

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



ĐẶNG THANH TUẤN

**XÂY DỰNG HỆ THỐNG NHẬN DẠNG THỨC ĂN VÀ DINH
DƯỠNG TRÊN NỀN TẢNG ANDROID**

CHUYÊN NGÀNH: KHOA HỌC MÁY TÍNH
MÃ SỐ: 60.48.01.01

TÓM TẮT LUẬN VĂN THẠC SĨ

HÀ NỘI - 2016

Luận văn được hoàn thành tại:
HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Người hướng dẫn khoa học: **Tiến sĩ Ngô Quốc Dũng**

Phản biện 1:

Phản biện 2:

Luận văn sẽ được bảo vệ trước Hội đồng chấm luận văn thạc sĩ tại Học viện Công nghệ Bưu chính Viễn thông

Vào lúc: giờ ngày tháng năm

Có thể tìm hiểu luận văn tại:

- Thư viện của Học viện Công nghệ Bưu chính Viễn thông

MỞ ĐẦU

1. Lý do chọn đề tài

Thức ăn hàng ngày là một phần tất yếu của cuộc sống. Thức ăn với chế độ dinh dưỡng hợp lý sẽ đem lại sức khỏe tốt cho con người. Ngược lại, chế độ dinh dưỡng không hợp lý có thể dẫn tới một số bệnh lý như tiểu đường, béo phì, mỡ trong máu, tim mạch.

Trong những năm gần đây, bệnh tiểu đường (hay còn gọi là đái tháo đường) đang trở thành vấn đề lo ngại lớn đối với giới y khoa và cả cộng đồng.

Chi phí chữa trị cho bệnh tiểu đường cũng ngốn ngân sách khá lớn ở các quốc gia.

Theo dự kiến người mắc tiểu đường sẽ tăng lên khoảng 7-8 triệu người vào năm 2025. Đây chính là con số đáng báo động cho thấy tốc độ gia tăng rất nhanh về số lượng người bệnh. Nghiêm trọng hơn, thực trạng bệnh tiểu đường và tim mạch đang bị trẻ hóa do béo phì gây ra nhiều biến chứng khó lường. Điều này cho thấy cần có một chế độ dinh dưỡng khoa học riêng cho từng cá nhân, đảm bảo tốt cuộc sống mà vẫn duy trì được sức khỏe cơ thể của mình.

Bên cạnh đó, hiện nay, để có một chiếc điện thoại thông minh không còn là một vấn đề khó khăn nữa, việc sử dụng điện thoại di động hỗ trợ chế độ dinh dưỡng hợp lý thông qua hình ảnh là một phương pháp đem lại hiệu quả đáng kể. Nó kết nối nhanh khoảng cách giữa sức khỏe cá nhân và công nghệ, người dùng và thông tin dinh dưỡng của họ, bằng cách giám sát thực phẩm dựa trên thiết bị tính toán cầm tay cá nhân thông qua chế độ ăn uống hàng ngày. Cho phép họ theo dõi tình trạng sức khỏe hàng ngày để giảm chi phí y tế tổng thể. Do đó, việc xây dựng một hệ thống nhận dạng thức ăn và dinh dưỡng hàng ngày dựa trên điện thoại di động với một chi phí thấp, tự động, mang tính cá nhân là cần thiết. Chính vì vậy, em quyết định chọn đề tài: **“Xây dựng hệ thống nhận dạng thức ăn và dinh dưỡng trên nền tảng Android”**.

2. Mục đích nghiên cứu

Hệ thống nhận dạng thức ăn và dinh dưỡng dựa trên điện thoại di động giúp cho người dùng thuận tiện hơn trong việc quản lý và theo dõi chế độ dinh dưỡng hàng ngày thông qua việc ăn uống, góp phần nâng cao hiệu quả trong việc giữ gìn, duy trì sức khỏe hoặc giảm cân.

3. Đối tượng và phạm vi nghiên cứu

Đối tượng nghiên cứu: Trên thực tế thức ăn hàng ngày rất đa dạng, phong phú và rất khó nhận dạng vì đặc điểm bên ngoài của chúng có thể khá giống nhau (ví dụ như thịt lợn và thịt bò). Nên trong khuôn khổ luận văn này đề tài tập trung vào nhận dạng thức ăn đơn.

Phạm vi nghiên cứu: Phát triển một hệ thống nhận dạng thức ăn đơn trên điện thoại di động thông qua việc phân loại hình ảnh của chúng.

4. Phương pháp nghiên cứu

Luận văn sử dụng thuật toán phân cụm K-means để phân đoạn ảnh rau quả và các món ăn thường ngày. Sau đó sử dụng thuật toán hỗ trợ máy vector SVM (support vector machine) để phân loại thức ăn. Dữ liệu về hình ảnh thức ăn (khoảng 300 - 400 ảnh từ 15 loại thức ăn khác nhau) sẽ được thu thập bằng cách chụp hình trên điện thoại di động với độ phân giải thấp và một số hình ảnh được tìm kiếm từ internet. Cuối cùng thuật toán nhận dạng sẽ được đánh giá trên tập dữ liệu này.

Nội dung của luận văn sẽ được trình bày trong 3 chương chính với những nội dung như sau :

Chương 1. Tổng quan về nhận dạng thức ăn và dinh dưỡng

Trong chương này, luận văn sẽ trình bày tính cấp thiết cần phải có một ứng dụng quản lý dinh dưỡng. Từ đó mục tiêu của luận văn là nhắm đến xây dựng một ứng dụng dựa trên nền tảng android nhằm quản lý chế độ dinh dưỡng cho người sử dụng đồng thời giới thiệu tổng quan về bài toán nhận dạng thức ăn và dinh dưỡng trên nền tảng Android.

Chương 2. Phương pháp nhận dạng thức ăn và dinh dưỡng

Chương này sẽ trình bày một số khái niệm về phân loại và xử lý ảnh, một số phương pháp liên quan đến các thuộc tính, kỹ thuật của việc nhận dạng thức ăn dựa trên hình ảnh thông qua các thuật toán trích chọn đặc trưng. Qua đó làm tiền đề để nghiên cứu sâu hơn các kỹ thuật trích chọn nội dung ảnh và đối sánh ảnh.

Chương 3. Cài đặt thử nghiệm và đánh giá

Chương này sẽ tiến hành cài đặt, xây dựng chương trình nhận dạng thức ăn và dinh dưỡng trên nền tảng Android, chạy thử, đưa ra các kết quả của hệ thống dựa trên một số thuật toán trích chọn đặc trưng ảnh, đồng thời sẽ có những đánh giá khách quan về các kết quả đạt được.

