG THU DẠ DÀY

Nội soi dạ dày là phương pháp thông dụng có độ chính xác cao để thăm khám các bệnh lý về dạ dày. Phương pháp này giúp chẩn đoán sớm các tổn thương dạ dày dù nhỏ, những tổn thương niêm mạc dạ dày nhìn chính xác hơn so với chụp X.quang. Nội soi kết hợp với sinh thiết, siêu âm, xét nghiệm giúp cho việc chẩn đoán bệnh chính xác hơn. Ngày nay, với sự phát triển của khoa học và kỹ thuật, nội soi dạ dày - tá tràng không chỉ để chẩn đoán các bệnh lý về dạ dày tá tràng mà còn để điều trị nhiều bệnh lý ở dạ dày tá tràng mà không cần phẫu thuật.

## 1.1 Khái niệm về nội soi

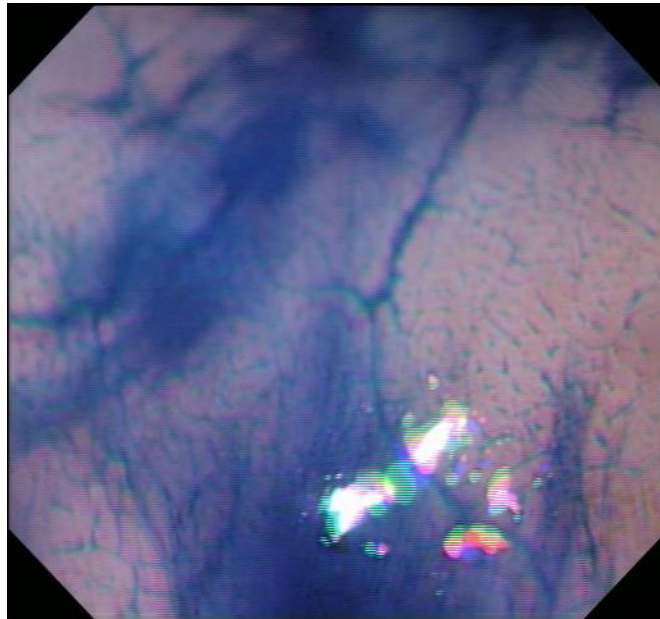
### 1.1.1 Khái niệm về nội soi

Nội soi (Endoscopy) là một kỹ thuật y học hiện đại được ứng dụng trong việc khám bệnh, chẩn đoán bệnh bằng việc sử dụng các loại dụng cụ chuyên biệt để quan sát trực tiếp bên trong các cơ quan của cơ thể. Với kỹ thuật nội soi, người ta có thể quay phim, chụp hình bên trong các cơ quan, lấy dị vật, sinh thiết và thậm chí là thực hiện phẫu thuật nội soi. Nội soi hiện nay được sử dụng trong hầu hết các chuyên khoa: tai mũi họng, tiêu hóa (thực quản, dạ dày, ruột già...) sản, ngoại, tiết niệu, xương khớp, thần kinh, thẩm mỹ...

Thiết bị nội soi là một thiết bị y tế có khả năng thu được hai hình ảnh đồng thời và thậm chí các mẫu mô từ đường tiêu hóa. Mặc dù có rất nhiều cấu hình nhưng các loại nội soi đều có cùng một số đặc điểm:

- Bao gồm một ống mềm dẻo.
- Có khả năng chiếu sáng và thu thập hình ảnh được kết hợp.
- Có một kênh phụ, có thể chụp lấy mẫu sinh thiết.

Trên Hình 1.1 miêu tả một ví dụ về một ảnh nội soi dạ dày.



**Hình 1. 1 Minh họa 1 ảnh nội soi dạ dày**

### **1.1.2 Các kỹ thuật nội soi**

Một số kỹ thuật nội soi tiêu biểu được trình bày dưới đây [1]:

*a) Nội soi độ phân giải chuẩn và cao:* Máy nội soi độ phân giải chuẩn được trang bị chip CCD, thu tín hiệu hình ảnh với độ nét 100-400 pixel và tỷ lệ khung hình là 4:3. Máy nội soi độ phân giải cao có cấu hình tương tự nhưng cho ảnh có độ nét cao hơn. Do đó, bác sĩ có thể phát hiện những thay đổi nhỏ trong niêm mạc dễ dàng hơn. Ảnh được lấy từ máy nội soi độ phân giải cao có độ nét cao gấp 10 lần, hai tỷ lệ khung hình khác nhau và màn hình của nó có thể hiển thị ảnh chất lượng cao trong khi đang tiến hành. Tốc độ khung hình 60 lần/giây làm giảm lượng nhiễu và cho phép chụp các chuyển động nhanh. Xem Hình 1.2a.

*b) Nội soi cho ảnh phóng đại lớn (High Magnification Endoscope - HME):* Công nghệ này sử dụng các ống kính chuyển động điện tử cho phép lấy được ảnh niêm mạc dạ dày chi tiết hơn trong thời gian thực. Ảnh nội soi được phóng to đến 150 lần, trong khi vẫn giữ được độ chi tiết và độ nét cao. Trong HME, ảnh được di chuyển gần màn hình, giảm lượng quan sát được và độ phân giải hình ảnh. Hầu hết các nội soi thông thường đều có khả năng phóng đại điện tử 1.5 lần đến 2 lần nhưng lại cần một bộ xử lý tương thích. Xem Hình 1.2b.

*c) Nội soi tự động phát quang (Autofluorescence Imaging – AFI):* Các máy nội soi gần đây sử dụng kỹ thuật tạo ảnh AFI. AFI phát hiện sự phát quang tự nhiên của các mô,

được phát ra bởi các phân tử kích thích ánh sáng cụ thể. Có thể ghi lại sự khác biệt về màu sắc trong phát xạ thời gian thực. Kỹ thuật này cho phép mô tả đặc điểm mô và có thể được sử dụng để phát hiện một số lượng đáng kể bệnh nhân mắc ung thư giai đoạn đầu. Xem Hình 1.2c.

*d) Nội soi nhuộm màu (Chromoendoscopy):* Kỹ thuật này dựa trên việc tăng cường bề mặt của niêm mạc bằng cách áp dụng các thuốc nhuộm khác nhau. Tùy thuộc vào các giao thức tiếp theo và các thuốc nhuộm khác nhau được sử dụng, làm cho các cấu trúc giải phẫu khác nhau dễ được quan sát hơn. Xem Hình 1.2d.

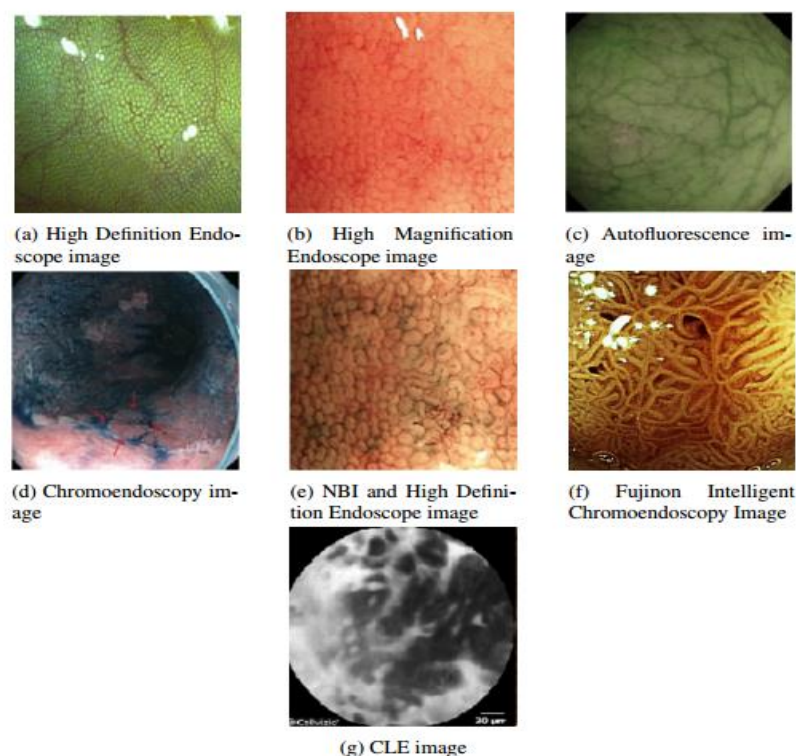
*e) Chụp quang tuyến Laser (CLE):* Công nghệ này cung cấp hình ảnh cực nhỏ độ phân giải cao ở độ phân giải tế bào nhỏ, luồn vào các lớp sâu nhất của niêm mạc dạ dày-ruột. Tuy nhiên phương pháp kết hợp này đòi hỏi người dùng có kinh nghiệm để thao tác chính xác để đạt được kết quả tốt. Xem Hình 1.2e.

*f) I-scan và Nội soi thông minh Fujinon (FICE):* Kỹ thuật này bao gồm ba loại thuật toán: Tăng cường bề mặt, Tăng cường độ tương phản và Tăng cường tông màu. Tăng cường bề mặt tăng cường độ tương phản sáng-tối trên hình ảnh, bằng cách lấy dữ liệu cường độ sáng cho từng pixel. Sau đó, các thuật toán để quan sát chi tiết cấu trúc bề mặt niêm mạc được áp dụng. Tăng cường độ tương phản kỹ thuật số thêm màu xanh lam ở các vùng tương đối tối: dữ liệu cường độ độ chói cho từng pixel được thu được và các bất thường tinh tế xung quanh bề mặt được tính toán tăng cường. Cả hai chức năng tăng cường hoạt động trong thời gian thực mà không làm suy yếu màu sắc ban đầu của cơ quan. Ngoài ra, cả hai đều thích hợp để sàng lọc nội soi để phát hiện khối u đường tiêu hóa ở giai đoạn đầu. Công nghệ i-scan giúp phát hiện, chẩn đoán và điều trị các bệnh đường tiêu hóa dễ dàng hơn.

FICE có thể mô phỏng vô số bước sóng trong thời gian thực. Hệ thống có 10 kênh được thiết kế để khám phá toàn bộ bề mặt niêm mạc. Mỗi kênh tương ứng với ba bộ lọc bước sóng RGB cụ thể, nhưng không có thiết lập nào được sử dụng cụ thể cho riêng tình trạng dạ dày nhất định Xem Hình 1.2f.

*g) Phẫu thuật nội soi viên không dây (WCE):* Việc kiểm tra ruột non là một công việc khó khăn do hình dạng dài và phức tạp. WCE được thiết kế để vượt qua giới hạn này và làm cho quy trình nội soi an toàn, ít xâm lấn hơn và thoải mái hơn cho bệnh nhân. Trong trường hợp này, nội soi không phải là ống linh hoạt mà là một viên nang mà bệnh

nhân nuốt. Viên nang nhỏ được trang bị nguồn sáng, ống kính, máy ảnh, máy phát tín hiệu radio và pin. Được vận chuyển bởi peristalsis, viên nang đi qua hệ thống tiêu hóa trong khoảng 8 giờ và tự động chụp được hơn 50 000 hình ảnh. Chúng được truyền qua không dây tới máy ghi bên trong cơ thể. Hiện tại, WCE không chỉ muốn kiểm tra ruột non mà còn các cơ quan khác như đại tràng hoặc thực quản. Những hạn chế chính của WCE bao gồm việc thiếu khả năng lấy mẫu sinh thiết, trái với các kỹ thuật nội soi khác. Xem Hình 1.2g.



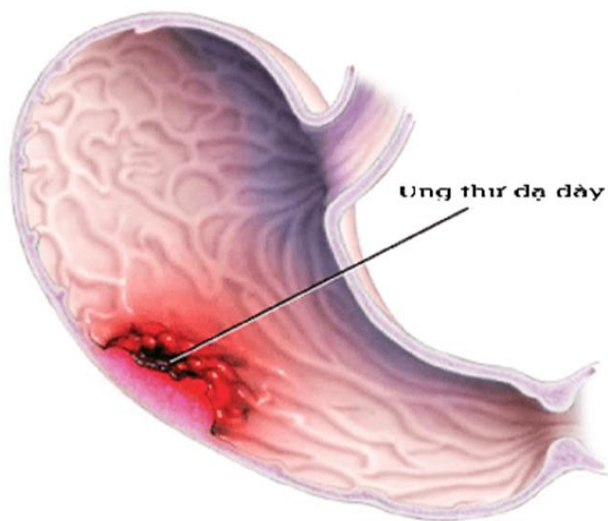
**Hình 1. 2 Minh họa 7 công nghệ nội soi [1]**

## 1.2 Khái niệm về ung thư dạ dày

Ung thư dạ dày là một trong số những loại ung thư phổ biến nhất hiện nay mỗi năm có thể gây ra 800.000 ca tử vong trên khắp thế giới [2]. Ung thư dạ dày là tình trạng phát triển khối u ác tính trong dạ dày. Bệnh xảy ra khi các tế bào trong dạ dày phát triển quá mức dẫn đến hình thành các khối u. Ung thư dạ dày là bệnh thường gặp và rất dễ di căn đến các bộ phận khác. Nếu không được chữa trị kịp thời, bệnh có thể dẫn đến tử vong.

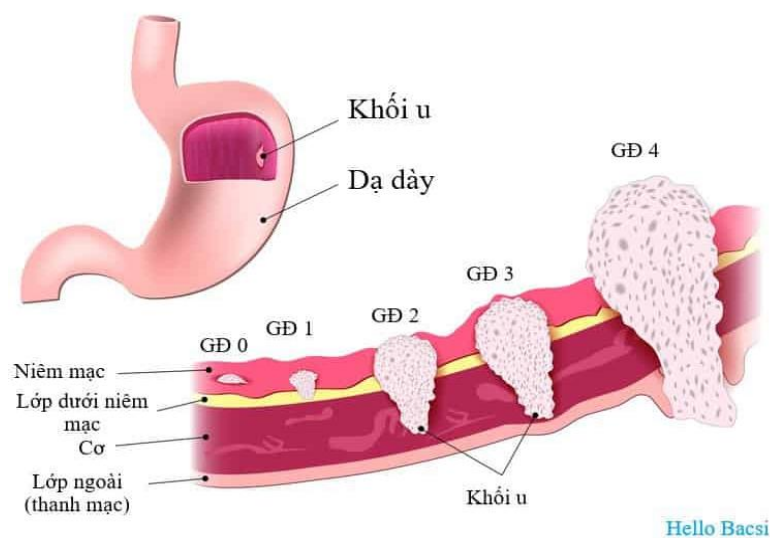
Các giai đoạn ung thư dạ dày [4]:

- Giai đoạn 0: Giai đoạn 0 là ung thư dạ dày giai đoạn đầu. Giai đoạn này còn gọi là ung thư biểu mô khi các tế bào ung thư nằm ở lớp niêm mạc dạ dày.



**Hình 1. 3 Ảnh ví dụ về ung thư dạ dày [3]**

- Giai đoạn 1: Giai đoạn này, các tế bào ung thư đã xâm lấn vào lớp thứ 2 của dạ dày. Giai đoạn này vẫn chưa có nguy hiểm gì và bệnh chưa lây qua các cơ quan khác.
- Giai đoạn 2 Ở ung thư dạ dày giai đoạn 2, các tế bào ung thư đã xâm lấn qua lớp niêm mạc. Giai đoạn này còn gọi là ung thư dưới cơ.
- Giai đoạn 3: Ở giai đoạn ung thư dạ dày này, các tế bào ung thư đã bắt đầu lan ra hạch bạch huyết và các cơ quan ở xa.
- Giai đoạn 4: Đây là ung thư dạ dày giai đoạn cuối, khi các tế bào ung thư đã lan ra khắp cơ thể. Khi người bệnh được chẩn đoán giai đoạn này, cơ hội sống sót là rất ít.



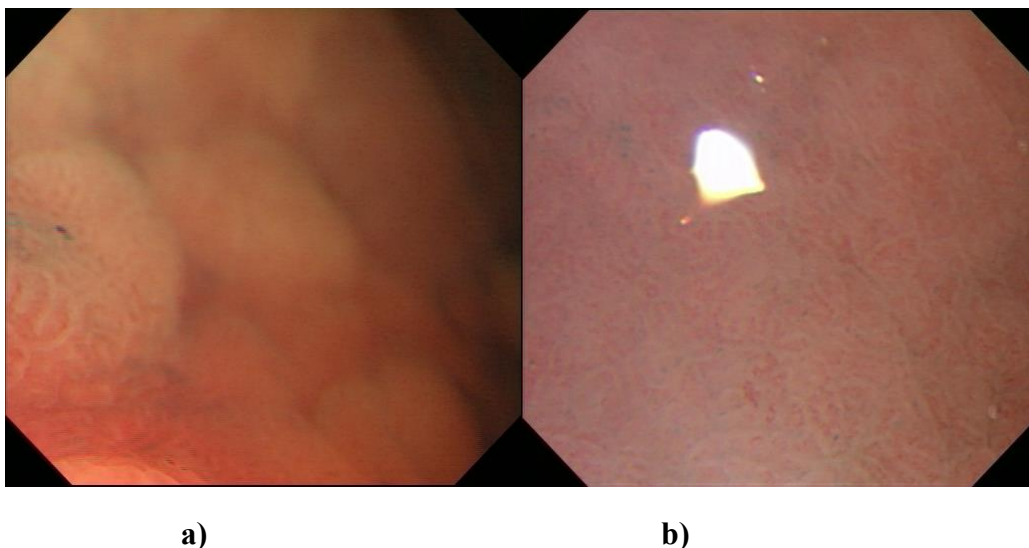
**Hình 1. 4 Các giai đoạn ung thư dạ dày [4]**

### 1.3 Tầm quan trọng và ý nghĩa của phân loại ảnh nội soi về ung thư dạ dày

Việc chẩn đoán từ ảnh nội soi hiện nay còn rất nhiều thách thức. Ảnh nội soi giữa dạ dày của người bị ung thư và người bình thường không có bệnh có nhiều nét tương đồng khó phân biệt. Kết quả chẩn đoán còn phụ thuộc nhiều vào kinh nghiệm và trạng thái của bác sĩ.

Vì vậy việc nghiên cứu và tạo ra một mô hình phân loại ảnh nội soi về ung thư đường tiêu hóa nói chung và ung thư dạ dày nói riêng để giúp các bác sĩ chẩn đoán chính xác hơn và các bác sĩ thiếu kinh nghiệm được loại bớt những trường hợp mập mờ là điều cần thiết trong y học hiện đại.

Việc phân loại ảnh nội soi ung thư dạ dày có thể đưa ra liệu pháp điều trị thích hợp cho từng loại ung thư và từng giai đoạn ung thư của từng bệnh nhân, hứa hẹn sẽ giảm bớt tỷ lệ tử vong trong tương lai.



**Hình 1. 5 Ví dụ hai ảnh nội soi dạ dày. Ảnh a) là ảnh nội soi dạ dày bị ung thư, ảnh b) là ảnh nội soi của người bình thường**

### 1.4 Tình hình nghiên cứu hiện nay

Để giải quyết bài toán phân loại ảnh nội soi dạ dày, các nghiên cứu thường theo hướng chính là trích chọn đặc trưng (features extraction) của ảnh.

Trong [5], các đặc trưng kết cấu dựa trên ma trận chênh lệch mức xám (Gray Level Difference Matrix - GLDM) đã được sử dụng để mô tả các biến thể khác nhau trong ảnh nội soi phóng đại. Riaz, trong [6], đã sử dụng các bộ lọc Gabor kết hợp với ma trận tự tương quan để mô tả đặc trưng kết cấu. Trong [7], sự phân tích ngữ nghĩa trong không



gian đặc trưng được áp dụng cho nội soi bằng laser đồng tiêu (Confocal Laser Endomicroscopy - CLE) đã được đề xuất. Trong [8], các tác giả đã đề xuất một phương pháp để tăng cường sự phân biệt trong đặc trưng SIFT (Scale-Invariant Feature Transform). Các tính năng LBP cũng đã được trích xuất để phân loại ảnh nội soi viên nang của ruột non bất thường (Wireless Capsule Endoscopy - WCE) từ bệnh nhân mắc bệnh Constantinescu trong [9]. Nghiên cứu của Hussam Ali [10] thì tập trung vào việc thiết kế một bộ trích xuất đặc trưng dựa trên các bộ lọc Gabor và thuật toán Genetic (GA). Các phương pháp trên đều có một khuyết điểm là độ chính xác còn chưa thực sự cao để có thể đưa vào áp dụng thực tế.

