



DOI:10.22144/ctujos.2023.193

XÁC ĐỊNH MÓN ĂN ĐẶC SẢN VIỆT NAM VỚI SỰ KẾT HỢP CỦA MẠNG HỌC SÂU VÀ BẢN THỂ HỌC

Mã Trường Thành^{1*}, Châu Ngân Khánh², Thạch Minh Hón³, Phạm Xuân Hiền¹ và Phan Bích Chung¹

¹Khoa Khoa học máy tính, Trường Công nghệ thông tin & Truyền thông, Trường Đại học Cần Thơ

²Khoa Công nghệ thông tin, Trường Đại học An Giang, Đại học Quốc gia Thành phố Hồ Chí Minh

³Trường Trung học phổ thông Trần Văn Bảy, Sóc Trăng

*Tác giả liên hệ (Corresponding author): mtthanh@ctu.edu.vn

Thông tin chung (Article Information)

Nhận bài (Received): 15/08/2023

Sửa bài (Revised): 18/08/2023

Duyệt đăng (Accepted): 21/08/2023

Title: Identification of Vietnamese Specialty Dishes via combining Deep Learning and EL Ontology

Author(s): Ma Truong Thanh^{1*}, Chau Ngan Khanh², Thach Minh Hon³, Pham Xuan Hien¹ and Phan Bich Chung¹

Affiliation(s): ¹Can Tho University; ²Vietnam National University Ho Chi Minh City; ³Tran Van Bay High School, Soc Trang

TÓM TẮT

Việc bảo tồn và phổ biến các giá trị của truyền thống văn hóa ẩm thực luôn là một thách thức không ngừng và cần được giữ gìn của mỗi quốc gia. Nhìn chung, có rất ít nghiên cứu và ứng dụng trí tuệ nhân tạo AI (Artificial intelligence) tập trung vào lĩnh vực này, hầu hết những nghiên cứu chỉ tập trung vào phân lớp hình ảnh và thiếu thông tin toàn diện của từng món ăn. Nhận thấy được tầm quan trọng của kiến thức cốt lõi và thông tin chi tiết của từng món ăn, đề xuất về khung AI cho các món ăn Việt Nam được giới thiệu trong bài báo này. Cụ thể, một bản thể luận (ontology) món ăn đặc sản Việt Nam để lưu trữ thông tin liên quan và mô hình phân lớp hình ảnh các món ăn đặc sản được trình bày. Đóng góp chính là phân lớp ảnh chụp món ăn với trên 96% và cung cấp các công thức nấu ăn tương ứng từ ontology.

Từ khóa: Ẩm thực, bản thể học, mạng học sâu, mô tả logic, trí tuệ nhân tạo

ABSTRACT

Conservation and dissemination of the cultural values inherent in culinary traditions constitute a persistent challenge requiring continuous preservation within each nation. There is a limited focus on research and AI (Artificial intelligence) applications in this domain, with most studies centering around image classification and lacking specific dish-related information. Realizing the significance of core knowledge and detailed insights into dishes that tourists seek, this article introduces an AI framework for Vietnamese cuisine. Specifically, it presents a specialized ontology for storing pertinent information regarding Vietnamese specialty dishes alongside a model for image classification. The essential contribution lies in achieving a dish image classification accuracy of over 96% and providing corresponding cooking recipes from the ontology.