CHƯƠNG 1 : TỔNG QUAN VỀ BÀI TOÁN NHẬN DẠNG THỨC ĂN VÀ DINH DƯỠNG

1.1. Đặt vấn đề

Vấn đề kiểm soát ăn uống trở nên vô cùng hệ trọng cho từng trường hợp cá biệt, nhất là những người có vấn đề không tốt về sức khỏe. Chính vì vậy mà áp dụng một chương trình dinh dưỡng hợp lý là một trong những biện pháp điều trị căn bản, cần thiết và quan trọng, không chỉ giúp phòng ngừa bệnh tiểu đường mà còn giúp phòng ngừa hàng loạt các bệnh mãn tính khác như tim mạch, cao huyết áp, ung thư,... Nhờ đó góp phần giúp cho cuộc sống của những thành viên trong gia đình và rộng hơn nữa là cuộc sống của cộng đồng xã hội ngày càng được cải thiện hơn.

Cùng với đó, thực phẩm cung cấp cho mỗi bữa ăn thường rất đa dạng, việc lựa chọn thực phẩm như thế nào cho phù hợp đang là những băn khoăn cho nhiều người nội trợ. Họ cần rất nhiều hiểu biết về vấn đề này cũng như cần có những hỗ trợ riêng để ước tính lượng Calo từ thực phẩm cung cấp cho các bữa ăn hàng ngày., nhu cầu năng lượng cho người già, người lao động, người trẻ, trẻ em, người bệnh,..

Để có được tư vấn thường xuyên của các chuyên gia Y tế là một hạn chế và thường gây ra sự tốn kém cho người cần tư vấn về sức khỏe.

Do đó, việc xây dựng một hệ thống nhận dạng thức ăn và dinh dưỡng trên mobile trợ giúp cho người bình thường ăn kiêng, giảm cân..., đặc biệt không chỉ cho những người bị bệnh tiểu đường mà những người bệnh khác một chế độ dinh dưỡng hợp lý là bắt buộc và cần thiết.

1.2. Những nghiên cứu trước đây về nhận dạng ảnh

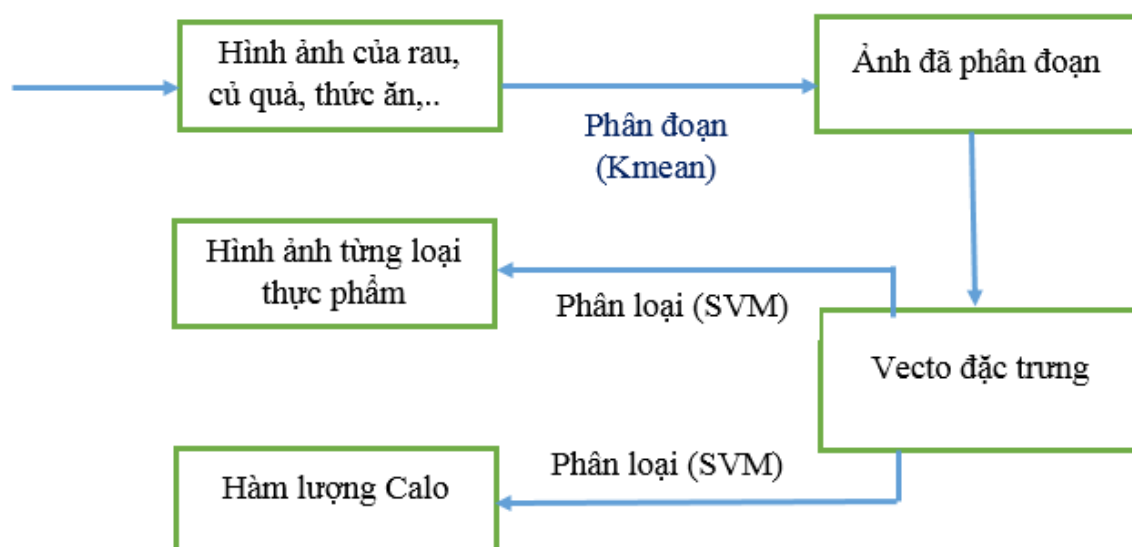
Xử lý ảnh, phân loại ảnh trong những năm gần đây là lĩnh vực đang được giới khoa học quan tâm không những trong mà còn ngoài nước.

1.2.1. Các nghiên cứu đáng chú ý trên thế giới.

1.2.2. Các nghiên cứu đáng chú ý tại Việt Nam.

1.3. Cách tiếp cận

Để thuận lợi cho việc theo dõi tình trạng sức khỏe hàng ngày, việc xây dựng một hệ thống nhận dạng thức ăn bằng hình ảnh là cần thiết, nó mang tính cá nhân với chi phí thấp, rất tiện lợi cho người dùng. Từ thực phẩm được nhận dạng sẽ ước tính lượng calo tương ứng cho mỗi loại. Đây cũng là giải pháp cho vấn đề quản lý dinh dưỡng đã được đề cập ở trên. Hình ảnh thức ăn có được từ người dùng chụp lại bằng điện di động với độ phân giải thấp. Từ những hình ảnh này hệ thống sẽ thực hiện phân loại thực phẩm qua hai bước: phân đoạn và làm mịn ảnh. Mỗi hình ảnh có được là từ bề mặt của các loại thức ăn, rau, quả, củ, số lượng màu sắc chưa xác định, một phương thức không có giám sát là cần thiết để phân vùng ảnh. Và thuật toán phân cụm K-mean được thực hiện cho việc phân đoạn màu sắc thực của ảnh. Khi chụp hình trong khung cảnh thực những hình ảnh chứa đựng những đối tượng xung quanh đầy ý nghĩa, làm phát sinh một số cụm (nhỏ) không có liên quan đến đối tượng quan tâm. Để loại bỏ những đối tượng này, ta áp dụng các toán tử hình thái (mathematical morphology) bao gồm open và close để làm mịn hình ảnh phân đoạn. Vùng chính cần quan tâm có thức ăn sẽ được trích xuất bằng việc sử dụng các thuộc tính màu sắc và SURF cho việc xác nhận thức ăn. Cuối cùng thực hiện sự phân lớp ảnh cho việc xác nhận từng loại thức ăn dựa trên đặc trưng SURF và thuộc tính màu sắc thông qua thuật toán máy hỗ trợ vector (SVM) [1][5].