### **1.5 Hướng tiếp cận bài toán phân loại ảnh nội soi dạ dày**

Đề tài này giải quyết vấn đề phân loại ảnh nội soi dạ dày, khi đó mỗi nhãn “bình thường” và “bất thường” sẽ được gán cho từng ảnh. Trong hướng tiếp cận này, hai tập ảnh sẽ được trích xuất ra vector đặc trưng của từng ảnh. Sau đó các vector đặc trưng này được huấn luyện mô hình phân loại và đưa ra dự đoán cho từng ảnh. Trong bài toán phân loại này, việc chọn được loại đặc trưng để dự đoán phân loại là vô cùng quan trọng. Vì vậy đề án đã chọn một vài đặc trưng khác nhau và kết hợp các đặc trưng để thử nghiệm. Việc này giúp so sánh và xác định lựa chọn loại đặc trưng nào sẽ cho độ chính xác cao hơn để đưa vào ứng dụng.

### **1.6 Mục đích và phạm vi nghiên cứu**

Trong phạm vi đề tài, luận văn này chỉ trình bày những nội dung liên quan đến trích chọn đặc trưng ảnh trên cơ sở dữ liệu trên ảnh nội soi dạ dày thu được từ máy nội soi Olympus CV-180 tại Bệnh viện Portuguese Institute of Oncology (IPO) ở Porto, Bồ Đào Nha, với các đặc trưng tăng tốc mạnh mẽ SURF (Speeded-up Robust Features) [11], kết hợp SURF và túi từ trực quan (Bag of Visual Words – BoW) [12], đặc trưng sai khác xác suất ngược khối BDIP (Block Difference Of Inverse Probabilities) [13], đặc trưng biến thiên hệ số tương quan cục bộ khối BVLC (Block Variation Of Local Correlation Coefficients) [13] và xây dựng mô hình phân loại máy vector hỗ trợ (Support Vector Machine – SVM) cho các đặc trưng. Từ những nội dung được nêu trên, mục đích của đề án nhằm đánh giá các đặc trưng trên ảnh nội soi dựa vào độ chính xác của quá trình phân loại ảnh.

## **Tổng kết chương**

Chương 1 đã trình bày tổng quan về ảnh nội soi, tầm quan trọng và ý nghĩa của phân loại ảnh nội soi đối với rất nhiều ứng dụng thực tế như theo dõi, đánh giá và giúp đỡ các bác sĩ trong việc chẩn đoán.... Chỉ ra được tình hình nghiên cứu hiện nay về bài toán phân loại ảnh nội soi dạ dày. Với những ý nghĩa đó, phân loại ảnh nội soi đang gặp phải rất nhiều những thách thức lớn đòi hỏi quá trình nghiên cứu lâu dài để giải quyết.

## CHƯƠNG 2. CƠ SỞ LÝ THUYẾT

Chương này trình bày về các phương pháp trích chọn đặc trưng, khái niệm về học máy (Machine Learning) và đưa ra mô hình phân loại máy vector hỗ trợ SVM trong việc phân loại ảnh nội soi dạ dày.

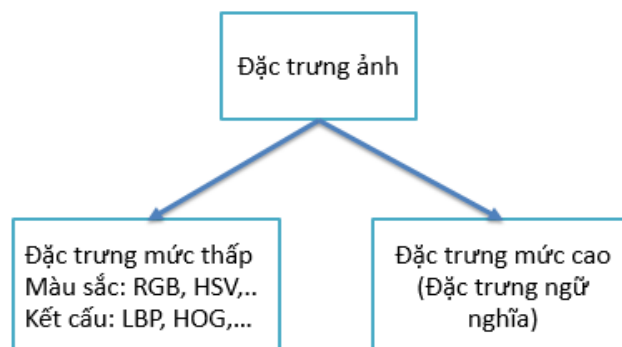
### 2.1 Trích chọn đặc trưng ảnh

#### 2.1.1 Khái niệm và phân loại đặc trưng ảnh

##### 2.1.1.1 Khái niệm

Trong phạm vi xử lý ảnh, đặc trưng của ảnh là một phần thông tin của ảnh thích hợp cho các nhiệm vụ tính toán liên quan đến một ứng dụng nhất định. Những đặc trưng đó có thể là các kết cấu đặc biệt trong ảnh như các điểm ảnh, các cạnh của một đối tượng hoặc một đối tượng nào đó có trong ảnh. Mặt khác, các đặc trưng của ảnh cũng có thể là kết quả của một phép biến đổi toàn diện hoặc là các phương pháp phát hiện điểm đặc trưng được áp dụng trên toàn bộ ảnh đó.

##### 2.1.1.2 Phân loại



**Hình 2. 1 Mô hình phân loại đặc trưng ảnh**

Theo Hình 2.1, đặc trưng ảnh được chia thành hai loại chính là: Đặc trưng mức thấp và đặc trưng mức cao.

##### a) Đặc trưng mức thấp

Trong các loại đặc trưng mức thấp, màu sắc được xem là một loại đặc trưng đơn giản nhưng rất hiệu quả được sử dụng rất rộng rãi trong các công trình nghiên cứu. Bên cạnh

đó, đặc trưng hình dáng cũng được giới thiệu và công bố. Đặc trưng dựa trên sinh học (Biologically Inspired Feature – BIF), loại đặc trưng khá mới này cũng nhận được nhiều sự quan tâm. Ngoài ra, các đặc trưng về kết cấu thường được sử dụng là Histogram of Oriented Gradients (HOG) được đề cập trong [14], Scale-Invariant Feature Transform (SIFT) [15], Speeded Up Robust Features (SURF) [11], PHOG, Haar-like, Local Binary Patterns (LBP). Một vài đặc trưng, chẳng hạn như HOG, có cách biểu diễn đặc trưng của riêng nó. Để tăng độ chính xác khi so sánh ảnh hai nhân vật với nhau, người ta thường chia ảnh nhân vật thành nhiều phần để so khớp từng phần.

*b) Đặc trưng mức cao (khái niệm, ngữ nghĩa)*

Các đặc trưng mức thấp thường rất nhạy cảm, dễ bị ảnh hưởng bởi nhiều yếu tố như ánh nền, sự thay đổi cường độ sáng qua các camera khác nhau,... nên thời gian gần đây, các đặc trưng mức cao đã được giới thiệu và sử dụng rộng rãi như trong [16]. Đặc trưng ngữ nghĩa là những đặc trưng gắn liền với các khái niệm gần gũi với con người và thường được con người sử dụng để nhận dạng các đối tượng. Các đặc trưng ngữ nghĩa cho nhân vật có thể là tóc ngắn, tóc dài, mang giày, quần short, áo thun ... Các đặc trưng ngữ nghĩa được phát hiện dựa trên những đặc trưng cấp thấp và các bộ phát hiện các đặc trưng ngữ nghĩa có độ chính xác cao, giúp nâng cao độ chính xác của bài toán tái nhận dạng nhân vật. Hình 2.2: Ví dụ đặc trưng ngữ nghĩa:



**Hình 2. 2 Ví dụ đặc trưng ngữ nghĩa:(1) quần jean; (2) tóc dài; (3) quần short [17]**

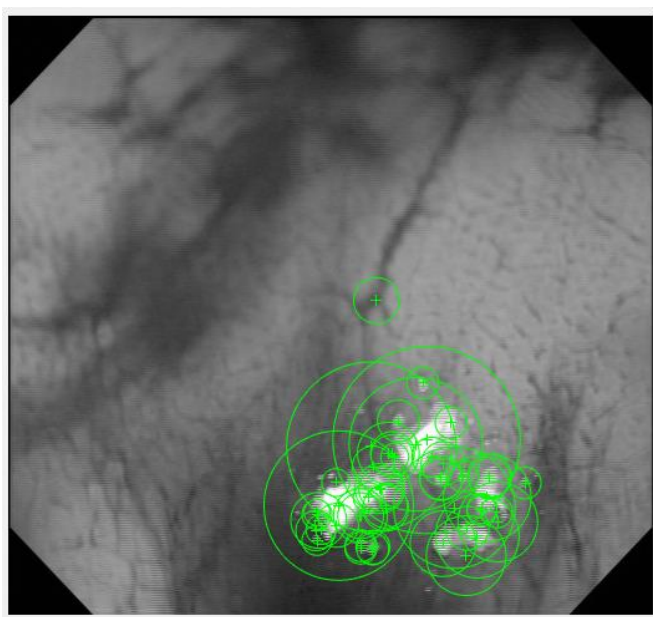
Hình 2.2 chỉ ra ví dụ về một số đặc trưng ngữ nghĩa nổi bật của người trong bức ảnh. Ta thấy ở ảnh thứ nhất có một người đàn ông với điểm nổi bật là mặc quần jean. Ảnh thứ hai chụp một cô gái có mái tóc đen dài, và ở ảnh thứ ba chụp một người đàn ông với điểm nổi bật là mặc quần short.

### 2.1.2 Đặc trưng tăng tốc mạnh mẽ SURF

Trong thị giác máy tính, đặc trưng SURF là một loại đặc trưng kết cấu có thể được sử dụng cho các bài toán như nhận dạng đối tượng, phân loại ảnh hoặc tái tạo ảnh 3D.

Trước khi SURF được phát minh và công nhận thì thuật toán anh em của nó là đặc trưng SIFT được coi là một trong những thuật toán trích chọn đặc trưng mạnh mẽ nhất. Dù vậy, tốc độ xử lý của SIFT vẫn còn rất chậm và không phù hợp với các ứng dụng thời gian thực. Để giải quyết bài toán này, người ta đã giới thiệu thuật toán trích chọn đặc trưng SURF (Speed Up Robust Features) có được sự cân bằng giữa yêu cầu tốc độ và sự chính xác.

Đặc trưng tối ưu cả hai giai đoạn phát hiện đặc trưng (detector) và mô tả đặc trưng (descriptor) về mặt thời gian tính toán nhưng vẫn giữ được tính bền vững của đặc trưng. Bộ phát hiện đặc trưng của SURF sử dụng phép xấp xỉ trên ma trận Hessian và ảnh tích hợp (integral image) để làm giảm thời gian tính toán một cách đáng kể. Bộ mô tả đặc trưng tương tự như đặc trưng SIFT, sử dụng vector 64 chiều chứa thông tin biến thiên trên ảnh dựa trên sự phân phối bậc nhất Haar wavelet tác động trên trục x và y, kết hợp với ảnh tích lũy làm tăng tốc độ tính toán. SURF được mô tả bởi vector có số chiều ít hơn SIFT nên tốc độ so khớp nhanh hơn, tuy nhiên độ bền vững vẫn được đảm bảo. Hơn thế nữa, bằng việc đánh chỉ mục dựa trên dấu của Laplacian, đặc trưng SURF không chỉ giữ tính bền vững cho đặc trưng mà còn làm tăng tốc độ so khớp (tăng gấp 2 trong trường hợp tốt nhất).



Hình 2. 3 Ví dụ về đặc trưng SURF của một ảnh nội soi dạ dày

Ảnh tích hợp có công thức sau:

$$I_z(X) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(i, j) \quad (2.1)$$

Thuật toán của kỹ thuật SURF để trích xuất đặc trưng gồm 2 bước dưới đây:

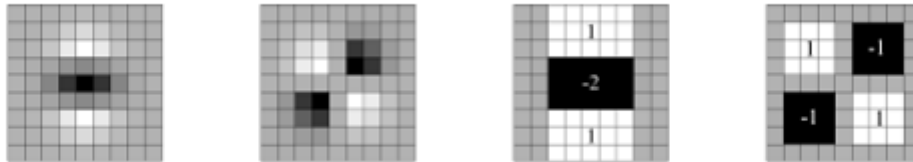
- Sử dụng bộ dò Fast-Hessian để xác định các điểm nổi bật: Bộ dò này được xây dựng dựa trên ma trận Hessian vì hiệu suất tốt của nó về thời gian tính toán và độ chính xác, kết hợp với việc sử dụng ảnh tích hợp để giảm thời gian tính toán đặc trưng.

Ma trận Hessian được định nghĩa như sau:

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \quad (2.2)$$

Với  $L_{xx}(x, \sigma) = \frac{\partial^2}{\partial x^2} g(\sigma)$  là tích của đạo hàm bậc hai hàm Gaussian với ảnh  $I$  tại điểm  $\mathbf{X}(x, y)$ , có tỷ lệ  $\sigma$ .

Nếu như SIFT xấp xỉ việc tính Laplacian của hàm Gaussian (LoG) bằng việc tính sai khác của hàm Gaussian (DoG) thì SURF xấp xỉ việc tính đạo hàm cấp 2 của hàm Gaussian bằng các hộp lọc (box filters). Dưới đây là một ví dụ của việc tính xấp xỉ đạo hàm cấp hai của hàm Gaussian với hệ số tỷ lệ thấp nhất bằng hộp lọc:



**Hình 2. 4 Xấp xỉ đạo hàm cấp 2 hàm Gaussian bằng hộp lọc**

Trong Hình 2.4 trên: Ảnh thứ nhất là đạo hàm ma trận đạo hàm cấp 2 Gaussian theo trục  $y$ , ảnh thứ hai theo trục  $x$  và trục  $y$ . Ảnh thứ ba và thứ tư lần lượt là các hộp lọc xấp xỉ với hai trường hợp của ảnh một và hai. Phép tích chập xoắn của ảnh  $I$  với các hộp lọc này được thực hiện rất nhanh bằng việc sử dụng kết hợp với ảnh tích lũy. Ta xác định vị trí và hệ số tỷ lệ tương ứng của điểm đặc trưng dựa trên định thức của ma trận Hessian. Công thức tính xấp xỉ định thức ma trận Hessian:

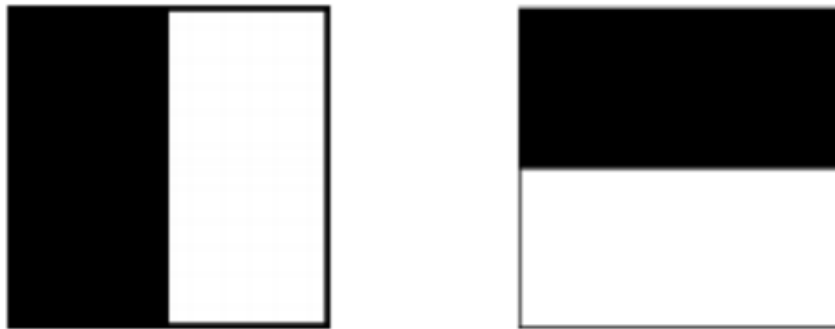
$$\det(H_{xấp xỉ}) = D_{xx}D_{yy} - (w \cdot D_{xy})^2 \quad (2.3)$$

Trong đó  $w$  là trọng số cân bằng của biểu thức định thức ma trận Hessian tùy thuộc vào hệ số tỷ lệ.  $D_{xx}, D_{xy}, D_{yy}$  là các hộp lọc xấp xỉ Gaussian như đã nói ở trên. Đối với tỷ lệ  $\sigma = 1.2$ , kích thước hộp lọc là  $9 \times 9$  thì giá trọng số được tính bằng:

$$w = \frac{|L_{xy}(1.2)|_F |D_{yy}(9)|_F}{|L_{yy}(1.2)|_F |D_{xy}(9)|_F} = 0.912 \approx 0.9 \quad (2.4)$$

Vị trí, tỷ lệ và không gian ảnh mà điểm đặc trưng được xác định một phép loại trừ phi cực đại trong một vùng  $3 \times 3 \times 3$  sẽ được áp dụng

Gán hướng cho các điểm nổi bật và mô tả đặc trưng SURF: Đầu tiên, ta phải xác định vùng hình xung quanh điểm đặc trưng vừa tìm được, gán một giá trị hướng duy nhất cho điểm đặc trưng. Kích thước của hình tròn phụ thuộc và hệ số tỷ lệ tương ứng trong không gian ảnh mà điểm đặc trưng tìm được. Ở đây các tác giả chọn bán kính của hình tròn là  $6s$ , trong đó  $s$  là tỷ lệ mà tại đó điểm đặc trưng được tìm thấy. Hướng của đặc trưng được tính bằng Haar wavelet tác động theo hai hướng  $x$  và  $y$  (Hình 3.5). Trong đó, vùng tối có trọng số  $-1$ , vùng sáng có trọng số  $+1$ . Kích thước của wavelet cũng phụ thuộc vào hệ số tỷ lệ  $s$ , ở đây các tác giả chọn  $4s$ .

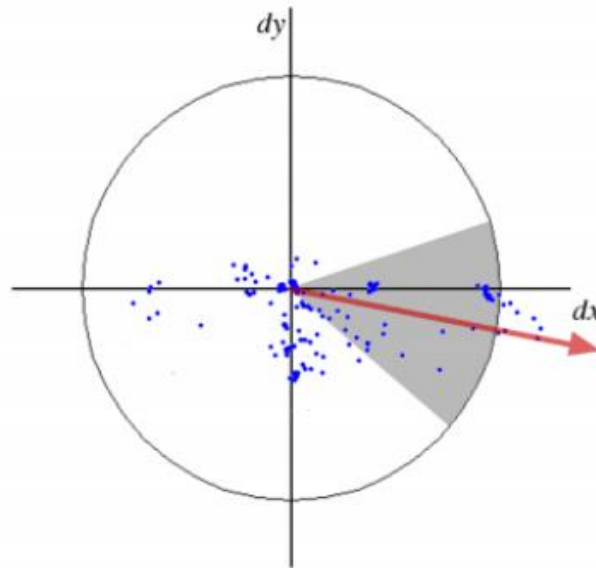


**Hình 2. 5 Lọc Haar wavelet để tính sự ảnh hưởng trên hai hướng  $x$  và  $y$**

Haar wavelet có thể được tính một cách nhanh chóng bằng cách sử dụng ảnh tích lũy tương tự như hộp lọc xấp xỉ của đạo hàm cấp 2 hàm Gaussian. Vector hướng nào trội nhất sẽ được ước lượng và gán vào thông tin của điểm đặc trưng. Hình dưới đây sẽ mô tả hướng và vùng ảnh hưởng của đặc trưng.

Tiếp theo, ta xây dựng các vùng hình vuông xung quanh điểm đặc trưng men theo vector hướng vừa ước lượng được ở bước trước đó. Vùng hình vuông này được chia nhỏ thành  $4 \times 4$  hình vuông con để ghi nhận thông tin của trên miền không gian ảnh lân cận.

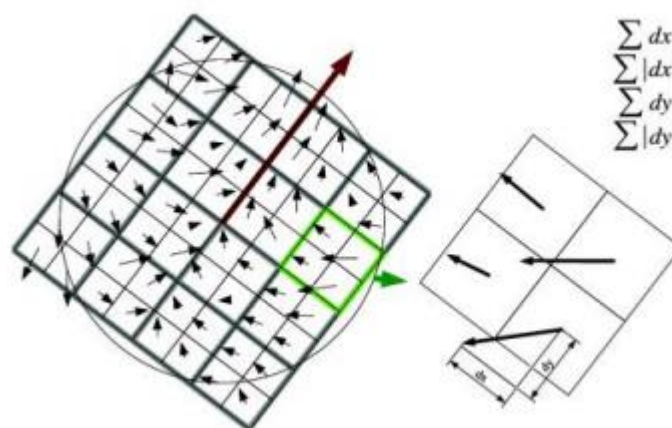
Haar wavelet được rút trích trên toàn bộ không gian điểm ảnh. Wavelet tác động trên hai hướng ngang và dọc được cộng dồn các giá trị  $d_x$  và  $d_y$  trên mỗi hình vuông con.