Keywords: Artificial intelligence, convolutional neural network, description logic, ontology, specialty dishes

1. GIỚI THIỆU

Văn hóa ẩm thực, xét trên nhiều khía cạnh, thể hiện những khía cạnh kế thừa phi vật thể của lịch sử hình thành và phát triển của loài người. Vì vậy, văn hóa ẩm thực của mỗi quốc gia không chỉ thể hiện qua những món ăn giàu tính vật chất mà còn thể hiện qua cách người ta dùng bữa, cách sử dụng nguyên liệu, gia vị và cách trình bày. Các cá nhân có thể hiểu được đạo đức, quy định và phong tục của từng nhóm dân tộc và hành vi ăn uống của đất nước thông qua ẩm thực. Mỗi món ăn có thể đại diện cho đặc trưng của một vùng, ví dụ như Pâté de Canard của Hauts de France, hay một quốc gia, chẳng hạn như Pizza của Ý. Do đó, việc bảo tồn và giới thiệu những tinh hoa này đến các cộng đồng khác nhau, tức là khách du lịch là hoàn toàn có thể trông chờ vào việc khám phá và khai thác chúng. Một số nghiên cứu đã thu hút sự chú ý đến khía cạnh du lịch tiềm năng này (An & Moon, 2019; Kim et al., 2021). Cụ thể, An và Moon đã áp dụng kỹ thuật deep neural network để phân tích tâm lý dựa trên dữ liệu thời tiết và mùa, sau đó gợi ý địa điểm du lịch. Nhóm tác giả Kim et al. (2021) đã phát triển mô hình CNN (Convolutional neural network) để dự báo nhu cầu du lịch nhằm hỗ trợ quản lý du lịch. Không nằm ngoài vấn đề nổi bật này, Việt Nam là quốc gia có nhiều dân tộc anh em với nhiều nét khác biệt về văn hóa. Trong đó, văn hóa ẩm thực Việt (Trần & Nguyễn., 2006) luôn là “món ăn” tinh thần không thể thiếu đối với cộng đồng người Việt. Tính “đồng quê”, mùi hương đặc trưng, ít chất béo và hương vị đậm đà của món ăn tăng cường bởi nhiều loại gia vị sẽ là một vài trong số những đặc trưng độc đáo, thú vị, đáng chú ý của ẩm thực Việt Nam.

Ở một góc nhìn khác, đời sống văn hóa của các vùng miền khác nhau, ý nghĩa lịch sử và ẩm thực địa phương luôn khơi gợi sự tò mò và khao khát chinh phục của hầu hết du khách. Hương vị của mỗi vùng miền sẽ luôn mang lại sự thích thú cho du khách. Thật vậy, trong cùng một món ăn và tên gọi, mỗi cá nhân bản địa sẽ có những nét khác nhau trong cách nêm nếm và chế biến, chẳng hạn lẩu mắm, bún cá. Hơn nữa, bản sắc tinh thần của người Việt Nam khi nói đến ẩm thực là thể hiện sự trang nhã, cách ứng xử trong bữa ăn và khả năng lấy lòng nhau bằng phép lịch sự. Đương nhiên, mỗi cá nhân, gia đình, cộng đồng, xã hội đều có những chuẩn mực, phong tục tập quán ăn uống riêng. Do đó, truyền thống ẩm thực Việt Nam luôn đặt ra nhiều kỳ vọng được khai thác và khám phá. Đó là kho tàng di sản vô giá của dân tộc. Hơn nữa, việc giới thiệu và thu thập dữ liệu về văn hóa ẩm thực là bàn đạp để hội nhập thế giới

thông qua văn hóa ẩm thực, đồng thời thúc đẩy phát triển kinh tế và du lịch. Tận dụng trí tuệ nhân tạo (AI), nhiều nghiên cứu được triển khai với học máy để phân lớp, nhận biết và đánh giá thực phẩm (Do et al., 2019; Schroeder et al., 2019; Shen et al., 2020; Tai et al., 2022). Cụ thể, Schroeder et al. (2019) đã thiết kế một hệ thống dựa trên các cảm biến hóa học trên ống nano carbon kết hợp với k-hàng xóm gần nhất và thuật toán rừng ngẫu nhiên để phân lớp một số loại thực phẩm dựa trên mùi của chúng. Shen et al. (2020) đã phát triển một hệ thống ước tính thuộc tính thực phẩm bằng cách sử dụng mô hình mạng thần kinh tích chập để giúp duy trì chế độ ăn uống cân bằng. Tai et al. (2022) đã giới thiệu một ứng dụng di động dựa trên kỹ thuật học chuyên giao để phân lớp thực phẩm Việt Nam.