Hình 1.2 : Sơ đồ nhận dạng thức ăn

1.4. Bài toán đặt ra

Luận văn xây dựng một ứng dụng nhằm giải quyết vấn đề trên. Thao tác của người dùng rất đơn giản, họ chỉ việc chụp ảnh món ăn, các thao tác còn lại như tính toán và quản lý về thông tin lượng dinh dưỡng của món ăn sẽ do ứng dụng đảm nhiệm.

1.5. Tổng quan về ứng dụng

Ứng dụng sẽ nhận dạng thức ăn và hàm lượng dinh dưỡng (calo) cho mỗi người dùng dựa trên cơ sở dữ liệu mà ứng dụng đã học được thông qua việc chụp ảnh thức ăn bằng camera của điện thoại Android và các hình ảnh về món ăn được sưu tầm thông qua internet. Từ đó giúp cho người dùng có thể tự tính toán, điều chỉnh được chế độ ăn theo từng bữa ăn cho phù hợp [12],[16].

Các thông tin nhận dạng bao gồm:

- Tên thức ăn
- Hàm lượng dinh dưỡng (Calo) trong thức ăn mà ứng dụng nhận dạng được.

1.6. Kết luận chương 1

Trong chương này, luận văn đã trình bày tính cấp thiết cần phải có một ứng dụng quản lý dinh dưỡng và một số hạn chế của các ứng dụng đã có. Từ đó mục tiêu luận văn là nhắm đến là xây dựng một ứng dụng trên nền tảng android nhằm quản lý chế độ dinh dưỡng cho người sử dụng.

Để nâng cao tính tương tác giữa ứng dụng và người sử dụng, luận văn sẽ nhắm đến một ứng dụng thời gian thực với độ chính xác cao.

CHƯƠNG 2 : PHƯƠNG PHÁP NHẬN DẠNG THỨC ĂN VÀ DINH DƯỠNG.

2.1. Giới thiệu

Chương này đi sâu vào trình bày hai thuật toán chính sử dụng trong hệ thống nhận dạng thực phẩm bằng hình ảnh là thuật toán K-mean và thuật toán máy hỗ trợ vector (SVM). Cách phân đoạn ảnh đầu vào dựa trên thuật toán K-mean để trích chọn các đặc trưng. Và sử dụng thuật toán SVM cho việc phân loại ảnh.

2.2. Đặt vấn đề

Phân loại hình ảnh là một trong nhiều lĩnh vực được chú ý nhất và đã được nghiên cứu trong những năm gần đây.

Một cách tổng quan, phân loại hình ảnh là quá trình gán các ảnh vào một hay nhiều lớp ảnh đã được xác định từ trước. Người ta có thể phân loại ảnh bằng cách thủ công, tức là nhìn vào nội dung tấm ảnh sau đó gán chúng vào một hay nhiều lớp cụ thể nào đó. Hệ thống quản lý tập ảnh gồm nhiều ảnh cho nên việc làm này sẽ tốn rất nhiều thời gian, công sức và do đó là không khả thi. Do vậy mà phải có các phương pháp phân loại tự động. Để phân loại tự động, người ta sử dụng các phương pháp máy học trong trí tuệ nhân tạo như cây quyết định, Naive Bayes, K láng giềng gần nhất, giải thuật clustering,...[14]

Một trong những ứng dụng quan trọng nhất của phân loại ảnh tự động là trong các hệ thống tìm kiếm ảnh. Từ một tập con các ảnh đã phân lớp sẵn, tất cả các ảnh trong miền tìm kiếm sẽ được xử lý rút ra những vector đặc trưng sau đó được đưa qua phân nhóm để hình thành các nhóm đặc trưng và được gán nhãn lớp tương ứng.

2.3. Phân đoạn và tiền xử lý ảnh

2.3.1. Giới thiệu

2.3.2. Phân đoạn ảnh bằng thuật toán K-MEAN .

K-Mean là thuật toán rất quan trọng và được sử dụng phổ biến trong kỹ thuật phân cụm. Tư tưởng chính của thuật toán K-Mean là tìm cách phân nhóm các đối tượng đã cho vào K cụm (K là số các cụm được xác định trước, K nguyên dương) sao cho tổng bình phương khoảng cách giữa các đối tượng đến tâm nhóm là nhỏ nhất.

Giải thuật K- mean:

1. Chọn ngẫu nhiên K tâm (centroid) cho K cụm (cluster). Mỗi cụm được đại diện bằng các tâm của cụm.
2. Tính khoảng cách giữa các đối tượng (objects) đến K tâm (thường dùng khoảng cách Euclidean)
3. Nhóm các đối tượng vào nhóm gần nhất
4. Xác định lại tâm mới cho các nhóm
5. Thực hiện lại bước 2 cho đến khi không có sự thay đổi nhóm nào của các đối tượng

2.3.3. Phân đoạn bằng phương pháp K-Nearest Neighbor(KNN)[14]

K-Nearest Neighbor là phương pháp truyền thống khá nổi tiếng theo hướng tiếp cận thống kê đã được nghiên cứu trong nhiều năm qua. KNN được đánh giá là một trong những phương pháp tốt nhất được sử dụng từ những thời kỳ đầu trong nghiên cứu về phân loại hình ảnh.

2.3.4. Tiền xử lý ảnh

Để việc nhận dạng ảnh chỉ tập trung vào đúng đối tượng quan tâm. Ta cần loại bỏ những đối tượng không liên quan tới nhận dạng như phong nền xung quanh đối tượng quan tâm, dụng cụ xử lý thực phẩm, vật đựng thực phẩm,... Nhằm tăng cường chất lượng ảnh, mà công đoạn tiền xử lý là bước đầu tiên nhằm loại bỏ nhiễu, khắc phục những khiếm khuyết do bước thu nhận ảnh không tốt là việc làm quan trọng.

2.4. Một số phương pháp trích chọn đặc trưng và đo độ tương đồng giữa các ảnh.

2.4.1. Khái niệm đặc trưng ảnh số

2.4.2. Tại sao phải trích chọn đặc trưng ảnh số ?

2.4.3. Trích chọn đặc trưng dựa trên điểm nổi bật

2.4.4. Các phương pháp trích chọn đặc trưng[2]

2.4.4.1. Phương pháp Harris

2.4.4.2. Phương pháp SIFT(Scale Invarian Feature Transform) [3],[5],[15]

- a. Phát hiện cực trị*
- b. Định vị chính xác điểm khóa*
- c. Gán hướng cho các điểm khóa*
- d. Xây dựng bộ mô tả cục bộ*

2.4.4.3. Phương pháp SURF(Speed Up Robust Feature) [6],[13]

SURF là một phương pháp phát hiện và mô tả hình ảnh mạnh mẽ, do Herbert Bay đưa ra lần đầu vào năm 2006, có thể được sử dụng để nhận dạng đối tượng, phục hồi 3D. Nó được lấy cảm hứng từ mô tả SIFT. Phiên bản tiêu chuẩn của SURF nhanh hơn nhiều so với SIFT và nó chống lại sự biến đổi hình ảnh khác nhau tốt hơn SIFT. SURF dựa trên tổng xấp xỉ các đặc trưng Haar wavelet 2D và sử dụng hiệu quả cho các ảnh tích hợp (integral image).