**Hình 2. 6 Vùng hình tròn xung quanh và hướng đại diện cho điểm đặc trưng**

Hơn thế nữa, các giá trị tuyệt đối  $|d_x|$  và  $|d_y|$  cũng được cộng dồn để lấy thông tin về độ lớn của sự thay đổi cường độ sáng trên ảnh. Như vậy mỗi hình vuông con sẽ được mô tả bởi một vector 4 chiều:

$$V = [\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|] \quad (2.5)$$



**Hình 2. 7 4x4 hình vuông con xung quanh điểm đặc trưng**

Như vậy vector mô tả cho tất cả 4x4 hình vuông con là một vector 64 chiều (4x4x4). Đây cũng chính là mô tả đặc trưng chuẩn của SURF (hay còn gọi là SURF-64). Ngoài



ra còn có các phiên bản khác dựa trên cách chia hình vuông con như SURF – 36, SURF – 128... Tuy nhiên thực nghiệm của các tác giả cho thấy rằng SURF – 64 cho tốc độ tính toán tốt nhất mà vẫn đảm bảo tính bền vững của đặc trưng. Haar wavelet bất biến với sự thay đổi của ánh sáng và sự tương phản khi ta chuẩn hóa vector mô tả đặc trưng về chiều dài đơn vị.

### 2.1.3 Túi từ trực quan BoW ( Bag Of Visual Words)

Mô hình túi từ (Bag Of Words) là một biểu diễn đơn giản hóa được sử dụng trong xử lý ngôn ngữ tự nhiên và truy vấn thông tin (IR). Trong mô hình này, một văn bản (chẳng hạn như một câu hoặc một tài liệu) được thể hiện dưới dạng túi (multiset) chứa các từ của nó, không quan tâm đến ngữ pháp và thậm chí trật tự từ nhưng vẫn giữ tính đa dạng. [18]

Mô hình túi từ thường được sử dụng trong các phương pháp phân loại tài liệu trong đó sự xuất hiện (tần suất) của mỗi từ được sử dụng như một đặc trưng để đào tạo bộ phân loại. Khi phân tích nội dung, các từ khóa (keywords) có trong đoạn văn bản sẽ được thu thập và đưa vào một tập hợp mà không cần quan tâm đến thứ tự xuất hiện của chúng. Những văn bản có nhiều từ khóa chung thì sẽ có mối liên hệ với nhau. Việc biểu diễn các văn bản dưới dạng một đồ thị histogram của tần suất xuất hiện của các từ khóa cho phép chúng ta phân tích và so sánh chúng một cách dễ dàng [19].

Ví dụ trong [18] là bài toán phân loại tin rác. Ta thấy rằng nếu một tin có chứa các từ khuyến mại, giảm giá, trúng thưởng, miễn phí, quà tặng, tri ân, ... thì nhiều khả năng đó là một tin nhắn rác. Vậy phương pháp đơn giản nhất là đếm xem trong tin đó có bao nhiêu từ thuộc vào các từ trên, nếu nhiều hơn 1 ngưỡng nào đó thì ta quyết định đó là tin rác. (Tất nhiên bài toán thực tế phức tạp hơn nhiều khi các từ có thể được viết dưới dạng không dấu, hoặc bị cố tình viết sai chính tả, hoặc dùng ngôn ngữ teen). Với các loại văn bản khác nhau thì lượng từ liên quan tới từng chủ đề cũng khác nhau. Từ đó có thể dựa vào số lượng các từ trong từng loại để làm các vector đặc trưng cho từng văn bản.

Sau đây là ví dụ về cách tạo ra vector đặc trưng cho mỗi văn bản dựa trên BoW. Giả sử ta có 2 văn bản đơn giản:

- (1) John likes to watch movies. Mary likes movies too.
- (2) John also likes to watch football games.

Dựa trên hai văn bản này, ta có danh sách các từ được sử dụng, được gọi là từ điển với 10 từ như sau:

- “John”, “likes”, “to”, “watch”, “movies”, “also”, “football”, “games”, “Mary”, “too”

Với mỗi văn bản, ta sẽ tạo ra một vector đặc trưng có số chiều bằng 10, mỗi phần tử đại diện cho số từ tương ứng xuất hiện trong văn bản đó. Với hai văn bản trên, ta sẽ có hai vector đặc trưng là:

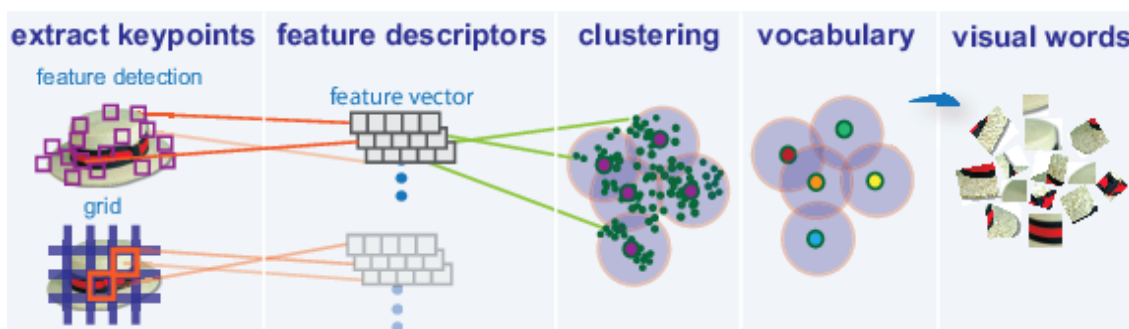
(1) [1, 2, 1, 1, 2, 0, 0, 0, 1, 1]

(2) [1, 1, 1, 1, 0, 1, 1, 1, 0, 0]

Trong thị giác máy tính, mô hình túi từ có thể được áp dụng để phân loại hình ảnh, bằng cách coi các đặc trưng của hình ảnh như các từ, khi đó ta sẽ gọi là mô hình túi từ trực quan (Bag Of Visual Words) và “một túi các từ trực quan” là một vector đếm tần suất xuất hiện của các đặc trưng của hình ảnh.

Tuy nhiên việc áp dụng BoW trong thị giác máy tính không đơn giản như việc phân tích một văn bản. Vấn đề ở đây là chúng ta không có một “tiêu chuẩn” để có thể đếm tần suất xuất hiện của các đặc trưng có trong bức ảnh, vì không có đặc trưng nào giống nhau 100% cả.

Để giải quyết vấn đề trên, chúng ta cần phải xây dựng một bộ “từ điển” (codebook) nhằm gom nhóm các vector đặc trưng vào các cụm (cluster) phù hợp, với các trung tâm của cụm được gọi là từ mã trực quan (visual vocabulary). Lúc này chúng ta có thể biểu diễn các bức ảnh dưới dạng đồ thị histogram tần suất xuất hiện của các từ mã.



**Hình 2. 8 Mô tả tạo túi từ đặc trưng BoW [12]**

Các bước để tạo túi từ trực quan được mô tả sau đây. Để thể hiện một hình ảnh bằng mô hình BoW, coi một ảnh là một văn bản. Tương tự, "từ" trong ảnh cũng cần được định

nghĩa. Để có được một mô hình BoW, thường bao gồm ba bước sau: trích chọn đặc trưng cho tập dữ liệu, xây dựng bộ từ điển nhằm tổng quát hóa tập vector đặc trưng vừa nhận được, biểu diễn các bức ảnh dưới dạng histogram tần xuất xuất hiện của các đặc trưng.

Sau khi phát hiện đặc trưng, mỗi ảnh được biểu diễn thành các đặc trưng dưới dạng các vector đặc trưng. Tiếp đó, chuyển đổi các vector thành từ khóa tương tự như các từ trong tài liệu văn bản; cũng tạo ra một từ điển. Một từ mã được coi là đại diện của một số vector đặc trưng. Để thực hiện việc này ta sử dụng phân cụm k-means trên tất cả các vector. Từ mã sẽ được định nghĩa là trung tâm của các cụm đã học. Số lượng các cụm là kích thước từ điển. Do đó hình ảnh có thể được biểu thị bằng đồ thị của từ mã, hay đồ thị đặc trưng. Tuy nhiên việc chọn số lượng cụm để gom nhóm cũng không hề đơn giản. Nếu số lượng cụm quá nhỏ thì sẽ không đủ để phân biệt nội dung giữa các bức ảnh. Ngược lại, nếu số lượng cụm quá lớn cũng ảnh hưởng xấu đến hiệu suất của hệ thống. Trên thực tế, để chọn được số lượng cụm phù hợp cần dựa vào kinh nghiệm và quá trình đánh giá hiệu năng.

Để biểu diễn đồ thị đặc trưng, sau khi có được đặc trưng của ảnh, tìm ra các từ mã tương đồng với nó nhất có trong từ điển bằng cách áp dụng khoảng cách Euclidean . Dựa vào số lượng từ mã tương đồng nhau đó, xây dựng các đồ thị tần suất của chúng, đồ thị này sẽ có kích thước bằng số lượng từ mã có trong từ điển.

Khi các bức ảnh đã được biểu diễn ở dạng đồ thị - cũng chính là một vector đặc trưng, chúng ta có thể sử dụng các thuật toán học máy (Machine Learning) để phân loại ảnh. Ở trong luận văn này, đặc trưng được đưa vào mô hình BoW là đặc trưng SURF.



**Hình 2. 9 Minh họa tạo đồ thị từ mã – vector đặc trưng [12]**

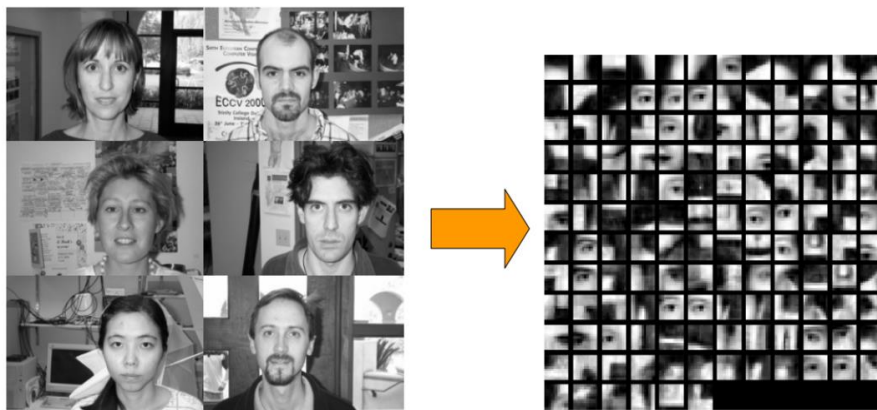
Để hiểu rõ hơn về mô hình BoW ta xét 2 ví dụ trong [20] sau:

- Ví dụ 1: Có hai class ảnh, một class là ảnh các khu rừng, một class là ảnh các sa mạc. Phân loại một bức ảnh là rừng hay sa mạc (giả sử ta biết rằng nó thuộc một trong hai loại này) một cách trực quan nhất là dựa vào màu sắc. Màu xanh nhiều

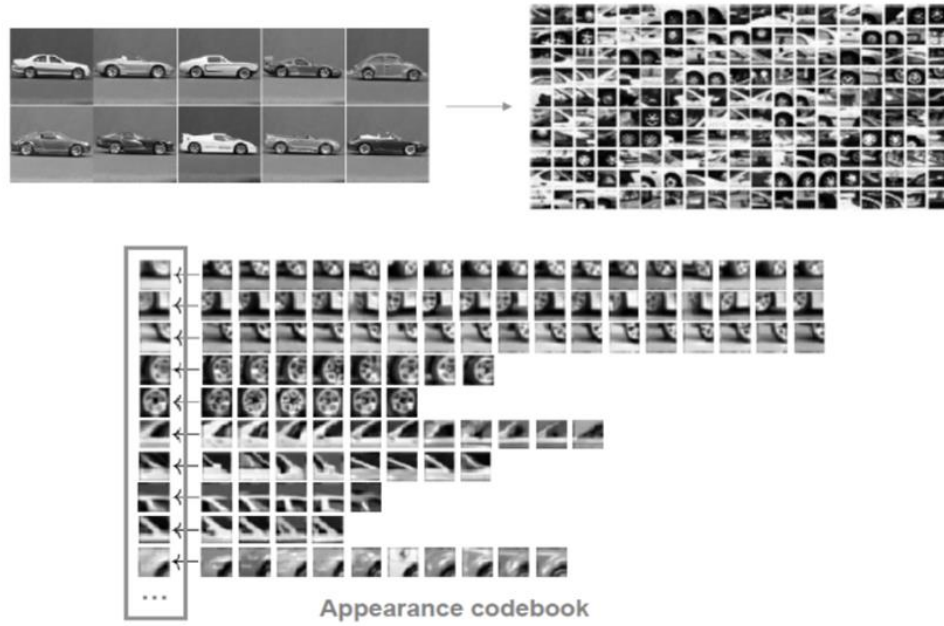
là rừng, màu đỏ và vàng nhiều là sa mạc. Vậy chúng ta có thể có một mô hình đơn giản để trích chọn đặc trưng như sau:

- Với một bức ảnh, chuẩn bị một vector  $X$  có số chiều bằng 3, đại diện cho 3 màu xanh ( $x_1$ ), đỏ ( $x_2$ ), và vàng ( $x_3$ ).
  - Với mỗi điểm ảnh trong bức ảnh đó, xem nó gần với màu xanh, đỏ hay vàng nhất dựa trên giá trị của pixel đó. Nếu nó gần điểm xanh nhất, tăng  $x_1$  lên 1; gần đỏ nhất, tăng  $x_2$  lên 1; gần vàng nhất, tăng  $x_3$  lên 1.
  - Sau khi xem xét tất cả các điểm ảnh, dù cho bức ảnh có kích thước thế nào, ta vẫn thu được một vector có độ dài bằng 3, mỗi phần tử thể hiện việc có bao nhiêu pixel trong bức ảnh có màu tương ứng. Vector cuối này còn được gọi là vector histogram của bức ảnh tương ứng với ba màu xanh, đỏ, vàng. Dựa vào vector này, ta có thể quyết định bức ảnh đó là ảnh rừng hay sa mạc.
- Ví dụ 2: Trên thực tế, các bài toán xử lý ảnh không đơn giản như ví dụ 1 trên đây. Mắt người thực ra nhạy với các đường nét, hình dáng hơn là màu sắc. Một cái (ảnh) cây dù không có màu vẫn là một cái (ảnh) cây! Vì vậy, xem xét giá trị từng điểm ảnh một không mang lại kết quả khả quan vì lượng thông tin bị mất quá nhiều. Có một cách khắc phục là thay vì xem xét một điểm ảnh, ta xem xét một cửa sổ nhỏ trong ảnh; trong thị giác máy tính, cửa sổ này được gọi là patch; là một hình chữ nhật chứa nhiều điểm ảnh gần nhau. Cửa sổ này đủ lớn để có thể chứa được các bộ phận có thể mô tả được vật thể trong ảnh.

Ví dụ với mặt người, các patch nên đủ lớn để chứa được các phần của khuôn mặt như mắt, mũi, miệng như Hình 2.10 ở trên. Và với ảnh là ô tô, các patch thu được có thể là bánh xe, khung xe, cửa xe,... như trong Hình 2.11.



**Hình 2. 10 Mô hình BoW chứa mặt người [20]**



Hình 2. 11 Mô hình BoW cho ảnh ô tô [20]

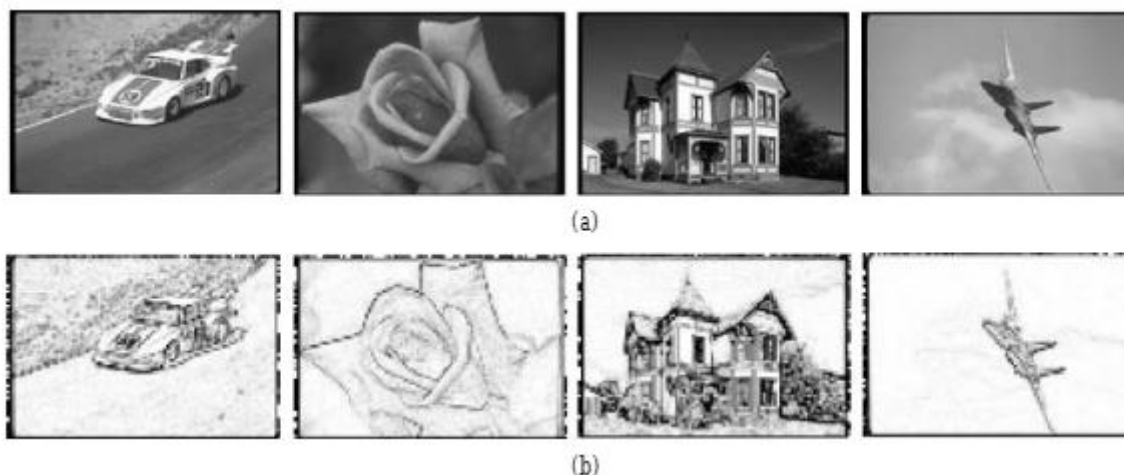
#### 2.1.4 Đặc trưng sai khác xác suất ngược khối BDIP (Block Difference Of Inverse Probabilities)

Toán tử sai khác xác suất ngược DIP (Difference of Inverse Probabilities) là toán tử trích xuất các đặc trưng của các vùng trung và đường viền tùy theo cường độ cục bộ [13]. Trong toán tử DIP, tỷ lệ cường độ pixel trong một cửa sổ ảnh với tổng tất cả các cường độ pixel trong cửa sổ đó là xác suất cần quan tâm. Vì thế, DIP nghĩa là sự khác biệt giữa nghịch đảo xác suất của pixel trung tâm trong một cửa sổ ảnh với pixel có cường độ lớn nhất trong cùng cửa sổ đó. BDIP là phiên bản dạng khối của DIP. Nó được định nghĩa là sự sai khác giữa số lượng pixel trong một khối và tỷ lệ của tổng các pixel trên pixel lớn nhất trong khối.

$$BDIP = M^2 - \frac{\sum_{(i,j) \in B} I(i,j)}{\max_{(i,j) \in B} I(i,j)} \quad (2.6)$$

trong đó  $I(i,j)$  là giá trị mức xám của pixel tọa độ  $(i,j)$  và  $B$  là khối có kích thước  $M \times M$ . Cường độ pixel trong khối biến đổi càng lớn thì giá trị BDIP sẽ càng lớn.

Trong Hình 2.10 là một vài ảnh gốc và ảnh BDIP tương ứng. Kích thước khối được chọn là  $2 \times 2$  và giá trị BDIP càng cao thì trên ảnh càng tối. Ở ảnh BDIP, bên trong vật thể và nền có độ sáng hơn, trong khi viền và các vùng trung tối màu hơn. Do đó có thể thấy hiệu quả của đặc trưng BDIP.



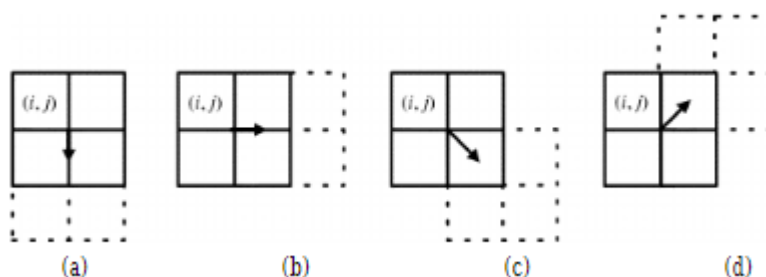
Hình 2. 12 Ảnh xám và ảnh BDIP tương ứng: (a) Ảnh gốc, (b) Ảnh BDIP [13]

### 2.1.5 Đặc trưng biến thiên hệ số tương quan cục bộ khối BVLC (Block Variation Of Local Correlation Coefficients)

Hệ số biến thiên tương quan cục bộ VLCC (Variation of Local Correlation Coefficients) được biết đến là một cách đo hiệu quả độ mịn của kết cấu. Nó được định nghĩa bởi sự sai khác hệ số tương quan cục bộ lớn nhất và nhỏ nhất theo 4 hướng. BVLC là phiên bản dạng khối của VLCC. Mỗi hệ số tương quan cục bộ được định nghĩa bởi:

$$\rho(k,l) = \frac{\frac{1}{M^2} \sum_{(i,j) \in B} I(i,j)I(i+k,j+l) - \mu_{0,0}\mu_{k,l}}{\sigma_{0,0}\sigma_{k,l}} \quad (2.7)$$

trong đó B là khối có kích thước M x M,  $\mu_{0,0}$  và  $\sigma_{0,0}$  là trung bình và độ lệch chuẩn cục bộ của khối B. Cặp giá trị (k,l) là độ dịch theo chiều ngang và dọc của khối theo 4 hướng (-90°, 0°, -45°, 45°).  $\mu_{k,l}$  và  $\sigma_{k,l}$  là trung bình và độ lệch chuẩn của khối sau khi bị dịch (k,l).