Nhìn chung, ẩm thực truyền thống là một trong những đặc điểm nổi bật được du khách và các nhà nghiên cứu quan tâm tìm hiểu, khám phá. Thông thường, các nhà khoa học tập trung vào việc xác định, đánh giá và truy xuất hình ảnh dựa trên các đặc điểm hình ảnh của chúng (He et al., 2013; Wang et al., 2021). He et al. (2013) đã phân tích hình ảnh thực phẩm để xác định thành phần dinh dưỡng của thực phẩm thông qua hệ thống đánh giá chế độ ăn uống. Sau đó, để đánh giá dữ liệu về thực phẩm, chẳng hạn như công thức nấu ăn và nguyên liệu, Wang et al. (2021) đã tạo ra các mạng dựa trên sự chú ý và nhất quán về ngữ nghĩa. Ngoài ra, một số quốc gia đã nhấn mạnh cách tiếp cận này để khai thác, Giovany et al. đã áp dụng các tính năng SIFT (Lowe, 2006) kết hợp với kỹ thuật máy học để xác định các món ăn của Indonesia.

Mặt khác, ẩm thực Việt Nam vẫn còn rất nhiều điều để khám phá vì nó luôn đa dạng và nhiều màu sắc trên khắp các vùng miền của đất nước (Ung et al., 2020; Tran et al., 2021; Tai et al., 2022). Hơn nữa, một số nhà nghiên cứu đã áp dụng các kỹ thuật AI cho thực phẩm trong du lịch. Vajirakachorn et al. (2017) và Razail et al. (2021) đã nghiên cứu về ẩm thực địa phương để phân tích và thúc đẩy sự phát triển của ngành du lịch. Nhìn chung, những nghiên cứu này về cơ bản tập trung vào việc phân lớp và nhận biết thực phẩm, đặc biệt, việc áp dụng một trong các cách tiếp cận sẽ không bao hàm các thông tin liên quan. Nó sẽ làm mất đi dữ liệu hữu ích và hấp dẫn thu hút khách du lịch. Một tình huống gợi ý rằng một khách du lịch nên đến một địa điểm và thử đồ ăn ở khu vực lân cận (không có dữ liệu về địa điểm hiện tại). Hình ảnh thực từ máy ảnh điện thoại thông minh của họ sau đó sẽ là thông tin chính. Một kết quả dự đoán là một số kiến thức về ẩm thực khu vực. Khi sử dụng các kỹ thuật máy học, chỉ có một

số phân lớp nhất định (chẳng hạn như tên món ăn) là kết quả. Vì vậy, kết hợp một số kỹ thuật AI mới để giải quyết vấn đề này sẽ được đề cập trong bài báo này.

Ngoài mong muốn tìm hiểu hương vị, tên gọi đặc trưng của các món ăn đặc sản vùng miền, du khách cũng tò mò về cách sơ chế, hướng dẫn (cách nấu), gia vị nêm nếm và các thông tin liên quan khác. Nhìn chung, hiện nay chưa có nghiên cứu nào ở nước ta hỗ trợ các món ăn đặc sản khác nhau của Việt Nam trong lĩnh vực du lịch. Theo đó, các lợi ích của ontology là nâng cao khả năng sử dụng, khả năng tái sử dụng và khả năng bảo trì của hệ thống thông tin. Hơn nữa, nó cung cấp một nền tảng phù hợp và đáng tin cậy để bảo quản dữ liệu. Do đó, việc sử dụng các bản thể luận để lưu trữ các khái niệm và sự kiện thiết yếu về ẩm thực đặc sản Việt Nam là cần thiết và khả thi. Đối với nền tảng, mạng tích chập là một trong những kỹ thuật học máy nổi tiếng nhất để xử lý các vấn đề phân lớp hình ảnh. Chất lượng đặc biệt của nó là tạo ra các dự báo động phức tạp. Trong những thập kỷ gần đây, nhiều nghiên cứu khác nhau về phân lớp hình ảnh đã được thực hiện bằng cách sử dụng học sâu (Jin et al., 1986; Xue et al., 2020). Cụ thể, Jin et al. (1986) đã phát triển một biểu đồ nhằm lần để xác định sự nhầm lẫn của hình ảnh có độ tương tự trực quan cao và cải thiện độ chính xác của việc phân lớp hình ảnh. Xue et al. (2020) đã đề xuất mô hình RestoreNet để cải thiện kỹ thuật học một lần để phân lớp hình ảnh. Ở đây, việc tận dụng phương pháp này để nhận ra việc thu thập thực phẩm và điều tra kiến thức về ẩm thực sẽ được vận dụng.