Phương pháp này dựa trên các ma trận Hessian, nhưng sử dụng một xấp xỉ cơ bản, cũng giống như DoG là một phát hiện cơ bản dựa trên Laplacian. Nó dựa trên hình ảnh tích hợp để giảm thời gian tính toán và được gọi nó là phát hiện 'Fast-Hessian'. Mặt khác, mô tả này mô tả một phân bố các đặc trưng Haarwavelet trong các láng giềng của điểm quan tâm. Hơn nữa, chỉ sử dụng kích thước 64, để giảm thời gian tính toán các đặc trưng và so khớp, đồng thời tăng sự chắc chắn.

- a. Ảnh tích hợp(integral image)*
- b. Phát hiện Fast_Hessian*
- c. Sự biểu diễn không gian tỷ lệ*
- d. Định vị điểm quan tâm*
- e. Bộ mô tả điểm quan tâm và so khớp*
- f. Gán hướng*

g. Bộ mô tả dựa trên tổng các đặc trưng Haar Wavelet

h. Lập chỉ mục và so khớp

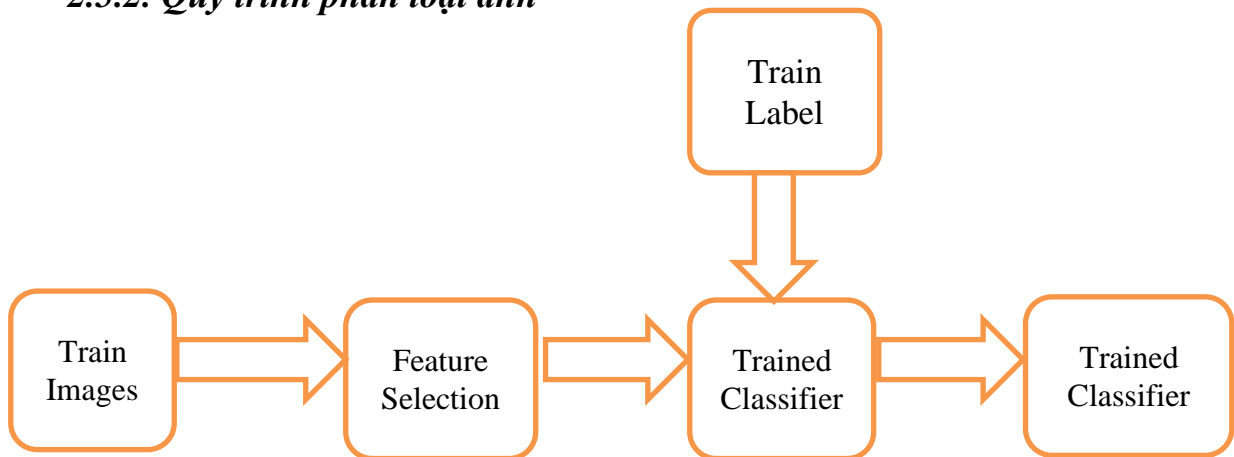
Nếu độ tương phản giữa 2 điểm quan tâm khác nhau (tối trên nền sáng và sáng trên nền tối), ứng viên sẽ không được xem là so khớp có giá trị.

2.5. Các phương pháp phân loại ảnh

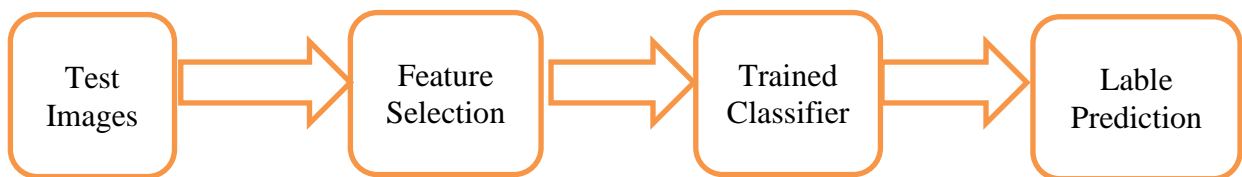
2.5.1. Tổng quan

Phân loại ảnh là một quá trình phức tạp và đòi hỏi phải xem xét nhiều yếu tố. Nhìn chung các bước chính của quá trình phân loại ảnh có thể bao gồm: xác định một hệ thống phân loại cụ thể và phù hợp, lựa chọn mẫu đào tạo, tiền xử lý ảnh, khai thác tính năng, lựa chọn các phương pháp tiếp cận phân loại phù hợp, và đánh giá độ chính xác. Nhu cầu của người sử dụng, quy mô của công trình nghiên cứu, điều kiện kinh tế, và kỹ năng của nhà phân tích là những yếu tố quan trọng ảnh hưởng đến sự lựa chọn dữ liệu, thiết kế của các thủ tục phân loại, và kết quả của các thủ tục phân loại.

2.5.2. Quy trình phân loại ảnh



Hình 2.13 : Quy trình huấn luyện ảnh



Hình 2.14 : Quy trình kiểm thử ảnh

2.5.3. Phân loại ảnh bằng phương pháp Support Vector Machine(SVM)

2.5.3.1. Tổng quát về SVM

Phân lớp (classification) và dự đoán (prediction) là hai bài toán cơ bản và có rất nhiều ứng dụng trong tất cả các lĩnh vực. Có nhiều phương pháp đã được nghiên cứu và ứng dụng cho các bài toán dạng này như: mạng Noron nhân tạo, phương pháp học thống kê,... Trong phần này, chúng ta sẽ đi sâu nghiên cứu phương pháp Support Vector Machines, một phương pháp rất hiệu quả hiện nay.

SVM được coi là công cụ mạnh cho những bài toán phân lớp phi tuyến. Phương pháp này thực hiện phân lớp dựa trên nguyên lý cực tiểu hóa rủi ro có cấu trúc SRM (Structural Risk Minimization), được xem là một trong các phương pháp phân lớp giám sát không tham số tinh vi nhất cho đến nay. Các hàm công cụ đa dạng của SVM cho phép tạo không gian chuyển đổi để xây dựng mặt phẳng phân lớp.