Hình 2. 13 Khối kích thước 2 x 2 (nét liền) và các khối sau khi dịch (nét đứt) theo 4 hướng: (a)  $\rho(0,1)$ , (b)  $\rho(1,0)$ , (c)  $\rho(1,1)$  và (d)  $\rho(1,-1)$  [13]



Hình 2.11 biểu diễn các khối 2 x 2 và các khối được dịch đi 4 hướng tương ứng để tính  $\rho(0,1)$ ,  $\rho(1,0)$ ,  $\rho(1,1)$  và  $\rho(1,-1)$ , Giá trị BVLC được định nghĩa bởi:

$$\text{BVLC} = \max_{(k,l) \in O_4} [\rho(k,l)] - \min_{(k,l) \in O_4} [\rho(k,l)]$$

$$O_4 = \{(0,1), (1,0), (1,1), (1,-1)\}. \quad (2.8)$$

Từ đó có thể thấy, độ biến thiên kết cấu của vùng cục bộ càng lớn thì giá trị BVLC sẽ càng lớn. Hình 3.5 là các ảnh BVLC cho các ảnh gốc ở Hình 3.3. Từ đó ta có thể thấy, cường độ của ảnh BVLC được xác định bởi độ mượt của kết cấu. Ví dụ, những kết cấu mượt như bầu trời sẽ có màu sáng, trong khi những kết cấu nhám hơn như con đường hoặc cánh hoa sẽ có màu tối hơn.



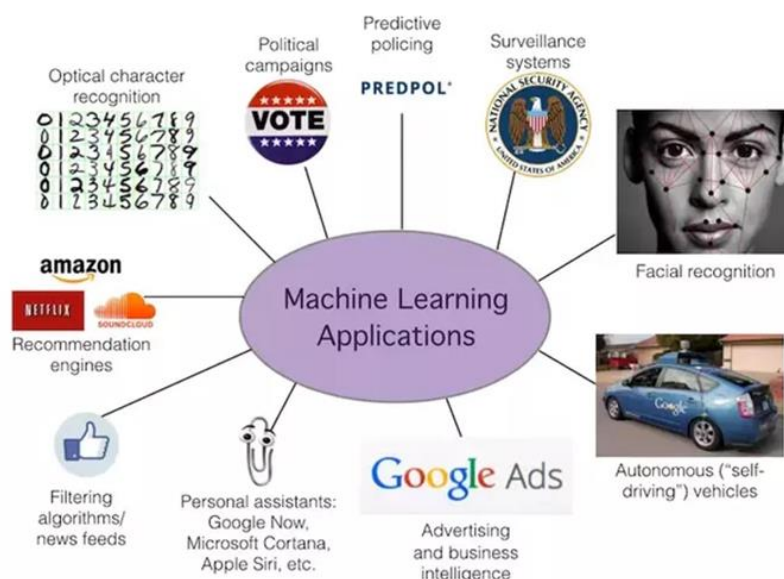
Hình 2. 14 Ảnh BVLC cho các ảnh gốc ở Hình 2.10 [13]

## 2.2 Giới thiệu về học máy (Machine Learning) và thuật toán máy vector hỗ trợ SVM

### 2.2.1 Giới thiệu về học máy (Machine Learning)

Học máy (Machine Learning) [21] là một lĩnh vực của trí tuệ nhân tạo liên quan đến việc nghiên cứu và xây dựng các kỹ thuật cho phép các hệ thống "học" tự động từ dữ liệu để giải quyết những vấn đề cụ thể. Ví dụ như các máy có thể "học" cách phân loại thư điện tử xem có phải thư rác (spam) hay không và tự động xếp thư vào thư mục tương ứng. Học máy rất gần với suy diễn thống kê (statistical inference) tuy có khác nhau về thuật ngữ.

Học máy có liên quan lớn đến thống kê, vì cả hai lĩnh vực đều nghiên cứu việc phân tích dữ liệu, nhưng khác với thống kê, học máy tập trung vào sự phức tạp của các giải thuật trong việc thực thi tính toán. Nhiều bài toán suy luận được xếp vào loại bài toán NP-khó (nondeterministic polynomial time), vì thế một phần của học máy là nghiên cứu sự phát triển các giải thuật suy luận xấp xỉ mà có thể xử lý được.



**Hình 2. 15 Các ứng dụng sử dụng Machine Learning [22]**

Học máy hiện nay được áp dụng rộng rãi bao gồm máy truy tìm dữ liệu, chẩn đoán y khoa, phát hiện thẻ tín dụng giả, phân tích thị trường chứng khoán, phân loại các chuỗi DNA, nhận dạng tiếng nói và chữ viết, dịch tự động, chơi trò chơi và cử động rô-bốt (robot locomotion).

### 2.2.2 Một số phương pháp học máy phổ biến

Hai phương pháp của học máy được chấp nhận rộng rãi chính là học có giám sát (Supervised Learning - SL) và học không giám sát (Unsupervised Learning - UL) nhưng cũng có những phương pháp khác như học bán giám sát (Semisupervised Learning), học tăng cường (reinforcement learning). Dưới đây là khái niệm chung về 2 phương pháp phổ biến nhất [23]:

- Học có giám sát (SL): Là một kỹ thuật học máy để học tập từ tập dữ liệu được gán nhãn cho trước. Tập dữ liệu cho trước sẽ chứa nhiều bộ dữ liệu. Mỗi bộ dữ liệu có cấu trúc theo cặp  $\{x, y\}$  với  $x$  được xem là dữ liệu thô (raw data) và  $y$  là nhãn của dữ liệu đó. Nhiệm vụ của SL là dự đoán đầu ra mong muốn dựa vào giá trị đầu vào. Để nhận ra, học có “*giám sát*” tức là máy học dựa vào sự trợ giúp của con người, hay nói cách khác con người dạy cho máy học và giá trị đầu ra mong muốn được định trước bởi con người. Tập dữ liệu huấn luyện hoàn toàn được gán nhãn dựa vào con người. Tập càng nhỏ thì máy tính học càng ít. SL cũng được áp dụng cho 2 nhóm bài toán chính là bài toán dự đoán (Regression



Problem) và bài toán phân lớp (Classification Problem). Kỹ thuật SL thực chất là để xây dựng một hàm có thể xuất ra giá trị đầu ra tương ứng với tập dữ liệu. Ta gọi hàm này là hàm  $h(x)$  và mong muốn hàm này xuất ra đúng giá trị  $y$  với một hoặc nhiều tập dữ liệu mới khác với dữ liệu được học. Hàm  $h(x)$  cần các loại tham số học khác nhau tùy thuộc với nhiều bài toán khác nhau. Việc học từ tập dữ liệu (training) cũng chính là tìm ra bộ tham số học cho hàm  $h(x)$ .

- Học không giám sát (UL): Là một kỹ thuật của máy học nhằm tìm ra một mô hình hay cấu trúc bị ẩn bởi tập dữ liệu “không” được gán nhãn cho trước. UL khác với SL là không thể xác định trước đầu ra từ tập dữ liệu huấn luyện được. Tùy thuộc vào tập huấn luyện kết quả đầu ra sẽ khác nhau. Trái ngược với SL, tập dữ liệu huấn luyện của UL không do con người gán nhãn, máy tính sẽ phải tự học hoàn toàn. Có thể nói, học “không giám sát” thì giá trị đầu ra sẽ phụ thuộc vào thuật toán UL.

Ứng dụng: Ứng dụng phổ biến nhất của học không giám sát là gom cụm. Ứng dụng này dễ nhận ra nhất là Google và Facebook. Google có thể gom nhóm các bài báo có nội dung gần nhau, hoặc Facebook có thể gợi ý kết bạn có nhiều bạn chung cho bạn. Các bài báo có cùng nội dung sẽ được gom lại thành một nhóm phân biệt với các nhóm khác. Dữ liệu huấn luyện là các bài báo từ quá khứ tới hiện tại và tăng dần theo thời gian. Dễ nhận ra rằng dữ liệu không thể gán nhãn bởi con người. Khi một bài báo mới được cho vào đầu vào, nó sẽ tìm cụm gần nhất với bài báo đó và gợi ý những bài liên quan.

## **2.2.3 Thuật toán máy vector hỗ trợ SVM phân lớp dữ liệu**

### **2.2.3.1 Giới thiệu về thuật toán máy vector hỗ trợ**

Máy vector hỗ trợ (Support Vector Machine - SVM) là một khái niệm trong thống kê và khoa học máy tính cho một tập hợp các phương pháp học có giám sát liên quan đến nhau để phân loại và phân tích hồi quy. SVM dạng chuẩn nhận dữ liệu vào và phân loại chúng vào hai lớp khác nhau. Do đó SVM là một thuật toán phân loại nhị phân. Với một bộ các ví dụ luyện tập thuộc hai thể loại cho trước, thuật toán luyện tập SVM xây dựng một mô hình SVM để phân loại các ví dụ khác vào hai thể loại đó. Một mô hình SVM là một cách biểu diễn các điểm trong không gian và lựa chọn ranh giới giữa hai thể loại sao cho khoảng cách từ các ví dụ luyện tập tới ranh giới là xa nhất có thể. Các

ví dụ mới cũng được biểu diễn trong cùng một không gian và được thuật toán dự đoán thuộc một trong hai thể loại tùy vào ví dụ đó nằm ở phía nào của ranh giới. [24]

Một máy vector hỗ trợ xây dựng một siêu phẳng hoặc một tập hợp các siêu phẳng trong một không gian nhiều chiều hoặc vô hạn chiều, có thể được sử dụng cho phân loại, hồi quy, hoặc các nhiệm vụ khác. Một cách trực giác, để phân loại tốt nhất thì các siêu phẳng nằm ở càng xa các điểm dữ liệu của tất cả các lớp (gọi là hàm lề) càng tốt, vì nói chung lề càng lớn thì sai số tổng quát hóa của thuật toán phân loại càng bé.

Trong nhiều trường hợp, không thể phân chia các lớp dữ liệu một cách tuyến tính trong một không gian ban đầu được dùng để mô tả một vấn đề. Vì vậy, nhiều khi cần phải ánh xạ các điểm dữ liệu trong không gian ban đầu vào một không gian mới nhiều chiều hơn, để việc phân tách chúng trở nên dễ dàng hơn trong không gian mới. Để việc tính toán được hiệu quả, ánh xạ sử dụng trong thuật toán SVM chỉ đòi hỏi tích vô hướng của các vector dữ liệu trong không gian mới có thể được tính dễ dàng từ các tọa độ trong không gian cũ. Tích vô hướng này được xác định bằng một hàm hạt nhân  $K(x, y)$  phù hợp. Một siêu phẳng trong không gian mới được định nghĩa là tập hợp các điểm có tích vô hướng với một vector cố định trong không gian đó là một hằng số. Vector xác định một siêu phẳng sử dụng trong SVM là một tổ hợp tuyến tính của các vector dữ liệu luyện tập trong không gian mới với các hệ số  $\alpha_i$ . Với siêu phẳng lựa chọn như trên, các điểm  $x$  trong không gian đặc trưng được ánh xạ vào một siêu mặt phẳng là các điểm thỏa mãn:

$$\sum_i \alpha_i K(X_i, X) = \text{hằng số} \quad (2.9)$$

Chú ý rằng nếu  $K(x, y)$  nhận giá trị ngày càng nhỏ khi  $y$  xa dần khỏi  $x$  thì mỗi số hạng của tổng trên được dùng để đo độ tương tự giữa  $x$  với điểm  $x_i$  tương ứng trong dữ liệu luyện tập. Như vậy, tác dụng của tổng trên chính là so sánh khoảng cách giữa điểm cần dự đoán với các điểm dữ liệu đã biết. Lưu ý là tập hợp các điểm  $x$  được ánh xạ vào một siêu phẳng có thể có độ phức tạp tùy ý trong không gian ban đầu, nên có thể phân tách các tập hợp thậm chí không lồi trong không gian ban đầu.

SVM thực chất là một bài toán tối ưu, mục tiêu của thuật toán này là tìm được một không gian  $F$  và siêu phẳng quyết định  $f$  trên  $F$  sao cho sai số phân loại là thấp nhất.

Cho tập mẫu  $D = \{(x_i, y_i) | x_i \in \mathbf{R}^n, y_i \in \{-1, 1\}\}_{i=1}^n$  (-1 biểu thị lớp I, 1 biểu thị lớp II). Ta có phương trình siêu phẳng chứa vector  $\vec{x}_i$  trong không gian:

$$\vec{x}_i \vec{w} + b = 0 \quad (2.10)$$

$$\text{Đặt } f(\vec{x}_i) = \text{sign}(\vec{x}_i \vec{w} + b) = \begin{cases} +1, & \vec{x}_i \vec{w} + b > 0 \\ -1, & \vec{x}_i \vec{w} + b < 0 \end{cases}$$

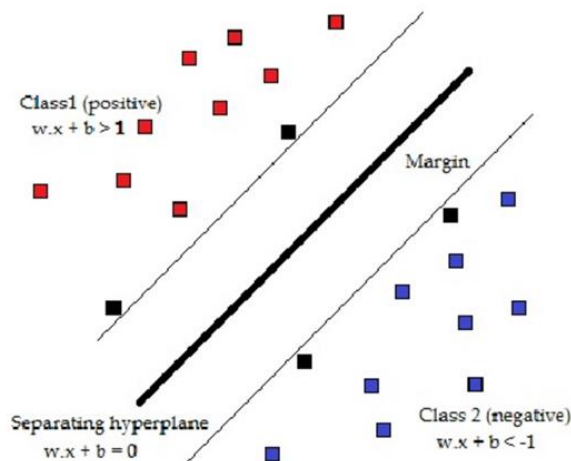
Như vậy,  $f(\vec{x}_i)$  biểu diễn sự phân lớp của  $\vec{x}_i$  vào 2 lớp như đã nêu. Ta nói  $y_i = +1$  nếu  $\vec{x}_i$  thuộc lớp I và  $y_i = -1$  nếu  $\vec{x}_i$  thuộc lớp II.

### 2.2.3.2 Bài toán phân 2 lớp với SVM

Bài toán đặt ra là: Xác định hàm phân lớp để phân lớp các mẫu trong tương lai, nghĩa là với một mẫu dữ liệu mới  $x_i$  thì cần phải xác định  $x_i$  được phân vào lớp +1 hay -1. [25]

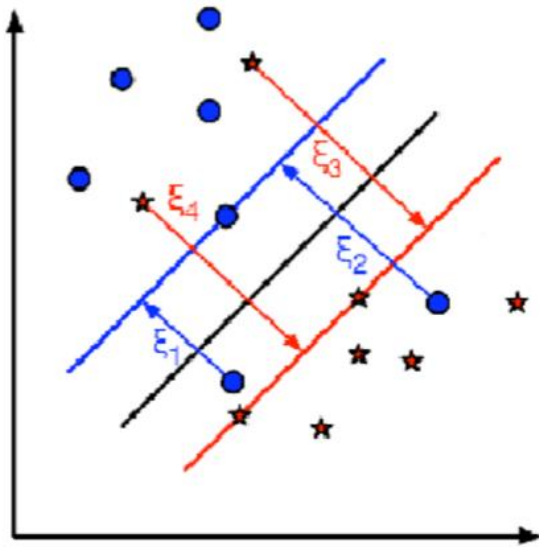
Ta xét 3 trường hợp, mỗi trường hợp sẽ có 1 bài toán tối ưu, giải được bài toán tối ưu đó ta sẽ tìm được siêu phẳng cần tìm.

- Trường hợp 1: Tập D có thể phân chia tuyến tính được mà không có nhiễu (tất cả các điểm được gán nhãn +1 thuộc về phía dương của siêu phẳng, tất cả các điểm được gán nhãn -1 thuộc về phía âm của siêu phẳng). Xem Hình 2.16.



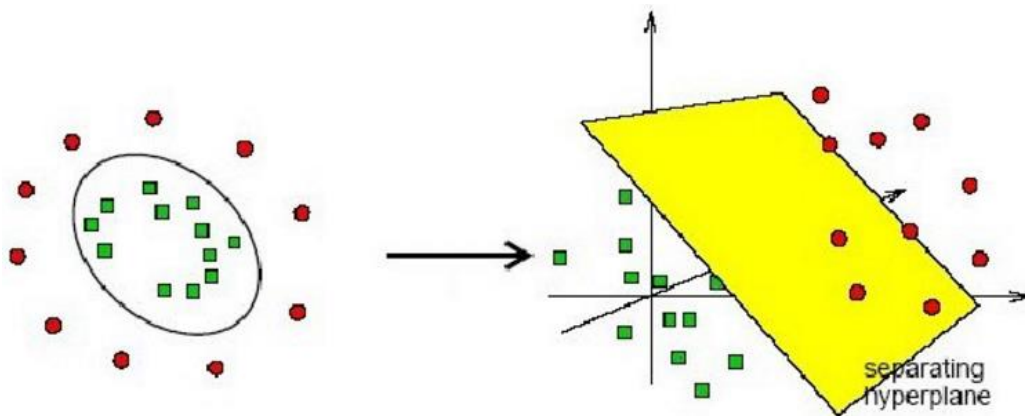
**Hình 2. 16 Tập dữ liệu được phân chia tuyến tính [25]**

- Trường hợp 2: Tập dữ liệu có thể phân chia tuyến tính được nhưng có nhiễu. Trong trường hợp này, hầu hết các điểm đều được phân chia đúng bởi siêu phẳng. Tuy nhiên có một số điểm bị nhiễu, nghĩa là: Điểm có nhãn dương nhưng lại thuộc phía âm của siêu phẳng, và ngược lại, điểm có nhãn âm nhưng lại thuộc phía dương của siêu phẳng. Xem Hình 2.17.



**Hình 2. 17 Tập dữ liệu phân chia tuyến tính nhưng có nhiễu [25]**

- Trường hợp 3: Tập dữ liệu D không thể phân chia tuyến tính được, ta sẽ ánh xạ các vector dữ liệu từ không gian n chiều vào một không gian m chiều ( $m > n$ ), sao cho trong không gian m chiều, D có thể phân chia tuyến tính được.



**Hình 2. 18 Tập dữ liệu không thể phân chia tuyến tính [25]**

### 2.2.3.3 Bài toán phân đa lớp với SVM

Để phân đa lớp thì kỹ thuật SVM sẽ chia không gian dữ liệu thành 2 phần và tiếp tục với không gian đã được phân chia. Khi đó hàm quyết định phân dữ liệu vào lớp thứ  $i$  sẽ là:

$$f_i(x) = w_i^T x + b_i \quad (2.11)$$

Những phần tử  $x$  là vector hỗ trợ nếu thỏa mãn điều kiện:

$$f_i(x) = \begin{cases} 1, & \epsilon i \\ -1, & \notin i \end{cases} \quad (2.12)$$

Giả sử bài toán phân loại  $k$  lớp ( $k > 2$ ), ta sẽ tiến hành  $k(k - 1)/2$  lần phân lớp nhị phân sử dụng phương pháp SVM. Mỗi lớp sẽ tiến hành phân tách với  $k-1$  lớp còn lại để xác định  $k-1$  hàm phân tách.