Đóng góp chính của bài báo này bao gồm:

- (1) Mô hình phân lớp đặc sản ẩm thực Việt Nam ở đồng bằng sông Cửu Long phục vụ du lịch;
- (2) Bản thể luận đặc sản ẩm thực Việt Nam;
- (3) Bộ dữ liệu hình ảnh ẩm thực Việt Nam được thu thập;
- (4) Ứng dụng AI để khám phá ẩm thực ở đồng bằng sông Cửu Long.

Lưu ý rằng, một bản thể luận nhẹ (về các món ăn truyền thống của Việt Nam) được phân lớp theo vùng và tỉnh. Việc triển khai bản thể luận được mã hóa bằng logic mô tả EL. Ở đây, lý do lựa chọn cách cài đặt này vì nó khá phù hợp với ứng dụng thực tế (độ phức tạp tính theo thời gian đa thức). Hơn nữa, sẽ tận dụng một mạng tích chập sâu để phát hiện và nhận ra tên món ăn. Để tăng độ chính xác của phân lớp, lý do lựa chọn ba món ăn có độ chính xác tốt

nhất để tiến hành lựa chọn duy nhất thông qua bản thể học. Cụ thể, tận dụng khả năng lập luận (lý luận) của bản thể học để đưa ra quyết định với một số thông tin bổ sung hữu ích, tức là vị trí của món ăn và mô tả các đặc điểm của nó. Cuối cùng, một truy vấn thông tin bằng SPARQL để có tên thực phẩm cuối cùng được thực hiện. Sau khi có tên món ăn, hệ thống cung cấp thêm thông tin chi tiết về nhà cung cấp, cách chế biến, video nấu ăn và nguyên liệu. Ngoài ra, hệ thống còn cung cấp ứng dụng chatbot để hỗ trợ khách du lịch tìm kiếm thông tin. Để xây dựng ứng dụng này, mô hình BERT (Sakata et al., 2019) và mô hình PhoBERT (Nguyen et al., 2020) được sử dụng. Một điểm cần lưu ý là nghiên cứu này tập trung chủ yếu vào việc đề xuất các ứng dụng thực tế (một khung) có thể áp dụng tốt cho du lịch và bảo tồn văn hóa truyền thống ở Việt Nam. Vì vậy, việc cải tiến thuật toán để nhìn thẳng vào việc tìm ra một mô hình phù hợp cho bài toán cơ bản của món ăn Việt Nam được tập trung nghiên cứu.