SVM dạng chuẩn lấy một tập hợp các dữ liệu đầu vào và dự báo mỗi dữ liệu đầu vào ứng với một lớp (class) trong số hai lớp mà dữ liệu ấy có khả năng rơi vào. Điều này làm cho dạng chuẩn SVM trở nên không có tính xác suất mà là một công cụ binary tuyến tính.

Bài toán phân lớp, sử dụng SVM nhằm mục đích tìm một siêu phẳng có biên cực đại giữa lớp mẫu âm và mẫu dương, đồng thời cực tiểu hóa các mẫu không phân chia được trong tập huấn luyện. SVM dựa trên cơ sở toán học vững chắc. Tuy nhiên việc huấn luyện mẫu sử dụng SVM đòi hỏi phải giải bài toán tối ưu nhiều biến. Ban đầu, SVM được phát triển để giải các bài toán phân lớp, về sau do tính ưu việt, nó còn được ứng dụng rộng rãi để giải các bài toán hồi quy.

2.5.3.2. Bài toán phân lớp

Phân lớp là quá trình nhóm các đối tượng “giống” nhau vào “một lớp” dựa trên các đặc trưng dữ liệu của chúng. Tuy nhiên, phân lớp là một hoạt động tiềm ẩn trong tư duy con người khi nhận dạng thế giới thực, đóng vai trò quan trọng làm cơ sở đưa ra các dự báo, các quyết định. Phân lớp và cách mô tả các lớp giúp cho tri thức được định dạng và lưu trữ trong đó.

Việc giải bài toán phân lớp là quá trình xây dựng một mô hình (các hàm, các luật...) để quyết định mỗi đối tượng, vật thể thuộc vào lớp nào dựa trên các đặc trưng dữ liệu của chúng. Quá trình phân lớp sẽ trở nên khó khăn hơn khi đối tượng có nhiều đặc trưng dữ liệu, và quá trình phân lớp cũng cần xác định đặc trưng nào cần thiết cho việc phân lớp, đặc trưng nào dư thừa.

2.5.3.3. 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 dữ liệu mới xt thì cần phải xác định xt được phân vào lớp +1 hay lớp -1?

2.5.3.4. Bài toán phân nhiều lớp với SVM

Để phân nhiều lớp thì kỹ thuật SVM nguyên thủy sẽ chia không gian dữ liệu thành 2 phần và quá trình này lặp lại nhiều lần.

Như vậy, bài toán phân nhiều lớp sử dụng phương pháp SVM hoàn toàn có thể thực hiện giống như bài toán hai lớp. Bằng cách sử dụng chiến lược “một đối một” (one – against – one).

Giả sử bài toán cần phân loại có k lớp ($k > 1$), chiến lược “một đối một” 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 dựa vào bài toán phân hai lớp bằng phương pháp SVM.

2.6. Phương pháp đề xuất

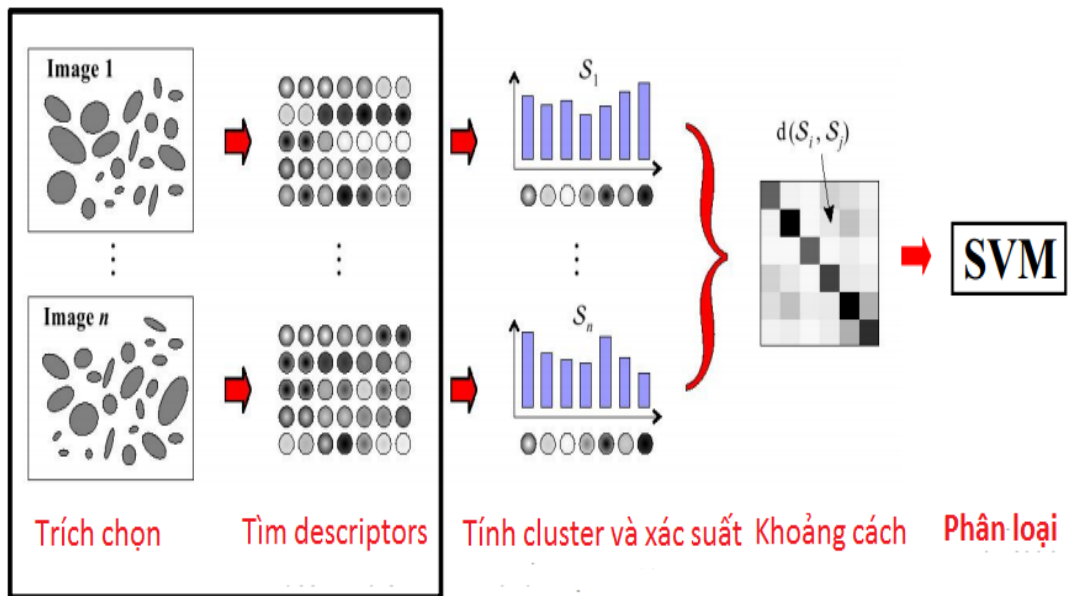
Khi xử lý ảnh cho bước đầu tiên trong phân đoạn ảnh ta sẽ gặp một số trường hợp như:

- *Vị trí thực phẩm trong hình*
- *Sự đa dạng về hình dạng*
- *Sự thay đổi trong kết cấu hình ảnh*

Với những vấn đề gặp phải, việc lựa chọn xây dựng điểm cục bộ bất biến là cần thiết. Luận văn áp dụng phân tích thành phần đặc trưng bất biến của tấm ảnh cụ thể là thuật toán SURF phục vụ cho việc tiến hành rút lấy những đặc trưng bất biến

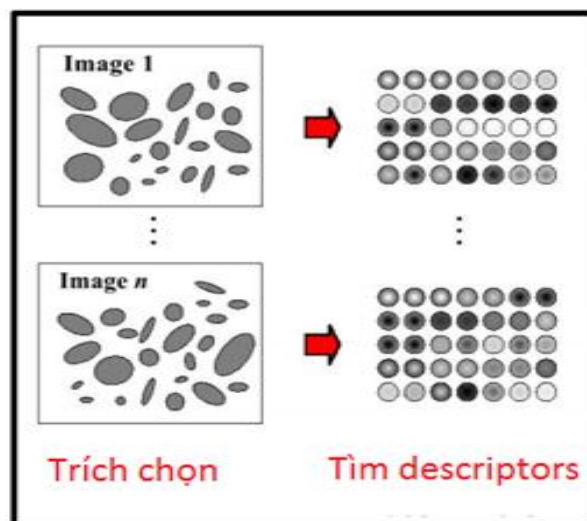
của ảnh sau đó dùng giải thuật SVM để tiến hành phân lớp. Thực nghiệm cho thấy độ chính xác cao, phù hợp để áp dụng vào hệ thống phân lớp tự động.