Kỹ thuật phân đa lớp bằng phương pháp SVM hiện vẫn đang được tiếp tục nghiên cứu và phát triển.

#### *2.2.3.4 Các bước chính của phương pháp phân loại bằng SVM*

- Tiền xử lý dữ liệu: Phương pháp SVM yêu cầu dữ liệu được diễn tả như các vector của các số thực. Như vậy nếu đầu vào chưa phải là số thực thì ta cần phải tìm cách chuyển chúng về dạng số của SVM. Tránh các số quá lớn, thường nên co giãn dữ liệu để chuyển về đoạn  $[-1, 1]$  hoặc  $[0, 1]$ .
- Chọn hàm hạt nhân: Cần chọn hàm hạt nhân phù hợp tương ứng cho từng bài toán cụ thể để đạt được độ chính xác cao trong quá trình phân lớp.
- Thực hiện việc kiểm tra chéo để xác định các tham số cho ứng dụng.
- Sử dụng các tham số cho việc huấn luyện tập mẫu.
- Kiểm thử tập dữ liệu test.

#### **2.2.4 Một số thuật toán phân loại khác được sử dụng trong thử nghiệm**

##### *2.2.4.1 Thuật toán K phần tử lân cận gần nhất (K nearest neighbors – KNN)*

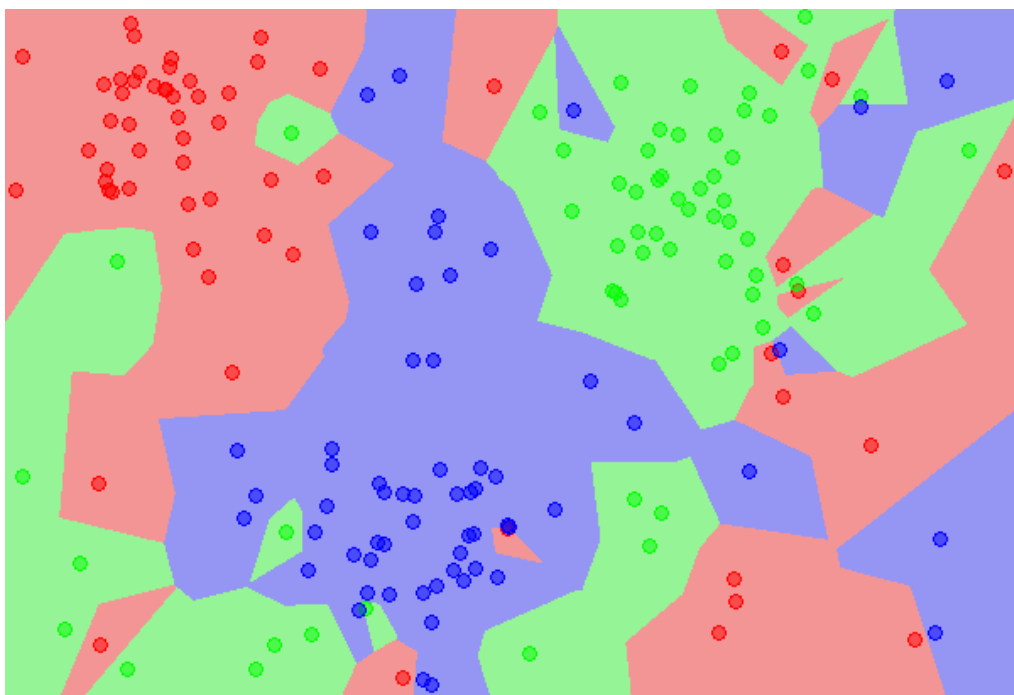
Trong bài toán nhận dạng, thuật toán K phần tử lân cận gần nhất (K nearest neighbors – KNN) là một phương pháp không tham số được sử dụng để phân loại và hồi quy [26]. Trong cả hai trường hợp, đầu vào bao gồm  $k$  mẫu huấn luyện gần nhất trong không gian đặc trưng. Đầu ra phụ thuộc vào việc KNN được sử dụng để phân loại hay hồi quy.

Với KNN, trong bài toán phân loại, nhãn của một điểm dữ liệu mới được suy ra trực tiếp từ  $K$  điểm dữ liệu gần nhất trong tập dữ liệu huấn luyện. Nhãn của một dữ liệu kiểm định có thể được quyết định bằng bầu chọn theo số phiếu (major voting) giữa các điểm gần nhất, hoặc nó có thể được suy ra bằng cách đánh trọng số khác nhau cho mỗi trong các điểm gần nhất đó rồi suy ra nhãn [27].

Trong bài toán hồi quy, đầu ra của một điểm dữ liệu sẽ bằng chính đầu ra của điểm dữ liệu đã biết gần nhất (trong trường hợp  $K=1$ ), hoặc là trung bình có trọng số của đầu ra của những điểm gần nhất, hoặc bằng một mối quan hệ dựa trên khoảng cách tới các điểm gần nhất đó.

Một cách ngắn gọn, KNN là thuật toán đi tìm đầu ra của một điểm dữ liệu mới bằng cách chỉ dựa trên thông tin của K điểm dữ liệu trong training set gần nó nhất (K-lân cận), không quan tâm đến việc có một vài điểm dữ liệu trong những điểm gần nhất này là nhiều. Thuật toán KNN là một trong những thuật toán đơn giản nhất trong tất cả các thuật toán học máy. Hình 2.19 là một ví dụ về KNN trong classification với  $K = 1$ .

Ví dụ trong Hình 2.19 là bài toán phân loại với 3 lớp: Đỏ, Lam, Lục. Mỗi điểm dữ liệu mới (test data point) sẽ được gán nhãn theo màu của điểm mà nó thuộc về. Trong hình này, có một vài vùng nhỏ xem lẫn vào các vùng lớn hơn khác màu. Ví dụ có một điểm màu Lục ở gần góc 11 giờ nằm giữa hai vùng lớn với nhiều dữ liệu màu Đỏ và Lam. Điểm này rất có thể là nhiễu. Dẫn đến nếu dữ liệu test rơi vào vùng này sẽ có nhiều khả năng cho kết quả không chính xác.



**Hình 2. 19 Bản đồ phân loại của 1NN [26]**

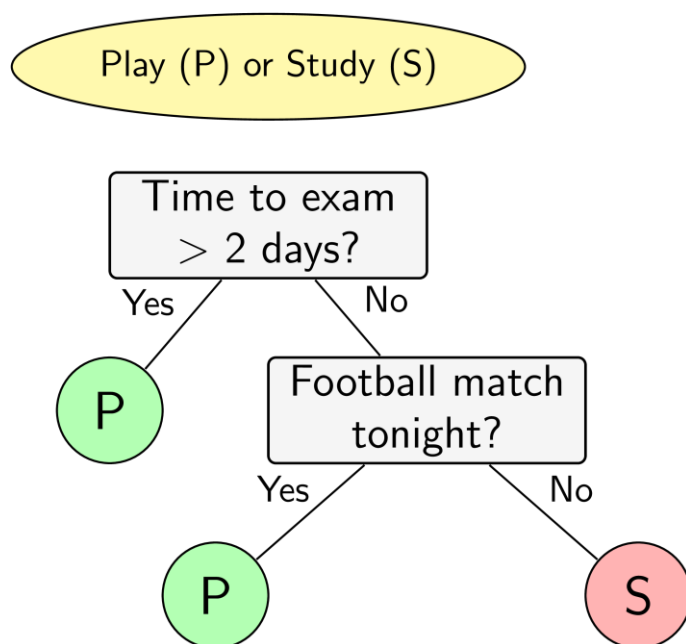
#### 2.2.4.2 Thuật toán cây quyết định hay cây phân loại (Decision or Classification Tree-CT)

Trong lĩnh vực máy học, cây quyết định, hay còn gọi là cây phân loại (trong mô hình phân loại) là một kiểu mô hình dự đoán (predictive model), nghĩa là một ánh xạ từ các quan sát về một sự vật/hiện tượng tới các kết luận về giá trị mục tiêu của sự vật/hiện tượng. Mỗi một nút trong (internal node) tương ứng với một biến; đường nối giữa nó với nút con của nó thể hiện một giá trị cụ thể cho biến đó. Mỗi nút lá đại diện cho giá trị

dự đoán của biến mục tiêu, cho trước các giá trị của các biến được biểu diễn bởi đường đi từ nút gốc tới nút lá đó. Kỹ thuật học máy dùng trong cây quyết định được gọi là học bằng cây quyết định, hay chỉ gọi với cái tên ngắn gọn là cây quyết định.

Học bằng cây quyết định cũng là một phương pháp thông dụng trong khai phá dữ liệu. Khi đó, cây quyết định mô tả một cấu trúc cây, trong đó, các lá đại diện cho các phân loại còn cành đại diện cho các kết hợp của các thuộc tính dẫn tới phân loại đó. [28] Một cây quyết định có thể được học bằng cách chia tập hợp nguồn thành các tập con dựa theo một kiểm tra giá trị thuộc tính. Quá trình này được lặp lại một cách đệ quy cho mỗi tập con dẫn xuất. Quá trình đệ quy hoàn thành khi không thể tiếp tục thực hiện việc chia tách được nữa, hay khi một phân loại đơn có thể áp dụng cho từng phần tử của tập con dẫn xuất. Một bộ phân loại rừng ngẫu nhiên (random forest) sử dụng một số cây quyết định để có thể cải thiện tỷ lệ phân loại.

Cây quyết định cũng là một phương tiện có tính mô tả dành cho việc tính toán các xác suất có điều kiện.



**Hình 2. 20 Ví dụ về một mô hình cây quyết định [29]**

Một ví dụ về cây quyết định trong [29]: Sắp đến kỳ thi, một cậu sinh viên tự đặt ra quy tắc học hay chơi của mình như sau. Nếu còn nhiều hơn hai ngày tới ngày thi, cậu ra sẽ đi chơi. Nếu còn không quá hai ngày và đêm hôm đó có một trận bóng đá, cậu sẽ sang nhà bạn chơi và cùng xem bóng đêm đó. Cậu sẽ chỉ học trong các trường hợp còn lại.

Việc ra quyết định của cậu sinh viên này có thể được mô tả trên sơ đồ trong Hình 2.20. Hình ellipse nền vàng thể hiện quyết định cần được đưa ra. Quyết định này phụ thuộc vào các câu trả lời của các câu hỏi trong các ô hình chữ nhật màu xám. Dựa trên các câu trả lời, quyết định cuối cùng được cho trong các hình tròn màu lục (chơi) và đỏ (học). Sơ đồ trong Hình 2.20 còn được gọi là một cây quyết định.

### **Tổng kết chương**

Chương 2 đã trình bày chi tiết về các phương pháp trích chọn đặc trưng ảnh, mô hình bộ phân loại SVM và một số thuật toán phân loại khác trong học máy. Trong chương 3 sẽ đi sâu vào đề xuất thử nghiệm bài toán phân loại ảnh nội soi dạ dày được trích chọn đặc trưng từ các phương pháp đó.

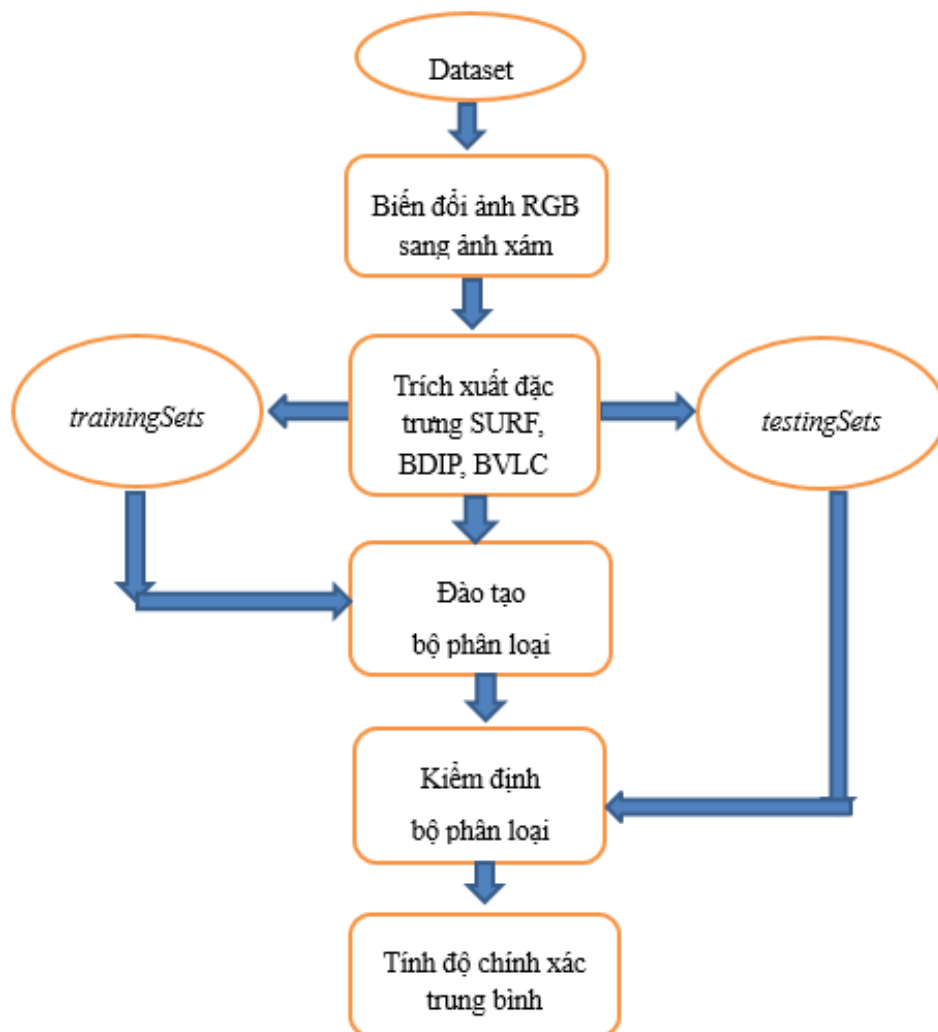


## CHƯƠNG 3. ĐỀ XUẤT PHƯƠNG PHÁP PHÂN LOẠI ẢNH

### NỘI SOI DẠ DÀY

Trong chương 3 này trình bày về đề xuất sử dụng các đặc trưng khác nhau để trích xuất cho ảnh nội soi dạ dày, đưa ra sơ đồ khối mô tả quá trình thực hiện giải quyết bài toán phân loại ảnh nội soi dạ dày.

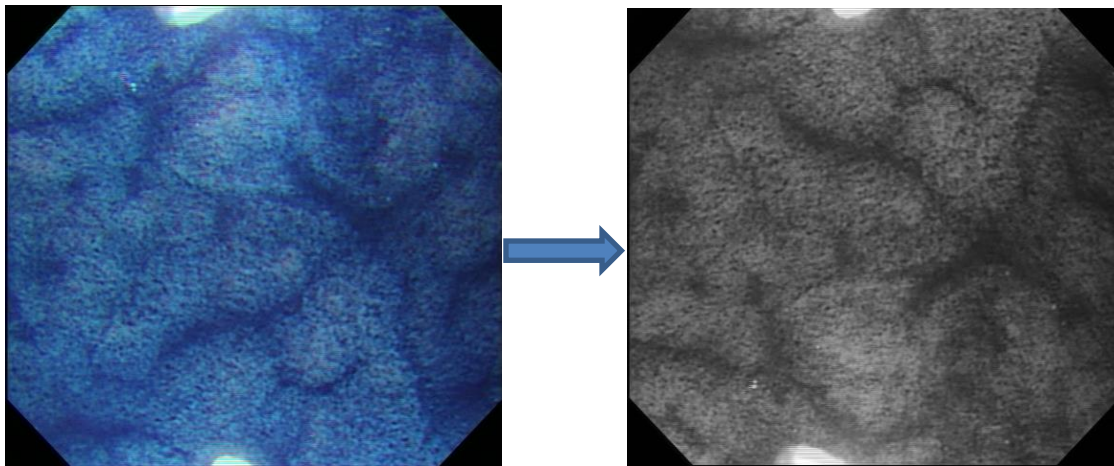
#### 3.1 Đề xuất các phương pháp trích chọn đặc trưng SURF, BDIP, BVLC



Hình 3. 1 Sơ đồ khối thực hiện bài toán phân loại ảnh nội soi dạ dày

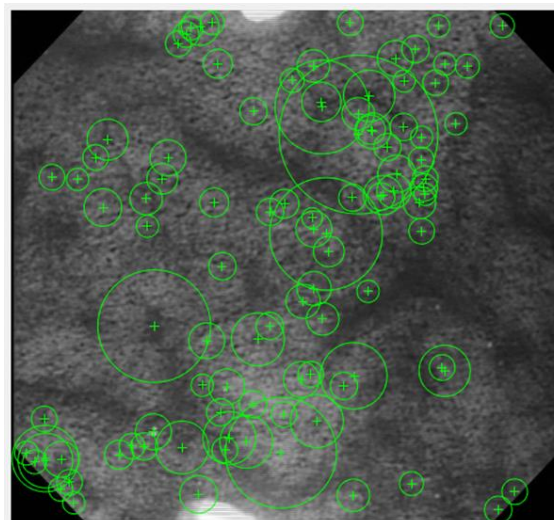
Theo Hình 3.1, cấu trúc thực hiện phương pháp trích chọn đặc trưng SURF, BDIP, BVLC cho ảnh để phân loại bao gồm các bước chính sau:

- Bước 1: Biến đổi ảnh RGB thành ảnh xám.



**Hình 3.2 Một ảnh trong bộ dữ liệu được chuyển thành ảnh xám**

- Bước 2: Trích xuất đặc trưng SURF, hoặc BDIP hay BVLC. Hình 3.3 mô tả một ảnh trong bộ dữ liệu đã được trích chọn ra 100 đặc trưng SURF mạnh nhất



**Hình 3.3 Ảnh xám trong Hình 3.2 được lấy 100 vector đặc trưng SURF**

- Bước 3 : Đào tạo và kiểm định bộ phân loại, tính độ chính xác.

## **3.2 Chi tiết các bước chính trong mô hình**

### **3.2.1 Biến đổi ảnh RGB thành ảnh xám**

Yêu cầu để trích xuất các đặc trưng SURF, BDIP, BVLC của ảnh là ảnh phải là ảnh xám. Ảnh xám hay còn gọi là ảnh đơn sắc (monochromatic), là ảnh mà tại mỗi điểm ảnh có một giá trị mức xám. Ảnh 8 mức xám sẽ có giá trị mỗi điểm ảnh nằm trong đoạn  $[0, 7]$ . Ảnh 256 mức xám sẽ có giá trị mỗi điểm ảnh nằm trong đoạn  $[0, 255]$ .



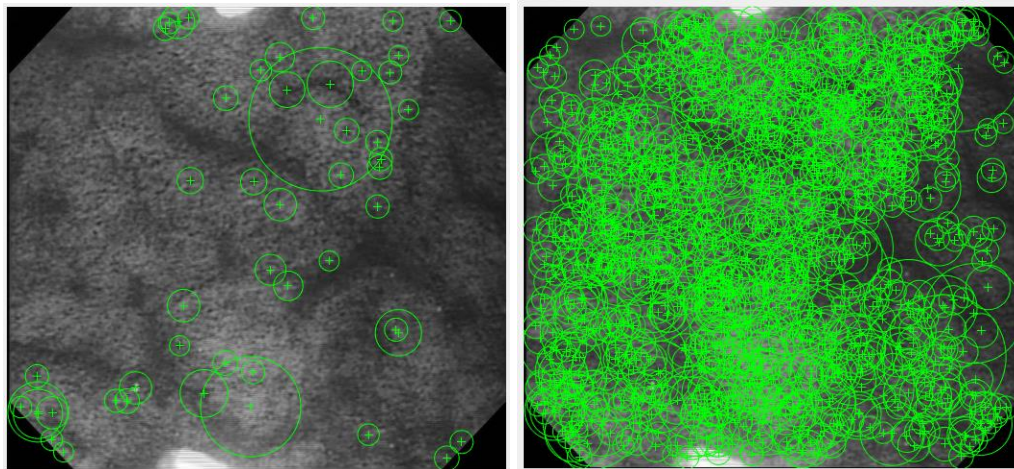
**Hình 3. 4 Thang màu xám của ảnh**

Giá trị điểm ảnh = 0 nghĩa là điểm ảnh đó tối (đen), giá trị điểm ảnh lớn nhất nghĩa là điểm ảnh đó trắng. Nói cách khác, giá trị mỗi điểm ảnh càng lớn thì điểm ảnh đó càng sáng. Cường độ sáng được tính theo công thức (chuyển đổi từ RGB sang ảnh xám): độ sáng =  $0.2989R + 0.5870G + 0.1140B$ . Các hệ số có thể được làm tròn thành 0.3, 0.59 và 0.11.