2.6.1. Các bước cơ bản



Hình 2.16 : Tổng quan phương pháp phân nhiều lớp với SVM

Bước 1: Trích chọn đặc trưng



Hình 2.17: Trích chọn đặc trưng

Phần này gồm 2 bước:

- **Detector:** Phát hiện Interested Point(Keypoint)

- **Descriptor:** Sau khi phát hiện Interested Point thì cần một cách để biểu diễn Interested Point đó thành một feature vector, và đó là nhiệm vụ của Descriptor. Descriptor của SURF cho ra một feature vector có 64 phần tử.

Bản chất của SURF là chọn ra một số keypoint, rồi dùng descriptor để đặc trưng cho một vùng xung quanh keypoint. Nhờ đó các feature vector của SURF là bất biến với một local image region.

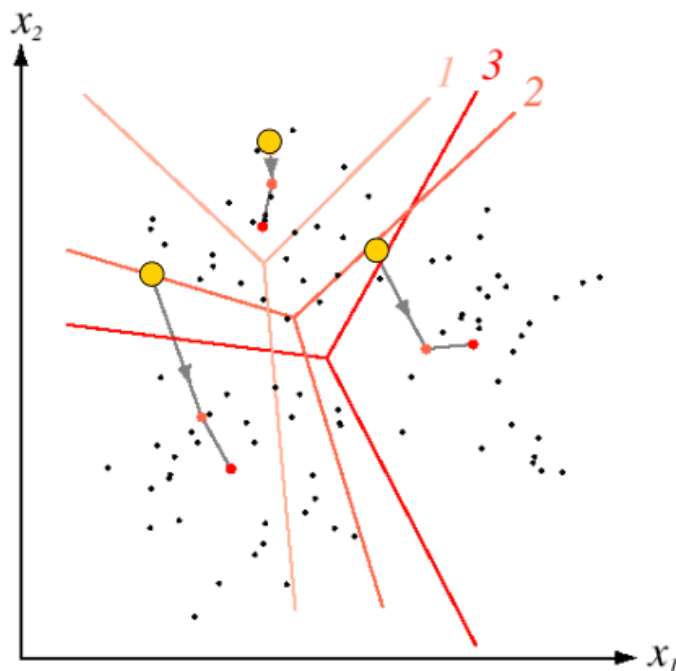
SURF bất biến với scale là do khi tính descriptor, SURF tính ở scale của keypoint. Tức là trong bước detect, nó phát hiện keypoint ở điểm nào, thì sẽ dùng ảnh ở scale đó để tính descriptor.

Bước 2:

Phần này gồm 3 bước chính:

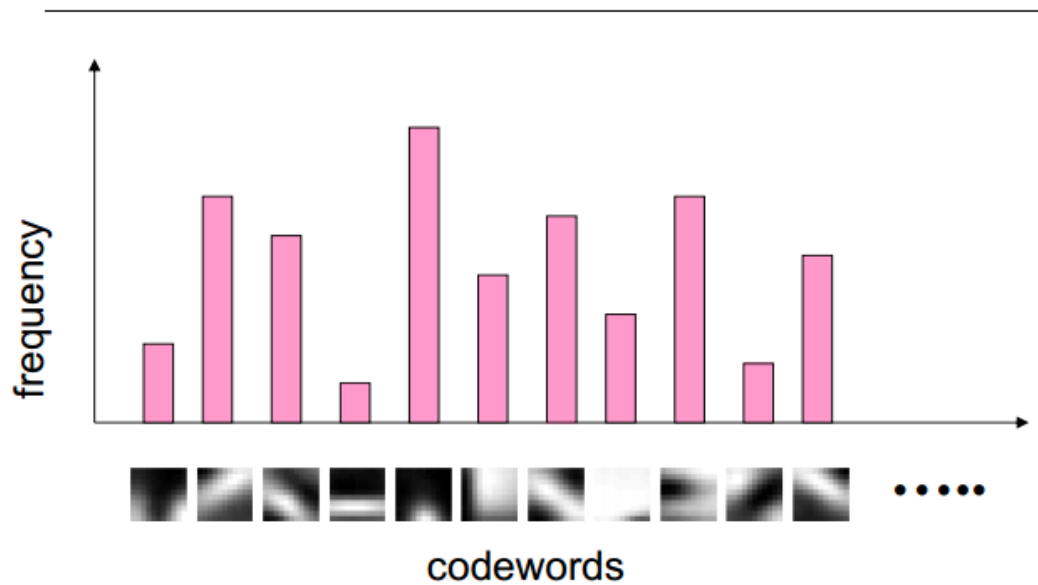
Nhóm descriptor phát hiện ở bước 1 vào các nhóm bằng thuật toán KMeans.

Ví dụ: KMeans với $K=3$



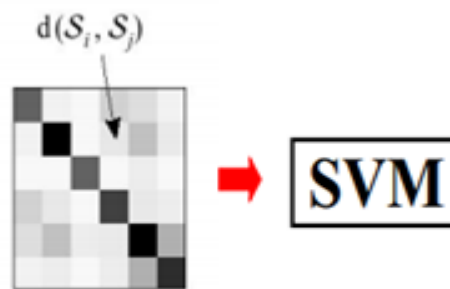
Hình 2.18: Nhóm các descriptor.

Xây dựng biểu đồ xác xuất :



Hình 2.19: Biểu đồ xác suất

Bước 3: Phân loại với SVM



Hình 2.20: Phân loại với SVM

Dựa vào dữ liệu thu được từ bước trước và thuật toán SVM, xác định nhãn đúng cho đối tượng cần nhận diện.

2.7. Kết luận chương 2

Qua chương 2, luận văn đã làm rõ một số yếu tố quan trọng ảnh hưởng đến quá trình phân loại ảnh:

Cần một tập dữ liệu huấn luyện chuẩn và đủ lớn để cho thuật toán học phân loại. Nếu chúng ta có được một tập dữ liệu chuẩn và đủ lớn thì quá trình huấn luyện sẽ tốt và khi đó chúng ta sẽ có kết quả phân loại tốt sau khi đã được học.

Các phương pháp trên hầu hết đều sử dụng mô hình vector để biểu diễn ảnh, do đó phương pháp rút trích đặc trưng trong ảnh đóng vai trò quan trọng quá trình biểu diễn ảnh bằng vector.

Thuật toán sử dụng để phân loại phải có thời gian xử lý hợp lý, thời gian này bao gồm: thời gian học, thời gian phân loại ảnh, ngoài ra thuật toán này phải có tính tăng cường (incremental function) nghĩa là không phân loại lại toàn bộ tập ảnh khi thêm một số ảnh mới vào tập dữ liệu mà chỉ phân loại các ảnh mới mà thôi, khi đó thuật toán phải có khả năng giảm độ nhiễu (noise) khi phân loại ảnh.

CHƯƠNG 3: CÀI ĐẶT THỬ NGHIỆM VÀ ĐÁNH GIÁ

Trong chương này, đề tài sẽ trình bày các phần sau đây:

- 3.1. Xây dựng bộ dữ liệu.
- 3.2. Giao diện chương trình chạy trên Android
- 3.3. Những vấn đề gặp phải
- 3.4. Kết luận chương 3

3.1. Xây dựng bộ dữ liệu

3.1.1. So sánh giữa các thuật toán trích chọn đặc trưng

Bảng 3.1: So sánh các thuật toán trích chọn đặc trưng

Phương pháp	Kết quả	Thời gian(giây)
SURF	38173	4
SIFT	40907	5
Hessian-Laplace	30524	3

Từ bảng kết quả, phương pháp Hessian-Laplace tuy thời gian thực hiện là nhanh nhất nhưng số lượng điểm ảnh rút ra được lại thấp hơn hết. Điều này ảnh hưởng không nhỏ tới tính chính xác của chương trình, còn về hai phương pháp còn lại, SIFT tuy có kết quả nhanh hơn SURF nhưng lại thời gian thực hiện lại chậm hơn, ảnh hưởng đến sự tương tác của ứng dụng với người dùng. Với những điều này em nhận thấy là khả thi và quyết định chọn SURF làm phương pháp rút ra đặc trưng ảnh cho chương trình.

3.1.2. Xây dựng bộ dữ liệu đánh giá độ chính xác của thuật toán *SURF*

Bộ dữ liệu gồm 15 phần, trong mỗi phần gồm 20 ảnh chụp ở các góc độ khác nhau.

Ảnh test phải đảm bảo khác với ảnh đầu vào.

Số lượng ảnh đầu vào bắt đầu từ 1 đến 5.

Bảng 3.2: Tỷ lệ xác suất Test ảnh

Input	Test	Tỉ lệ
1	3	60.25 %
2	3	72.6
3	5	83.2 %
4	5	91.5 %
5	5	92 %

3.1.3. Xây dựng bộ dữ liệu của ứng dụng

Trong phạm vi luận văn, em xây dựng bộ dữ liệu gồm 15 món ăn quen thuộc với người dân Việt Nam.

- Gà quay.
- Thịt xiên nướng.
- Tôm hùm.
- Bánh giò
- Na.
- Cà chua.
- Khoai tây chiên.
- Phở bò.
- Đậu rán.
- Cá rán.
- Trứng luộc.
- Rau muống xào.
- Chuối.
- Quả táo.
- Bánh mì

Xây dựng 15 bộ dữ liệu cho 15 món ăn, mỗi bộ gồm 20 ảnh; trong đó có 15 ảnh để huấn luyện và 05 ảnh để kiểm tra.

Kích thước ảnh của mỗi ảnh trong bộ dữ liệu là 100*250.

Mỗi bộ dữ liệu được gán một nhãn cố định để xử lý và nhận dạng

3.1.3.1. Phần huấn luyện:

Khởi tạo biến:

Điều chỉnh lại kích thước của ảnh về 100*250:

Tìm ra và mô tả điểm đặc trưng của tất cả các ảnh

Phân cụm tất cả các điểm đặc trưng dựa vào thuật toán KMeans.

Ghi lại tâm của các nhóm để sử dụng bên phần nhận dạng:

Gán nhãn cho mỗi nhóm, nhãn chính là tên thư mục chứa món ăn.

Tính tần xuất xuất hiện:

Sử dụng SVM xây dựng bộ huấn luyện lưu dưới file xml để sử dụng bên phần nhận dạng.

3.1.3.2. Phần nhận dạng:

Thực hiện các thao tác tương tự như với phần huấn luyện nhưng với tập ảnh nhận dạng.

Xác định nhãn của ảnh nhận dạng:

Kết quả result là nhãn của ảnh cần nhận dạng.

3.1.3.3. Kết quả:

3.2. Giao diện chương trình chạy trên Android

3.3. Những vấn đề gặp phải

Xây dựng bộ dữ liệu huấn luyện.

Khác với món ăn nước ngoài, món ăn Việt Nam thường là sự pha trộn của nhiều loại thực phẩm.

Sự đa dạng của món ăn. Ví dụ cùng món bò xào giá, tùy từng nơi mà tỉ lệ của bò và giá khác nhau, dẫn đến lượng dinh dưỡng cũng khác nhau.

Món ăn của Việt Nam rất đa dạng, dẫn tới khó khăn trong việc cập nhật món ăn cho ứng dụng.

Bộ dữ liệu trong luận văn chưa phong phú.

Chưa xác định được khối lượng món ăn, người dùng vẫn phải nhập khối lượng món ăn, gây ra một số bất tiện.

3.4. Kết luận chương 3

Tìm ra các điểm đặc trưng bất biến là một vấn đề quan trọng trong nhận dạng, vì các đặc trưng bất biến với các yếu tố như co giãn, ánh sáng, góc nhìn,... giúp việc nhận dạng tăng độ chính xác. Có một số phương pháp tìm ra các đặc trưng bất biến như Hessian Laplace, SIFT, SURF. Không có phương pháp nào tối ưu, mỗi phương pháp đều có ưu nhược điểm riêng của riêng mình. Phương pháp Hessian Laplace nhanh nhưng tìm được ít keypoint nên độ chính xác không cao, phương pháp SIFT có độ chính xác cao nhưng thời gian thực hiện khá lâu, ảnh hưởng đến thời gian nhận dạng. Trong phạm vi luận văn, em thấy phương pháp SURF vừa đảm bảo được độ chính xác, vừa đảm bảo được thời gian xử lý.