Vì vậy hàm `rgb2gray` trong Matlab được sử dụng để biến đổi ảnh của bộ dữ liệu từ ảnh 3 kênh màu RGB thành ảnh xám mà không thay đổi kích thước ban đầu của ảnh là 518x481 pixel.

### **3.2.2 Trích chọn đặc trưng SURF, BDIP và BVLC**

#### **3.2.2.1 Trích chọn đặc trưng SURF**



**a)**

**b)**

**Hình 3. 5 Ảnh xám trong Hình 3.2 được lấy 50 đặc trưng SURF mạnh nhất (a) và 1000 đặc trưng SURF mạnh nhất (b)**

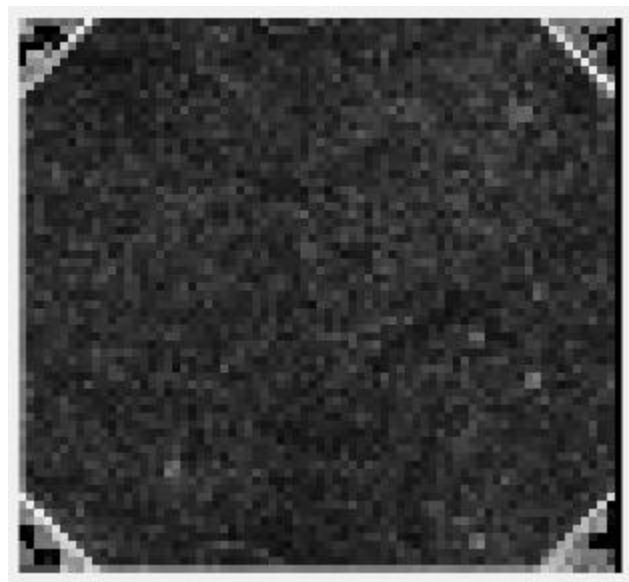
Ảnh có kích thước là 518x481 pixel vì vậy khi trích xuất đặc trưng SURF thì vector đặc trưng có kích thước khá lớn khoảng 2000x64, việc lấy vector đặc trưng lớn như vậy sẽ làm ảnh hưởng đến thời gian của quá trình phân loại, đồng thời khi phân loại sẽ không cho độ chính xác cao. Vì vậy, trong quá trình thử nghiệm và lựa chọn, các tác giả quyết định sẽ lựa chọn một số lượng nhất định các đặc trưng mạnh nhất trong ảnh để tạo thành vector đặc trưng đưa vào bộ phân loại. Thử nghiệm lần lượt với lựa chọn số lượng đặc

trung mạnh nhất là 50, 100 và 1000 đặc trưng SURF mạnh nhất của ảnh, với kích thước mỗi đặc trưng là 64. Hình 3.5a và 3.5b mô tả lần lượt cùng một ảnh trong bộ dữ liệu đã được trích chọn ra 50 và 1000 đặc trưng SURF mạnh nhất.

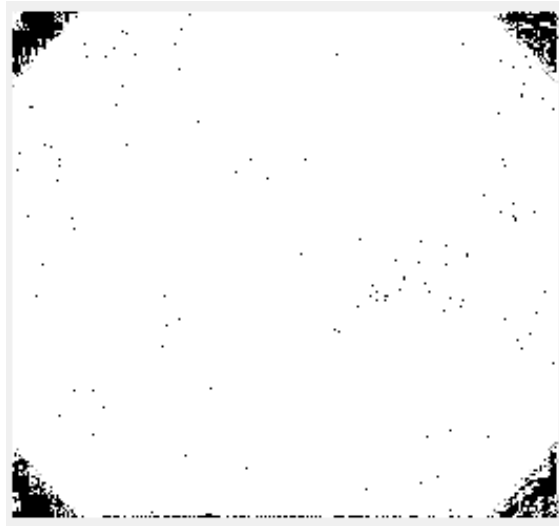
#### 3.2.2.2 Trích chọn đặc trưng BDIP và BVLC

Ta biết ảnh đầu vào có kích thước là 518x481, vì vậy câu hỏi đặt ra là kích thước khối cần thiết là bao nhiêu để đạt hiệu quả phân loại, có cần chồng lấp nhau hay không (overlapping/non-overlapping). Với mỗi loại đặc trưng, ta chia ảnh đầu vào thành các khối có kích thước bằng nhau, tính toán giá trị đặc trưng trên mỗi khối đó và lưu kết quả. Chi tiết cách làm như sau:

- Sử dụng hàm `imread` trong Matlab để đọc ảnh xám đã có được ở trên.
- Xác định các tham số: `block_size` – là kích thước khối và số pixel chồng lấp lẫn nhau giữa các khối.
- Tính toán số lượng đặc trưng tối đa trên một ảnh để tạo vector lưu trữ (do số lượng đặc trưng sẽ thay đổi theo hai tham số ở trên).
- Tính giá trị BDIP và BVLC trên mỗi khối theo công thức đã có (với đặc trưng BVLC, do tính chất dịch theo 4 hướng nên đối với những khối nằm ở biên của ảnh, cần phải tạo thêm các giá trị để đảm bảo tồn tại giá trị tính toán).
- Tiến hành lưu kết quả tính toán đặc trưng dưới dạng ma trận (file “.mat”).



**Hình 3. 6 Ảnh xám trong Hình 3.2 được lấy đặc trưng BDIP**



**Hình 3. 7** Ảnh xám trong Hình 3.2 được lấy đặc trưng BVLC

### ***3.2.3 Đào tạo và kiểm định bộ phân loại, tính độ chính xác của mô hình phân loại***

Từ vector đặc trưng SURF hoặc BDIP hay BVLC ta đã lấy ở trên và dataset được chia làm 2 tập là trainingsets và testingsets đào tạo ra một bộ phân loại SVM của trainingsets, bằng cách sử dụng hàm `fitcsvm` trong Matlab. Hàm này trả về một bộ phân loại SVM được gọi là “svm\_model”. Sau khi dùng tập trainingsets để đào tạo bộ phân loại, bộ phân loại được kiểm định bằng hàm `predict` của Matlab để dự đoán kết quả của tập testingsets. Sau đó ta tính kết quả độ chính xác của tập testingsets để kiểm định bộ phân loại đã đạt yêu cầu đặt ra chưa. Và vẽ đồ thị AUC (Area Under the Curve) cho mô hình phân loại.

Đối với các bộ phân loại còn lại như CT hay KNN, phương pháp sử dụng cũng tương tự.

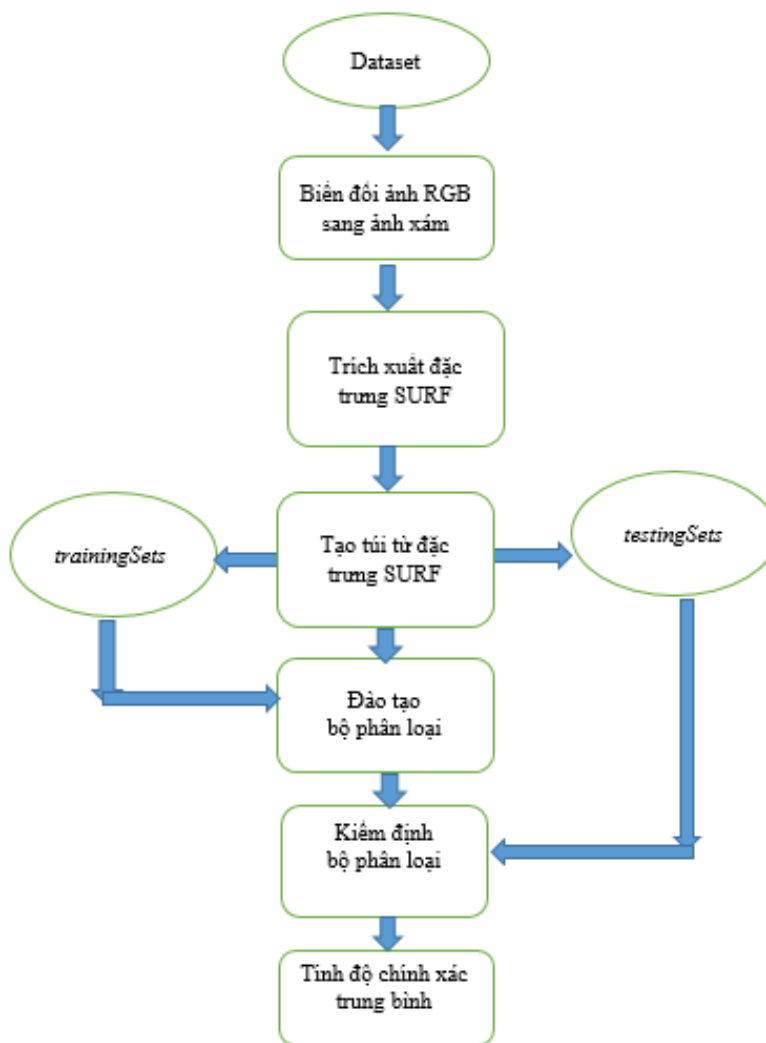
### **3.3 Đề xuất phương pháp túi từ đặc trưng SURF (SURF-based Bag of Words)**

Sơ đồ khối của phương pháp phân loại ảnh nội soi dạ dày bằng túi từ đặc trưng SURF chỉ khác sơ đồ khối Hình 3.1 trước đó là thêm một bước tạo túi từ BoW từ đặc trưng SURF đã tìm được.

Sau khi thu thập các đặc trưng SURF trong ảnh ta thu được một ma trận đặc trưng  $m \times n$  mỗi hàng là một vector đặc trưng. Các đặc trưng được trích chọn từ các ảnh trong cơ sở dữ liệu với số lượng rất lớn, lên đến vài trăm nghìn vector. Việc xử lý và tính toán với số lượng lớn đặc trưng như vậy, yêu cầu máy tính có bộ nhớ và tốc độ xử lý cao, thêm vào đó số lượng đặc trưng càng nhiều thì thời gian xử lý càng dài và hơn thế nữa

độ chính xác không hẳn là cao hơn. Vì vậy, việc làm giảm số lượng đặc trưng là công việc cần thiết.

Trong BoW ta cần giảm số lượng đặc trưng SURF mà vẫn chọn ra được đặc trưng nào là mạnh nhất có hiệu quả phân loại nhất trong ảnh, vì vậy việc chọn tham số lựa chọn đặc trưng mạnh nhất trong tổng số đặc trưng được trích chọn trong ảnh là vô cùng cần thiết. Ở trong luận văn này, các tác giả đề xuất lần lượt thử lựa chọn tăng dần số lượng đặc trưng được lựa chọn để tạo túi từ BoW để kiểm nghiệm là 10%, 20%, 30%, 50% và 70% tổng số đặc trưng SURF của ảnh có được.



Hình 3. 8 Sơ đồ khối mô hình phân loại sử dụng túi từ đặc trưng SURF

## Tổng kết chương

Chương 3 đã trình bày chi tiết quá trình thử nghiệm thuật toán phân loại ảnh nội soi dạ dày khi áp dụng các phương pháp trích chọn đặc trưng khác nhau: đặc trưng SURF,

BDIP, BVLC và túi từ đặc trưng SURF kết hợp cùng với các bộ phân loại khác nhau. Kết quả thử nghiệm của bài toán phân loại ảnh nội soi dạ dày sử dụng các phương pháp trích chọn đặc trưng sẽ được trình bày chi tiết trong chương 4.

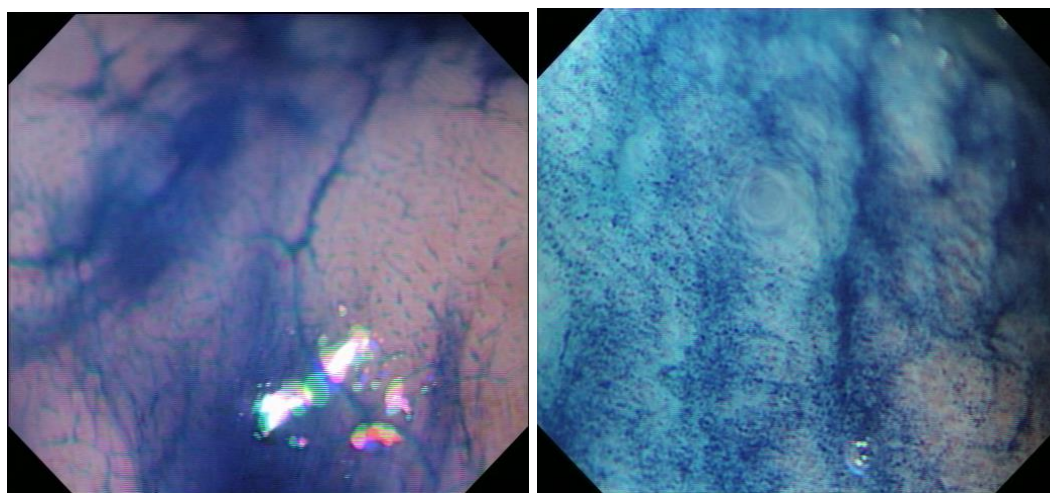


## CHƯƠNG 4. KẾT QUẢ THỬ NGHIỆM

Chương này trình bày các kết quả thử nghiệm bài toán tái định danh trên cơ sở dữ liệu của Bệnh viện Portuguese Institute of Oncology (IPO) ở Porto, Bồ Đào Nha. Nêu ra các kết quả thử nghiệm và so sánh kết quả thử nghiệm với kết quả với các bài báo trước đó.

### 4.1 Cơ sở dữ liệu

Cơ sở dữ liệu của bệnh viện Portuguese Institute of Oncology (IPO) ở Porto, Bồ Đào Nha là một bộ dữ liệu ảnh nội soi dạ dày nhuộm màu (Chromoendoscopy) đã được sử dụng để đánh giá mô hình phân loại trong các bài báo khoa học trước đó. Cơ sở dữ liệu gồm 176 ảnh, mỗi ảnh có kích thước 518x481, trong đó có 56 ảnh mô bình thường, và 120 ảnh mô bất thường (96 ảnh mắc chứng metaplasia và 24 ảnh thuộc triệu chứng dysplasia) cùng với vùng ảnh quan tâm (Region Of Interest) tương ứng với mỗi ảnh, được xác định bởi bác sĩ. Hình ảnh nội soi dạ dày nhuộm màu thu được bằng cách sử dụng một máy nội soi Olympus GIF-H180 tại Portuguese Institute of Oncology trong các nghiên cứu lâm sàng thông thường.



a)

b)

**Hình 4. 1 (a) Một ảnh nội soi thuộc lớp Normal, (b) Một ảnh nội soi thuộc lớp Abnormal**

Từ đó, các tác giả quyết định chọn cơ sở dữ liệu ảnh nội soi dạ dày này để thử nghiệm bài toán và có thể so sánh với kết quả nghiên cứu khác trên bộ dữ liệu này đã được công



bổ. Trong quá trình thử nghiệm phân loại, tập dữ liệu được chia thành 2 lớp ảnh là bình thường (Normal, gồm 56 ảnh) và bất bình thường (Abnormal, gồm 120 ảnh) để thực hiện. Trên Hình 4.1 minh họa 2 ảnh nội soi dạ dày bình thường và bất thường ở 2 lớp Normal và Abnormal trong cơ sở dữ liệu.

## 4.2 Phương pháp đánh giá

Có một loạt các cách thức hữu ích cho việc định lượng hiệu quả của một hệ thống phân loại ảnh nội soi dạ dày. Các cách thức phổ biến là ma trận nhầm lẫn (confusion matrix), độ chính xác trung bình (accuracy), đường cong ROC (Receiver Operating Characteristic) và AUC (Area Under the Curve).

Ma trận nhầm lẫn là thuật ngữ của học máy. Nó có bản chất là một bảng phân phối tần số 2 chiều (bảng chéo) cho phép trình bày tỷ lệ tương hợp và bất xứng giữa thực tế và kết quả phân loại của quy luật cần kiểm tra (mô hình). Trong bài này, ta chỉ xét trường hợp đơn giản nhất của ma trận nhầm lẫn áp dụng cho bài toán nhị phân (Binary classification). Khi đó, ma trận nhầm lẫn trình bày tần suất của 4 tổ hợp: TP (True Positive), TN (True Positive), FP (False positive) và FN (False negative). True (Đúng) chỉ sự tương hợp, khi kết quả phân loại phù hợp với giá trị thực tế; False (Sai) chỉ sự bất xứng hay nhầm lẫn, khi mô hình phân loại nhầm so với thực tế [33].

		Thực tế	
		Positive	Negative
Mô hình phân loại	Positive	TP	FP
	Negative	FN	TN

Hình 4. 2 Minh họa ma trận nhầm lẫn [33]

Độ chính xác trung bình [34] là loại số liệu đơn giản và thông dụng nhất để đánh giá một mô hình phân loại. Công thức tính độ chính xác được định nghĩa là tỷ lệ giữa số dự

đoán đúng trong tổng số các dự đoán. Với bài toán phân loại nhị phân, độ chính xác còn có thể được tính bằng :

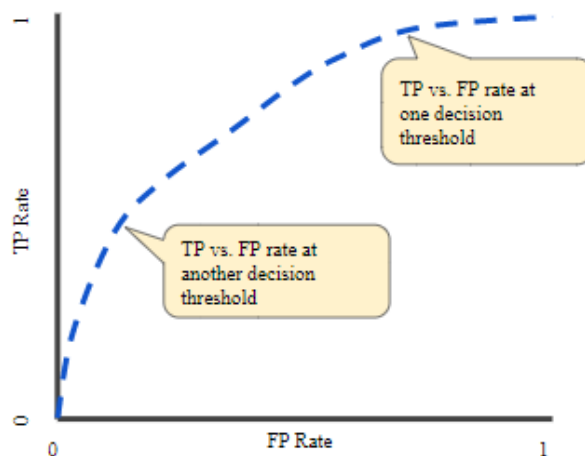
$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (4.1)$$

Với TP, TN, FP, FN đã được định nghĩa ở trên. Độ chính xác hoạt động hiệu quả khi số lượng mẫu ở mỗi lớp bằng nhau hoặc không chênh lệch nhiều. Ví dụ, nếu lớp A chiếm 90% tổng số lượng mẫu, lớp B chiếm 10% còn lại, vậy thì bộ phân loại sẽ dễ dàng đạt được độ chính xác huấn luyện 90% khi cho tất cả các mẫu vào lớp A. Đặc biệt, trong bài toán thực tế, cái giá của việc phân loại nhầm lớp thiểu số là rất lớn, ví dụ như bị ung thư mà phân loại thành không bị ung thư tạo thành hậu quả lớn hơn nhiều so với không bị ung thư mà phân loại thành có ung thư. Do đó cần thêm các tham số khác để đánh giá một bộ phân loại.

Đường cong ROC [34] là đồ thị biểu diễn hiệu quả của mô hình phân loại ở tất cả các ngưỡng. Đường cong này phụ thuộc vào hai tham số : TPR (True Positive Rate) và FPR (False Positive Rate), trong đó :

$$\text{TPR} = \frac{TP}{TP+FN} \quad (4.2)$$

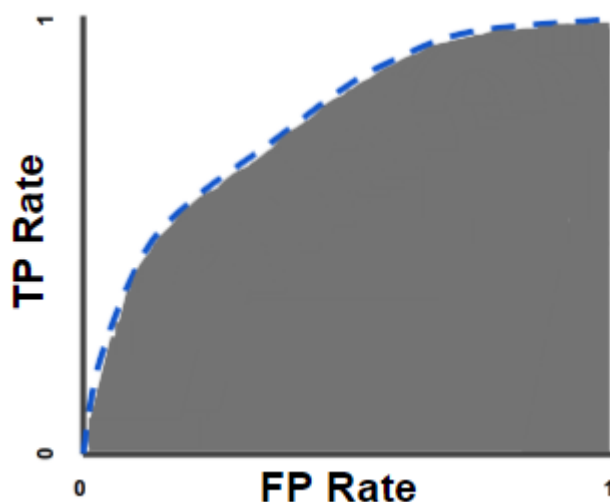
$$\text{FPR} = \frac{FP}{FP+TN} \quad (4.3)$$



**Hình 4. 3 Minh họa đường cong ROC [34]**

Đường cong ROC vẽ TPR và FPR ở các giá trị ngưỡng phân loại khác nhau. Ngưỡng phân loại càng thấp thì càng nhiều mẫu được phân loại vào Positive, do đó làm tăng cả FP và TP. Hình 4.3 mô tả một đường cong ROC điển hình.

Để tính giá trị các điểm trên đường cong ROC, có một cách hiệu quả hơn việc tính toán giá trị với từng ngưỡng khác nhau, đó là tính AUC - diện tích nằm bên dưới đường cong ROC. Giá trị AUC sẽ là một số dương nhỏ hơn hoặc bằng 1 và càng lớn thì mô hình càng tốt, và ít nhất từ 0.5 trở lên thì mới có thể chấp nhận.



Hình 4. 4 Minh họa AUC [34]

## 4.3 Kết quả thử nghiệm

### 4.3.1 Kết quả thử nghiệm đối với phân loại đặc trưng SURF

Cơ sở dữ liệu ban đầu được chia ngẫu nhiên thành hai tập trainingsets và testingsets với tỷ lệ 70%: 30%, tức là có 53 ảnh nội soi dạ dày trong testingsets còn lại 123 ảnh nằm trong trainingsets ; tỷ lệ giữa ảnh bình thường và bất thường trong trainingsets và testingsets giống cơ sở dữ liệu nguyên bản.

Bảng 4. 1 Ảnh hưởng của số lượng đặc trưng SURF được chọn lên kết quả phân loại [31]

Số lượng đặc trưng SURF lựa chọn	Độ chính xác trung bình (%)	AUC trung bình
50	78.5%	0.85
100	85.1%	0.92
1000	79.6%	0.91

Quá trình đào tạo và thử nghiệm được lặp lại 100 lần và kết quả cuối cùng là trung bình của 100 kết quả đó. Bảng 4.1 đưa ra độ chính xác trung bình và AUC trung bình tương ứng theo việc lựa chọn 50, 100 và 1000 tính năng SURF mạnh nhất với bộ phân

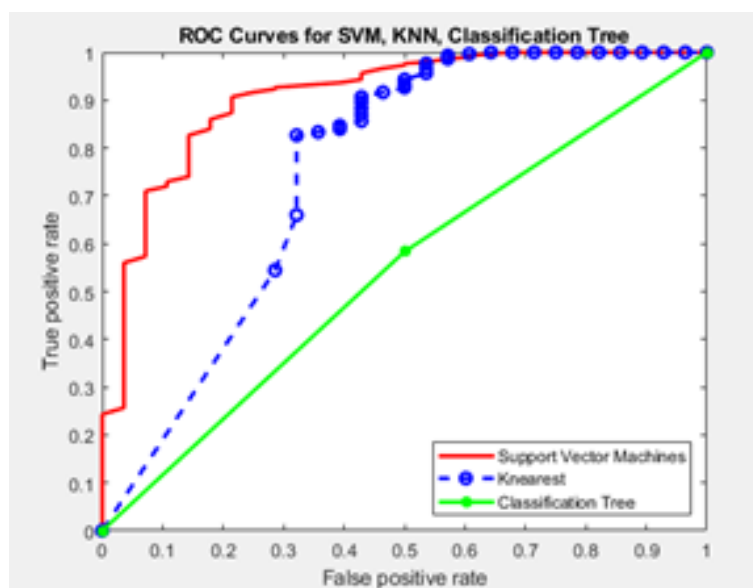
loại SVM. Từ kết quả nhận được và phân tích, các tác giả quyết định lựa chọn trích chọn 100 đặc trưng SURF mạnh nhất để phân loại.

Để so sánh hiệu suất của phương pháp được đề xuất với các phương pháp trích xuất đặc trưng hiện có, các tác giả quyết định thử nghiệm phân loại ảnh dựa vào đặc trưng KAZE [32].

**Bảng 4. 2 Kết quả phân loại của các bộ phân loại khác nhau kết hợp với đặc trưng SURF và KAZE [31]**

Đặc trưng + Bộ phân loại	Độ chính xác (%)	AUC
SURF+SVM	85.1%	0.92
SURF+CT	66.0%	0.59
SURF+KNN	71.7%	0.79
KAZE+SVM	81.1%	0.88
KAZE+CT	66.7%	0.62
KAZE+KNN	71.0%	0.68

Để xác nhận tính hiệu quả của phương pháp được đề xuất, các mô hình phân loại khác như cây phân loại CT và K phần tử lân cận gần nhất đã được thử nghiệm đối với đặc trưng SURF. Các kết quả được đưa ra trong Bảng 4.2 và Hình 4.1 chỉ ra rằng việc kết hợp đặc trưng SURF và bộ phân loại SVM là vượt trội so với bộ phân loại CT và KNN, với AUC trung bình của SVM đạt đến 0.92. Các tác giả cũng so sánh đặc trưng SURF với KAZE. Và nhận thấy rõ ràng là SURF cho kết quả tốt hơn KAZE.



**Hình 4. 5 Đường cong ROC cho 3 bộ phân loại ứng với đặc trưng SURF [31]**

#### 4.3.2 Kết quả thử nghiệm đối với phân loại túi từ đặc trưng SURF (SURF-based-BoW)

Đầu tiên thử nghiệm chọn số lượng đặc trưng SURF trên ảnh. Thử nghiệm lần lượt lấy 10%, 20%, 30%, 50% và 70% số lượng đặc trưng SURF mạnh nhất của ảnh đưa vào bộ phân loại. Bộ dữ liệu được chia ngẫu nhiên tỷ lệ đào tạo và kiểm định là 50%: 50%. Kết quả thử nghiệm được mô tả trong Bảng 4.3 dưới đây.

**Bảng 4. 3 Kết quả thay đổi số lượng đặc trưng trên ảnh**

Số lượng đặc trưng	Độ chính xác trung bình (%)	AUC trung bình
10%	98.5%	0.998
20%	93.5%	0.984
30%	94.8%	0.987
50%	91.9%	0.975
70%	90.4%	0.970

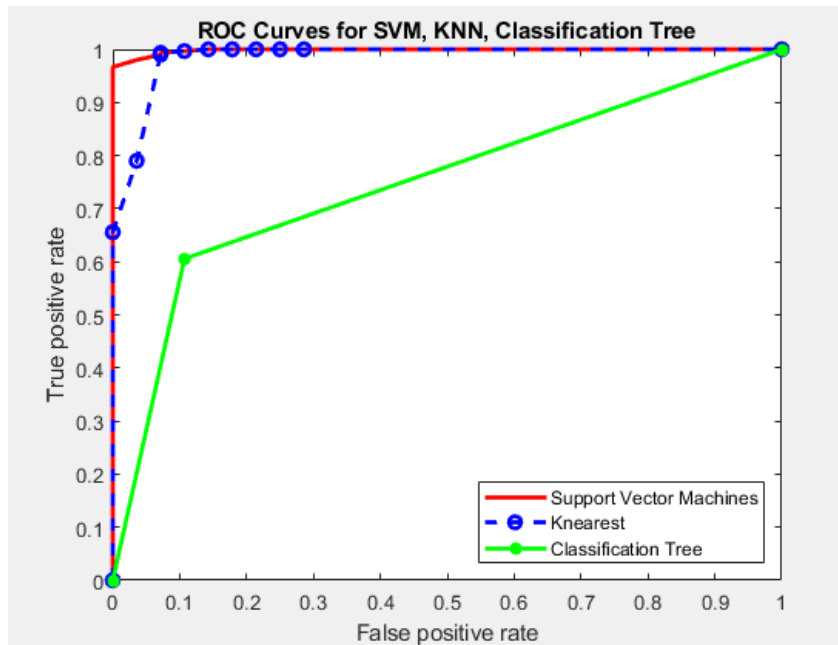
Từ kết quả của Bảng 4.3, quyết định chọn lấy số lượng đặc trưng là 10%. Thử nghiệm thay đổi tỷ lệ của hai tập huấn luyện và kiểm định lần lượt là 30%: 70%, 40%: 60%, 50%: 50%, 60%: 40%, 70%: 30% với các tỷ lệ sẽ lựa chọn ngẫu nhiên các ảnh nội soi dạ dày đưa vào hai tập. Bảng 4.4 mô tả kết quả thử nghiệm thay đổi tỷ lệ đào tạo - kiểm định.

Từ sự phân tích kết quả trong Bảng 4.4, các tác giả quyết định lựa chọn tỷ lệ đào tạo – kiểm định là 50%: 50%, cho độ chính xác phân loại ổn định, không sử dụng quá nhiều ảnh để đào tạo.

**Bảng 4. 4 Kết quả thay đổi tỷ lệ huấn luyện – kiểm định**

Đào tạo: Kiểm định (%)	Độ chính xác trung bình	AUC trung bình
30:70	97.5%	0.9980
40:60	98.0%	0.9985
50:50	98.5%	0.9984
60:40	98.8%	0.9982
70:30	98.7%	0.9993

Để xác nhận được tính hiệu quả của phương pháp đề xuất, chúng em thử sử dụng các bộ phân loại khác nhau như Cây phân loại (Classification Tree - CT) và K Nearest Neighbor (KNN). Các kết quả được đưa ra trong Hình 4.4 cho thấy rằng việc kết hợp SURF với SVM là vượt trội so với CT và cao hơn so với KNN.



Hình 4. 6 Đường cong ROC 3 bộ phân loại khác nhau đối với túi từ đặc trưng SURF

### 4.3.3 Kết quả thử nghiệm đối với phân loại đặc trưng BDIP và BVLC

#### 4.3.3.1 Thử nghiệm với đặc trưng BDIP

##### a) Kích thước khối khác nhau

Với cùng tỷ lệ trainingsets: testingsets = 7:3, ta lần lượt thử các kích thước khối khác nhau (không chồng lấp) là : 6x6, 7x7, 8x8 và 10x10 pixel để tìm ra kích thước cho hiệu quả tốt nhất. Kết quả thử nghiệm lần lượt theo thứ tự kích thước của các khối trong các Bảng 4.5, 4.6, 4.7, 4.8 dưới đây.

Bảng 4. 5 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 6x6 không chồng lấp

Lần thí nghiệm	Độ chính xác (%)	AUC
1	86.7925	0.872549
2	84.9057	0.908482
3	84.9057	0.9425
4	90.566	0.934783
5	83.0189	0.819473
6	77.3585	0.822857
7	83.0189	0.928571
8	84.9057	0.954023
9	83.0189	0.899038
10	81.1321	0.93125
<b>Trung bình</b>	<b>83.9623</b>	<b>0.901353</b>

**Bảng 4. 6 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 7x7 không chồng lấp**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	90.566	0.960784
2	88.6792	0.949346
3	92.4528	0.964052
4	90.566	0.949346
5	88.6792	0.934641
6	86.7925	0.908497
7	86.7925	0.95098
8	90.566	0.95098
9	83.0189	0.918301
10	83.0189	0.849673
<b>Trung bình</b>	<b>88.1132</b>	<b>0.93366</b>

**Bảng 4. 7 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 8x8 không chồng lấp**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	83.0189	0.916129
2	77.3585	0.836538
3	90.566	0.919118
4	81.1321	0.894624
5	81.1321	0.817204
6	84.9057	0.917647
7	88.6792	0.919913
8	79.2453	0.817284
9	79.2453	0.873047
10	86.7925	0.873272
<b>Trung bình</b>	<b>83.2075</b>	<b>0.878478</b>

**Bảng 4. 8 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 10x10 không chồng lấp**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	77.3585	0.849383
2	81.1321	0.9
3	69.8113	0.808333
4	79.2453	0.913131
5	83.0189	0.842742
6	81.1321	0.906173
7	79.2453	0.87619

8	84.9057	0.893145
9	88.6792	0.93287
10	83.0189	0.901786
<b>Trung bình</b>	80.7547	0.882375

**Kết luận:** từ những kết quả trên có thể thấy đặc trưng BDIP cho kết quả phân loại khá tốt (độ chính xác trung bình đạt trên 80%), và tốt nhất ở kích thước khối là 7x7 (đạt trên 88%).

*b) Tỷ lệ trainingsets: testingsets khác nhau*

Với cùng kích thước khối là 7x7, thử các tỷ lệ trainingsets: testingsets khác nhau để tìm ra tỷ lệ cho kết quả tốt nhất (tỷ lệ 7:3 đã có ở trên). Kết quả thử nghiệm lần lượt theo thứ tự kích thước của các khối trong các Bảng 4.9, 4.10, 4.11, 4.12 dưới đây.

**Bảng 4. 9 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 7x7 không chồng lấp với tỷ lệ trainingsests: testingsets=6:4**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	87.1429	0.939394
2	87.1429	0.91572
3	90	0.935606
4	87.1429	0.895833
5	87.1429	0.932765
6	91.4286	0.927083
7	88.5714	0.91572
8	90	0.941288
9	81.4286	0.909091
10	75.7143	0.86553
<b>Trung bình</b>	86.5714	0.917803

**Bảng 4. 10 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 7x7 không chồng lấp với tỷ lệ trainingsests: testingsets=5:5**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	87.5	0.901786
2	85.2273	0.892857
3	93.1818	0.9625
4	85.2273	0.885119
5	85.2273	0.925595
6	92.0455	0.967857
7	87.5	0.927381
8	82.9545	0.883929
9	86.3636	0.896429



10	81.8182	0.855357
<b>Trung bình</b>	86.7045	0.909881

**Bảng 4. 11 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 7x7 không chồng lấp với tỷ lệ trainingsets: testingsets=4:6**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	85.8491	0.907271
2	88.6792	0.921977
3	83.9623	0.910131
4	84.9057	0.904412
5	85.8491	0.898693
6	88.6792	0.931781
7	85.8491	0.88317
8	85.8491	0.917892
9	88.6792	0.933415
10	86.7925	0.948529
<b>Trung bình</b>	86.5094	0.915727

**Bảng 4. 12 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 7x7 không chồng lấp với tỷ lệ trainingsets: testingsets=3:7**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	84.5528	0.892552
2	86.9919	0.913919
3	86.1789	0.934676
4	86.1789	0.932845
5	81.3008	0.881258
6	86.1789	0.903541
7	87.8049	0.884615
8	86.9919	0.924908
9	86.1789	0.925824
10	88.6179	0.938645
<b>Trung bình</b>	86.0976	0.913278

**Kết luận:** từ bảng kết quả trên có thể thấy việc thay đổi tỷ lệ trainingsets: testingsets không làm ảnh hưởng nhiều đến kết quả phân loại (độ chính xác đều xấp xỉ 86% và AUC đều khoảng 91%).

*c) Kích thước khối 7x7 có chồng lấp*

Với cùng tỷ lệ trainingsets: testingsets = 7:3 và kích thước khối 7x7, thử nghiệm các tỷ lệ chồng lấp khác nhau để so sánh hiệu quả với cùng kích thước khối đó nhưng không

chồng lấp. Kết quả thử nghiệm lần lượt theo thứ tự kích thước của các khối trong các Bảng 4.13, 4.14, 4.15, 4.16 dưới đây.

**Bảng 4. 13 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 7x7 chồng lấp 1 pixel**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	86.7925	0.992424
2	83.0189	0.880952
3	88.6792	0.985
4	83.0189	0.890306
5	90.5660	0.924479
6	84.9057	0.954667
7	81.1321	0.816667
8	71.6981	0.827273
9	83.0189	0.905172
10	88.6792	0.935961
<b>Trung bình</b>	<b>84.1509</b>	<b>0.91129</b>

**Bảng 4. 14 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 7x7 chồng lấp 3 pixel**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	79.2453	0.854911
2	84.9057	0.848558
3	86.7925	0.939732
4	79.2453	0.857988
5	90.5660	0.946429
6	86.7925	0.835938
7	83.0189	0.906667
8	92.4528	0.947115
9	86.7925	0.944444
10	88.6792	0.941667
<b>Trung bình</b>	<b>85.4891</b>	<b>0.902345</b>

**Bảng 4. 15 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 7x7 chồng lấp 5 pixel**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	92.4528	0.901961

2	84.9057	0.927734
3	84.9057	0.953333
4	84.9057	0.93956
5	86.7925	0.969388
6	75.4717	0.811224
7	90.5660	0.945313
8	83.0189	0.892308
9	84.9057	0.933498
10	75.4717	0.846667
<b>Trung bình</b>	84.3396	0.912099

**Bảng 4. 16 Kết quả thử nghiệm đặc trưng BDIP với kích thước khối là 7x7 chồng lấp 6 pixel**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	92.4528	0.987879
2	90.5660	0.94321
3	83.0189	0.860784
4	79.2453	0.903704
5	84.9057	0.88425
6	88.6792	0.991398
7	79.2453	0.848739
8	88.6792	0.941176
9	86.7925	0.928241
10	84.9057	0.860041
<b>Trung bình</b>	85.8491	0.914942

**Kết luận:** Với cùng kích thước khối, việc thay đổi số pixel chồng lấp ít làm thay đổi kết quả phân loại (độ chính xác vẫn xấp xỉ 85%) và cũng không chênh lệch nhiều so với không có pixel chồng lấp.

#### 4.3.3.2 Thử nghiệm với đặc trưng BVLC

##### a) Kích thước khối khác nhau

Với cùng tỷ lệ trainingsets: testingsets = 7:3, ta lần lượt thử các kích thước khối khác nhau (không chồng lấp) là : 6x6, 7x7, 8x8 và 10x10 pixel để tìm ra kích thước cho hiệu quả tốt nhất. Kết quả thử nghiệm lần lượt theo thứ tự kích thước của các khối trong các Bảng 4.17, 4.18, 4.19, 4.20, 4.21, 4.22 dưới đây.

**Bảng 4. 17 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 10x10 không chồng lấp**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	67.9245	0.712418
2	71.6981	0.722222
3	56.6038	0.611111
4	66.0377	0.718954
5	66.0377	0.681373
6	73.5849	0.761438
7	71.6981	0.740196
8	69.8113	0.787582
9	66.0377	0.689542
10	62.2642	0.637255
<b>Trung bình</b>	67.1698	0.7062091

**Bảng 4. 18 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 15x15 không chồng lấp**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	60.3774	0.73366
2	71.6981	0.75
3	71.6981	0.776144
4	54.717	0.633987
5	73.5849	0.705882
6	69.8113	0.825163
7	67.9245	0.736928
8	60.3774	0.686275
9	62.2642	0.691176
10	71.6981	0.75
<b>Trung bình</b>	66.4151	0.728922

**Bảng 4. 19 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 25x25 không chồng lấp**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	58.3202	0.496732
2	71.6981	0.779412
3	77.3585	0.826797
4	71.6981	0.849673
5	71.6981	0.802288
6	64.1509	0.810458

7	77.3585	0.764706
8	66.0377	0.820261
9	67.9245	0.812092
10	77.3585	0.800654
<b>Trung bình</b>	69.8113	0.776307

**Bảng 4. 20 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 35x35 không chồng lấp**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	73.5849	0.784314
2	83.0189	0.854575
3	77.3585	0.836601
4	83.0189	0.931373
5	60.3774	0.841503
6	75.4717	0.828431
7	77.3585	0.771242
8	67.9245	0.776144
9	81.1321	0.851307
10	77.3585	0.687908
<b>Trung bình</b>	75.6604	0.81634

**Bảng 4. 21 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 45x45 không chồng lấp**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	84.9057	0.838235
2	75.4717	0.712418
3	75.4717	0.854575
4	79.2453	0.862745
5	79.2453	0.678105
6	75.4717	0.836601
7	77.3585	0.888889
8	88.6792	0.936275
9	71.6981	0.784314
10	79.2453	0.718954
<b>Trung bình</b>	78.6792	0.811111

**Bảng 4. 22 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 50x50 không chồng lấp**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	75.4717	0.870915
2	81.1321	0.838235
3	79.2453	0.816993
4	71.6981	0.838235
5	67.9245	0.792484
6	81.1321	0.820261
7	66.0377	0.643791
8	83.0189	0.874183
9	75.4717	0.820261
10	83.0189	0.771242
<b>Trung bình</b>	<b>76.4151</b>	<b>0.80866</b>

**Kết luận:** từ những kết quả trên có thể thấy đặc trưng BVLC cho kết quả phân loại trung bình (độ chính xác trung bình đạt khoảng 70%), và tốt nhất ở kích thước khối là 45x45 (đạt trên 78%).