Sau khi trích trồn được đặc trưng, bước tiếp theo là phân loại ảnh. Có nhiều phương pháp phân loại ảnh, trong đó phương pháp SVM khá hiệu quả vì SVM có ưu điểm với các thuật toán khác như KNN, NB, tất cả dữ liệu trong tập huấn luyện đều được dùng để tối ưu hóa kết quả.

KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

1. Những kết quả đã đạt được

Luận văn đã nghiên cứu việc ứng dụng các phương pháp trích chọn đặc trưng và phân loại ảnh nhằm giải quyết bài toán nhận dạng thức ăn và dinh dưỡng trên nền tảng Android. Cụ thể, luận văn đã thực hiện được các nội dung sau:

- Nghiên cứu các thuật toán khác nhau trong việc trích chọn đặc trưng và phân loại ảnh. Tìm hiểu những điểm mạnh, yếu của các thuật toán này. Đề xuất phương pháp.
- Xây dựng ứng dụng nhận dạng thức ăn và dinh dưỡng trên nền tảng Android dựa vào thuật toán SURF và SVM.
- Phân tích yêu cầu và thiết kế ứng dụng.
- Xây dựng bộ dữ liệu với những món ăn quen thuộc người Việt Nam.

Tuy nhiên vẫn còn một số vấn đề mà luận văn chưa giải quyết được:

- Chưa tự động xác định được khối lượng món ăn do người dùng chụp lại.
- Bộ dữ liệu chưa phong phú.
- Gặp khó khăn với những món ăn là sự pha trộn của nhiều loại thực phẩm.

2. Hướng phát triển tiếp theo của luận văn

Tập trung cải thiện độ chính xác khi nhận dạng món ăn, tính được khối lượng và giải quyết những vấn đề khó khăn với những món ăn phức tạp.

Phát triển ứng dụng và xây dựng thành hệ thống quản lý dinh dưỡng dựa vào việc nhận dạng ảnh và hàm lượng dinh dưỡng của món ăn. Dựa vào các thông tin người dùng về độ tuổi, cân nặng, chiều cao,...., cùng với việc nhận dạng thức ăn và chỉ số dinh dưỡng của món ăn hiện tại để từ đó có thể tính toán, đưa ra lời khuyên cho người sử dụng nên hoặc không nên bổ sung thêm những thành phần dinh dưỡng nào cho phù hợp với thể trạng hiện tại. Từ đó có thể giúp người sử dụng có được một chế độ dinh dưỡng hợp lý cho mình.

DANH MỤC TÀI LIỆU THAM KHẢO

Tài liệu Tiếng Việt

- [1] Trần Cao Đệ, Phạm Nguyên Khang (2012), *Phân loại văn bản với máy học vector hỗ trợ và cây quyết định*, Số tạp chí 21a(2012) Trang: 52-63.
- [2] Trần Sơn Hải (2007), *Nghiên cứu phương pháp kết hợp các đặc trưng màu sắc hình dạng và vị trí để truy vấn ảnh*, Luận văn Thạc sĩ, Trường Đại học Khoa học tự nhiên, Thành phố Hồ Chí Minh, trang 14.
- [3] Nguyễn Thị Hoàn (2010), *Phương pháp trích chọn đặc trưng ảnh trong thuật toán học máy tìm kiếm ảnh áp dụng vào bài toán tìm kiếm sản phẩm*, Luận văn thạc sĩ, Trường Đại học công nghệ, Đại học Quốc gia Hà Nội.
- [4] Trần Phước Long, Nguyễn Văn Lượng (2003), *Nhận dạng người dựa vào thông tin khuôn mặt xuất hiện trên ảnh*, Đồ án tốt nghiệp, Trường Đại học Khoa học tự nhiên Thành phố Hồ Chí Minh.
- [5] Nguyễn Đức Năng (2013), *Nghiên cứu và xây dựng bộ lọc ảnh thông qua phân loại ảnh kết hợp với gom cụm*, Luận văn Thạc sĩ, Trường Đại học Lạc Hồng.

Tài liệu Tiếng Anh

- [6] Qian Du (2006), *Unsupervised real-time constrained linear discriminant analysis to hyperspectral image classification*, Department of Electrical and Computer Engineering, Mississippi State University.
- [7] Svetlana Lazebnik, Cordelia Schmid, Jean Ponce (2006), *Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories*, University of Illinois.
- [8] David G. Lowe (Presented by David Lee 3/20/2006), *Object Recognition from Local Scale-Invariant Features*, Department Of Computer Science, University of British Columbia.
- [9] Xuezheng Liu, Lei Zhang, Mingjing Li, Hongjiang Zhang, Dingxing Wang (2005), *Boosting Image Classification with LDA-based Feature Combination*

for Digital Photograph Management, Department of Computer Science and Technology, Tsinghua University, Beijing, China

- [10] Android Application Programming with OpenCV - Published by Packt Publishing Ltd. Livery, Place 35 Livery Street , Birmingham B3 2PB, UK. ISBN 978-1 – 84969 – 520 -6
- [11] Pouladzadeh, Parisa, Shervin Shirmohammadi, and Rana Al-Maghrabi (2014), *Measuring calorie and nutrition from food image*, Instrumentation and Measurement, IEEE Transactions on 63, no. 8 (2014): 1947-1956.
- [12] Kawano, Yoshihiro, and Katsuki Yanai (2013), *Real-time mobile food recognition system*, In Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 1-7.
- [13] Xin, ZHENG Zhaobao, ZHANG Haitao, YE Zhiwei (2007), *Texture classification of aerial image based on bayesian network augmented naive bayes*, China School of Computer Science, Hubei University of Technology, Wuhan 430068, China
- [14] Yang, Shulin, Mei Chen, Dean Pomerleau, and Rahul Sukthankar (2010), *Food recognition using statistics of pairwise local features*, In Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, pp. 2249-2256.

Web tham khảo

- [15] Bảng thông tin về dinh dưỡng các món ăn Việt Nam :
<http://www.webthehinh.com/community/threads/48570/>, truy cập 5/2016
- [16] Giới thiệu về mô hình SVM
<http://www.stdio.vn/articles/read/436/gioi-thieu-ve-mo-hinh-svm>, truy cập 4/2016
- [17] Máy học Vector
https://en.wikipedia.org/wiki/Support_vector_machine, truy cập 4/2016
- [18] Mô hình Bag-of-features
https://rangers.googlecode.com/files/bof_classification.pdf, truy cập 4/2016

- [19] <http://healthplus.vn/viet-nam-dung-dau-ve-toc-do-tang-benh-dai-thao-duong-d10960.html>, truy cập 4/2016