*b) Tỷ lệ trainingsets: testingsets khác nhau*

Với cùng kích thước khối là 7x7, thử các tỷ lệ trainingsets: testingsets khác nhau để tìm ra tỷ lệ cho kết quả tốt nhất (tỷ lệ 7:3 đã có ở trên). Kết quả thử nghiệm lần lượt theo thứ tự kích thước của các khối trong các Bảng 4.23, 4.24, 4.25, 4.26 dưới đây.

**Bảng 4. 23 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 45x45 không chồng lấp với tỷ lệ trainingsets: testingsets=6:4**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	72.8571	0.798295
2	74.2857	0.642992
3	72.8571	0.755682
4	74.2857	0.787879
5	78.5714	0.848485
6	82.8571	0.82197
7	74.2857	0.750947
8	71.4286	0.660985
9	78.5714	0.760417
10	71.4286	0.797348
<b>Trung bình</b>	<b>75.1429</b>	<b>0.7625</b>

**Bảng 4. 24 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 45x45 không chồng lấp với tỷ lệ trainingsets: testingsets=5:5**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	75	0.788095
2	70.4545	0.704167
3	70.4545	0.71369
4	73.8636	0.796429
5	75	0.781548
6	75	0.804167
7	71.5909	0.722024
8	73.8636	0.7625
9	76.1364	0.705357
10	73.8636	0.7875
<b>Trung bình</b>	<b>73.5277</b>	<b>0.756548</b>

**Bảng 4. 25 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 45x45 không chồng lấp với tỷ lệ trainingsets: testingsets=4:6**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	72.6415	0.72018
2	66.9811	0.649918
3	77.3585	0.802288
4	78.3019	0.716095
5	76.4151	0.758987
6	82.0755	0.802288
7	69.8113	0.743056
8	68.8679	0.727533
9	70.7547	0.747958
10	79.2453	0.755719
<b>Trung bình</b>	<b>74.2453</b>	<b>0.742402</b>

**Bảng 4. 26 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 45x45 không chồng lấp với tỷ lệ trainingsets: testingsets=3:7**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	70.7317	0.590049
2	76.4228	0.750305
3	73.1707	0.749084
4	79.6748	0.872405

5	75.6098	0.791819
6	73.9837	0.804029
7	76.4228	0.766178
8	75.6098	0.753663
9	68.2927	0.741148
10	65.8537	0.736569
<b>Trung bình</b>	<b>73.5722</b>	<b>0.755525</b>

**Kết luận:** từ bảng kết quả trên có thể thấy việc thay đổi tỷ lệ trainingsets: testingsets không làm ảnh hưởng nhiều đến kết quả phân loại (độ chính xác đều xấp xỉ 74% và AUC đều khoảng 0.75).

*c) Kích thước khối 45x45 có chồng lấp*

Với cùng tỷ lệ trainingsets: testingsets = 7:3 và kích thước khối 45x45, thử nghiệm các tỷ lệ chồng lấp khác nhau để so sánh hiệu quả với cùng kích thước khối đó nhưng không chồng lấp. Kết quả thử nghiệm lần lượt theo thứ tự kích thước của các khối trong các Bảng 4.27, 4.28, 4.29, 4.30 dưới đây.

**Bảng 4. 27 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 45x45 chồng lấp 5 pixel**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	66.0377	0.779412
2	79.2453	0.882353
3	75.4717	0.795752
4	73.5849	0.748366
5	67.9245	0.694444
6	81.1321	0.825163
7	84.9057	0.901961
8	92.4528	0.933007
9	81.1321	0.812092
10	77.3585	0.838235
<b>Trung bình</b>	<b>77.9245</b>	<b>0.821078</b>

**Bảng 4. 28 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 45x45 chồng lấp 10 pixel**

Lần thí nghiệm	Độ chính xác (%)	AUC
1	77.3585	0.79902
2	83.0189	0.872549
3	79.2453	0.80719



4	77.3585	0.803922
5	79.2453	0.852941
6	64.1509	0.70098
7	75.4717	0.816993
8	79.2453	0.77451
9	79.2453	0.864379
10	77.3585	0.821895
<b>Trung bình</b>	77.1698	0.811438

**Bảng 4. 29 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 45x45 chồng lớp 15 pixel**

<b>Lần thí nghiệm</b>	<b>Độ chính xác (%)</b>	<b>AUC</b>
1	79.2453	0.897059
2	71.6981	0.710784
3	69.8113	0.727124
4	67.9245	0.818627
5	79.2453	0.849673
6	71.6981	0.756536
7	81.1321	0.905229
8	83.0189	0.785948
9	79.2453	0.849673
10	77.3585	0.826797
<b>Trung bình</b>	76.0377	0.812745

**Bảng 4. 30 Kết quả thử nghiệm đặc trưng BVLC với kích thước khối là 45x45 chồng lớp 20 pixel**

<b>Lần thí nghiệm</b>	<b>Độ chính xác (%)</b>	<b>AUC</b>
1	66.0377	0.723856
2	83.0189	0.870915
3	81.1321	0.895425
4	67.9245	0.841503
5	67.9245	0.738562
6	77.3585	0.833333
7	67.9245	0.738562
8	77.3585	0.833333
9	81.1321	0.864379
10	84.9057	0.897059
<b>Trung bình</b>	75.4717	0.823693

**Kết luận:** với cùng kích thước khối, việc thay đổi số pixel chồng lấp ít làm thay đổi kết quả phân loại (độ chính xác vẫn xấp xỉ 75%) và cũng không chênh lệch nhiều so với không có pixel chồng lấp.

#### 4.3.4 So sánh các phương pháp đề xuất với các phương pháp khác

Từ những kết quả đạt được qua thử nghiệm, đồ án đưa ra Bảng 4.31 so sánh các kết quả của phương pháp đề xuất với các phương pháp khác đã được công bố trước đó. Ta nhận thấy những phương pháp đề xuất cho kết quả khá là khả quan, đặc biệt là áp dụng túi từ đặc trưng SURF đã cải thiện hơn so với chỉ áp dụng đặc trưng SURF không rất nhiều, và kết quả cũng cho thấy độ chính xác cùng AUC của phương pháp này cao hơn so với bài báo được công bố gần đây của Hussam Ali và cộng sự [10] năm 2017 khi thử nghiệm trên cùng 1 cơ sở dữ liệu. Một phần kết quả đạt được trong phân loại ảnh nội soi dạ dày bằng đặc trưng SURF cũng đã được chấp nhận trình bày tại hội thảo International Technical Conference on Circuits/Systems, Computers and Communications 2019 tổ chức tại Jeju Shinhwa World, Hàn Quốc từ ngày 23-26 tháng 6 năm 2019. Tất cả những kết quả này đều mở ra những hướng phát triển tiếp trong tương lai của đề tài, hứa hẹn những kết quả tốt hơn nữa.

**Bảng 4. 31 So sánh các đặc trưng và bộ phân loại**

<b>Đặc trưng và bộ phân loại sử dụng</b>	<b>Độ chính xác trung bình (%)</b>	<b>AUC trung bình</b>
Gabor + Genetic algorithm + SVM [10]	90%	0.93
Túi từ Random Forests + SVM [1]	87.7%	Không đề cập
Trích chọn đặc trưng từ Superpixel [30]	86.6%	Không đề cập
Local Self Similar + SVM [8]	88.5%	Không đề cập
<b>SURF +SVM [31]</b>	<b>81.5%</b>	<b>0.92</b>
<b>Túi từ đặc trưng SURF + SVM</b>	<b>98.5%</b>	<b>0.99</b>
<b>BDIP + SVM</b>	<b>88.1%</b>	<b>0.93</b>
<b>BVLC + SVM</b>	<b>78.7%</b>	<b>0.81</b>

### Tổng kết chương

Chương 4 đã trình bày chi tiết các kết quả đạt được trong quá trình thử nghiệm bài toán tái phân loại ảnh nội soi dạ dày sử dụng các phương pháp trích chọn đặc trưng khác nhau. Những kết quả này thu được với cùng một điều kiện đầu vào (cùng cơ sở dữ liệu

của viện Portuguese Institute of Oncology ở Porto, Bồ Đào Nha) hoàn toàn giống với các thử nghiệm của các bài báo trước đó nên có độ tin cậy cao. Các kết quả trên chứng minh được sự rằng việc áp dụng phương pháp túi từ đặc trưng SURF hay SURF hoặc BDIP đã giúp cho hiệu quả phân loại của của bài toán khá khả quan và đáng để đầu tư nghiên cứu thêm.

## KẾT LUẬN

Sau khoảng thời gian tích cực và nghiêm túc thực hiện đồ án tốt nghiệp “Phân loại ảnh ung thư dạ dày dựa vào các đặc trưng”, chúng em đã đạt được một số kết quả sau:

- Có một cái nhìn tổng quan về các nghiên cứu hiện nay trong lĩnh vực phân loại ảnh sử dụng Machine Learning, biết được những hướng tiếp cận và hiệu quả của từng đặc trưng được đề xuất.
- Hiểu được ý nghĩa quan trọng của phân loại ảnh trong các ứng dụng trong thực tế, những khó khăn và thách thức mà phân loại ảnh vẫn còn đang gặp phải.
- Hiểu được quy trình làm việc của một bài toán phân loại ảnh sử dụng Machine Learning.
- Thực hiện thử nghiệm thuật toán trên Matlab, đồng thời đã áp dụng và chứng minh phương pháp được sử dụng làm tăng hiệu quả phân loại.
- Hoàn thành mục tiêu đề ra: thực hiện thuật toán trích xuất đặc trưng trên MATLAB và sử dụng những đặc trưng đó cho bộ phân loại ảnh, so sánh hiệu quả giữa các loại đặc trưng cũng như với những phương pháp đã được đề xuất.

Qua quá trình thử nghiệm trên cùng một cơ sở dữ liệu ảnh nội soi dạ dày Chromoendoscope (Viện Ung thư Bồ Đào Nha), có thể thấy rằng việc sử dụng túi từ đặc trưng SURF cho kết quả phân loại 98.5% cao hơn kết quả của phương pháp trích chọn đặc trưng Gabor kết hợp Genetic algorithm của bài báo gốc 92% [10]. Như vậy, việc sử dụng đặc trưng đã thực hiện có ưu thế hơn. Hướng phát triển đề tài là tìm thêm các cơ sở dữ liệu khác về ảnh nội soi hoặc ảnh về loại bệnh khác để thử nghiệm, tăng tính khách quan trong việc đánh giá tính hiệu quả của bộ phân loại và đặc trưng trong phân loại ảnh; tìm thêm một số loại đặc trưng khác để thử nghiệm. Với hướng tiếp cận này, chúng em hy vọng sẽ cải thiện được tốt hơn kết quả của bài toán phân loại.

## TÀI LIỆU THAM KHẢO

- [1] Sara Isabel Moreira Francisco. (2015). “Recognition of Cancer using Random Forests as a Bag-of-Words Approach for Gastroenterology”. Available: . [https://web.fe.up.pt/~up200900631/Dissertation\\_up200900631.pdf](https://web.fe.up.pt/~up200900631/Dissertation_up200900631.pdf). Lần truy cập cuối: 7/6/2019.
- [2] [https://vi.wikipedia.org/wiki/Ung\\_th%C6%B0\\_d%E1%BA%A1\\_d%C3%A0y](https://vi.wikipedia.org/wiki/Ung_th%C6%B0_d%E1%BA%A1_d%C3%A0y). Lần truy cập cuối: 7/6/2019.
- [3] <http://www.tuvankhoe.com/suc-khoe/benh-ung-thu-da-day-giai-doan-dau-va-giai-doan-cuoi.html>. Lần truy cập cuối: 7/6/2019.
- [4] <https://helloworld.com/benh/ung-thu-da-day/>. Lần truy cập cuối: 7/6/2019.
- [5] K. Onji, S. Yoshida, S. Tanaka, R. Kawase, Y. Takemura, S. Oka, T. Tamaki, B. Raytchev, K. Kaneda, M. Yoshihara, and K. Chayama, “Quantitative analysis of colorectal lesions observed on magnified endoscopy images,” Journal of Gastroenterology, vol. 46, pp. 1382–1390, 2011, 10.1007/s00535-011-0459-x.
- [6] F. Riaz, F. Silva, M. Ribeiro, and M. Coimbra, “Invariant Gabor texture descriptors for classification of gastro enterology images,” IEEE Transactions on Biomedical Engineering, vol. PP, no. 99, p. 1, 2012.
- [7] R. Kwitt, N. Vasconcelos, N. Rasiwasia, A. Uhl, B. Davis, M. Hafner, and F. Wrba, “Endoscopic image analysis in semantic space,” Medical Image Analysis, vol. 16, no. 7, pp. 1415 – 1422, 2012.
- [8] R. Sousa, D. C. Moura<sup>1</sup>, M. Dinis-Ribeiro, and M. T. Coimbra<sup>1</sup>, “Local Self Similar Descriptors: Comparison and Application to Gastroenterology Images”, Conf Proc IEEE Eng Med Biol Soc. 2014; 2014:4635-8.
- [9] A. F. Constantinescu, M. Ionescu, I. Rogoveanu, M.E. Ciurea, C.T. Streba, V.F. Iovanescu, S.A. Artene, C.C. Vere, “Analysis of wireless capsule endoscopy images using local binary patterns”, Appl. Med. Inf. 36 (2) (2015) 31.
- [10] Hussam Ali, Muhammad Sharif, Mussarat Yasmin and Mubashir Husain Rehmani, “Computer-based classification of chromoendoscopy images using homogeneous texture descriptors”, Computers in Biology and Medicine, vol. 88, 2017, pp. 84–92.
- [11] H. Bay, A. Ess, T. Tuytelaars and L. Van Gool, “Speeded-up robust features,” Computer Vision and Image Understanding, vol.110, no.3, pp.346-359, 2008.

- [12] <https://www.mathworks.com/help/vision/ug/image-classification-with-bag-of-visual-words.html>. Lần truy cập cuối: 7/6/2019.
- [13] H. J. So, M. H. Kim, N. C. Kim (2009) “Texture classification using waveletdomain BDIP and BVLC features”. 17th European Signal Processing Conference (EUSIPCO 2009), 2009, pp. 1117-1120.
- [14] Dalal, Navneet, and Bill Triggs. “Histograms of oriented gradients for human detection.” Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. Vol. 1. IEEE, 2005.
- [15] Lowe, David G. “Object recognition from local scale-invariant features”. Computer vision, 1999. The proceedings of the seventh IEEE international conference on. Vol. 2. Ieee, 1999.
- [16] Ryan Layne, Timothy Hospedales and Shaogang Gong. “Person re-identification by attributes”. *CCTV* 2012.
- [17] Hirzer, M., Beleznai, C., Roth, P. M., & Bischof, H. (2011, May). “Person re-identification by descriptive and discriminative classification”. In Scandinavian conference on Image analysis (pp. 91-102).
- [18] [https://en.wikipedia.org/wiki/Bag-of-words\\_model](https://en.wikipedia.org/wiki/Bag-of-words_model). Lần truy cập cuối: 7/6/2019.
- [19] <https://codelungtung.com/2018/10/25/image-search-engine-mo-hinh-bag-of-visual-words/>. Lần truy cập cuối: 7/6/2019.
- [20] <https://machinelearningcoban.com/general/2017/02/06/featureengineering/>. Lần truy cập cuối: 7/6/2019.
- [21] [https://vi.wikipedia.org/wiki/H%E1%BB%8Dc\\_m%C3%A1y](https://vi.wikipedia.org/wiki/H%E1%BB%8Dc_m%C3%A1y). Lần truy cập cuối: 7/6/2019.
- [22] [https://www.researchgate.net/publication/323108787\\_Introduction\\_to\\_Machine\\_Learning/figures?lo=1](https://www.researchgate.net/publication/323108787_Introduction_to_Machine_Learning/figures?lo=1). Lần truy cập cuối: 7/6/2019.
- [23] <https://techtalk.vn/tat-tan-tat-moi-kien-thuc-co-ban-ve-machine-learning.html>. Lần truy cập cuối: 7/6/2019.
- [24] [https://vi.wikipedia.org/wiki/M%C3%A1y\\_vect%C6%A1\\_h%E1%BB%97\\_tr%E1%BB%A3](https://vi.wikipedia.org/wiki/M%C3%A1y_vect%C6%A1_h%E1%BB%97_tr%E1%BB%A3). Lần truy cập cuối: 7/6/2019.
- [25] [https://lib.hpu.edu.vn/bitstream/handle/123456789/18298/41\\_PhamVanSon\\_CT1201.pdf](https://lib.hpu.edu.vn/bitstream/handle/123456789/18298/41_PhamVanSon_CT1201.pdf). Lần truy cập cuối: 7/6/2019.
- [26] [https://en.wikipedia.org/wiki/K-nearest\\_neighbors\\_algorithm](https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm). Lần truy cập cuối: 7/6/2019.

- [27] <https://machinelearningcoban.com/2017/01/08/knn/>. Lần truy cập cuối: 7/6/2019.
- [28] [https://vi.wikipedia.org/wiki/C%C3%A2y\\_quy%E1%BA%BFt\\_%C4%91%E1%BB%8Bnh](https://vi.wikipedia.org/wiki/C%C3%A2y_quy%E1%BA%BFt_%C4%91%E1%BB%8Bnh). Lần truy cập cuối: 7/6/2019.
- [29] <https://machinelearningcoban.com/2018/01/14/id3/>. Lần truy cập cuối: 7/6/2019.
- [30] Davide Boschetto; Enrico Grisan. “Superpixel-based classification of gastric chromoendoscopy images”. Proceedings Volume 10134, Medical Imaging 2017: Computer-Aided Diagnosis; 101340W (2017).
- [31] Dung. V. N, Hien. T. T, “Speeded-Up Robust Feature Descriptor for Endochromoscopy Images”, The 34th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC 2019), Jeju, Hàn Quốc, 23-26.06.2019.
- [32] Alcantarilla, P.F., A. Bartoli, and A.J. Davison. "KAZE Features." *ECCV 2012*, Part VI, LNCS 7577. 2012, p. 214.
- [33] <https://rpubs.com/lengockhanhi/347941>. Lần truy cập cuối: 7/6/2019.
- [34] <https://developers.google.com/machine-learning/crash-course/classification/accuracy?fbclid=IwAR26Q0zkNMXetA-X-XPT-6zNVT0ZYdEaC3qUNmq0rng7aQ5fuDkWpgPrEKA>. Lần truy cập cuối: 7/6/2019.
- [35] Trịnh Anh Dương, Chữ Văn Thành, Trần Hồng Giang, Nguyễn Đức Thiện. Bài tập lớn “Phân loại ảnh ung thư dạ dày sử dụng mô hình túi từ trực quan”, 06.2017.