2. CƠ SỞ LÝ THUYẾT

2.1. Món ăn Việt

Ẩm thực là một phần không thể thiếu trong văn hóa của mỗi quốc gia. Từ bao đời nay, văn hóa ẩm thực Việt Nam đã được hình thành, gìn giữ và phát triển cùng với lịch sử dân tộc. Việt Nam là một quốc gia gồm nhiều dân tộc khác nhau, mỗi dân tộc đều có những nét văn hóa riêng biệt. Vì vậy, món ăn Việt Nam cũng mang những nét đặc trưng của đa dân tộc từ cách chế biến, sử dụng gia vị hay thói quen ăn uống, từ đó góp phần tạo nên bức tranh ẩm thực đa sắc màu (Dinh, 2013; Peters, 2021). Nhìn chung, văn hóa ẩm thực Việt Nam có những đặc điểm cơ bản sau:

Tính hòa đồng và đa dạng: Đặc điểm này thể hiện ở chỗ người Việt Nam sẵn sàng tiếp thu những tinh hoa, giá trị tốt đẹp của ẩm thực và văn hóa ẩm thực của các dân tộc khác. Tuy nhiên, đây là sự tiếp thu có chọn lọc thận trọng, chỉ vận dụng những đặc điểm phù hợp với lối sống, thuần phong mỹ tục của dân tộc.

Ít dầu mỡ: So với ẩm thực của các nước phương Đông như Trung Quốc và các nước Âu Mỹ, món ăn Việt Nam ít dầu mỡ.

Đậm đà về hương vị: Bằng cách sử dụng đa dạng các loại gia vị, các món ăn Việt Nam trở nên phong phú và hấp dẫn hơn. Trước khi chế biến, các nguyên liệu được tẩm ướp kỹ càng. Ngoài ra, mỗi món ăn đều có nước chấm riêng để tăng thêm độ ngon.

Tổng hợp nhiều chất, nhiều mùi: Ẩm thực Việt Nam được làm từ nhiều nguyên liệu khác nhau, từ

thịt, tôm, cua, cá, rau. Ngoài ra, trong mỗi món ăn không chỉ có một vị mà còn có một hương, là sự kết hợp của nhiều hương vị như chua, cay, mặn, ngọt, béo... Nhờ đó, món ăn Việt Nam đóng vai trò quan trọng trong việc phát triển văn hóa của dân tộc cũng như tạo động lực thúc đẩy tăng trưởng du lịch. Để tỏ chức bản thể học của các món ăn, khía cạnh vùng miền và các đặc trưng riêng của món ăn sẽ được sử dụng để phát triển ontology. Lưu ý, bài báo này sẽ chọn lọc, phân lớp và truy vấn dữ liệu về các món ăn độc đáo ở Tây Nam Bộ Việt Nam cho bài đăng này nhằm bảo tồn và phổ biến văn hóa thông qua các món ăn đặc sắc. Tổng quan về thí nghiệm đầu tiên các món ăn Việt Nam được thể hiện trong Hình 1.

2.2. Mạng học sâu

Một loại mạng học sâu (Deep neural network-DNN) (Heaton, 2021) được gọi là CNN hoặc ConvNet, vượt trội trong phân tích hình ảnh trực quan. Đây là một kỹ thuật học có giám sát kiểm tra dữ liệu bằng cách sử dụng perceptron, là các thành phần học máy. Hạt nhân hoặc bộ lọc là các thành phần cơ bản của CNN. Các hạt nhân sử dụng quá trình xử lý tích chập để trích xuất các đặc điểm chính xác và thích hợp từ đầu vào. Deep CNN được sử dụng trong nghiên cứu để phân lớp hình ảnh vì nó cực kỳ linh hoạt và hiệu quả trong việc xử lý dữ liệu hình ảnh.

2.3. Logic mô tả EL

EL là một loại DL (Description Logics), hoạt động trong hồ sơ Ngôn ngữ web Ontology và là một trong những cách quan trọng nhất để diễn đạt kiến thức thuật ngữ (Baader et al., 2005). Các cá nhân, khái niệm và vai trò là các thành phần chính của DL, tương ứng với các đối tượng và quan hệ giữa các đối tượng ở cấp độ ngữ nghĩa. Về mặt hình thức, chúng ta có N_I, N_R, N_C bao gồm ba bộ cặp rời rạc, với N_I đại diện cho đối tượng riêng lẻ, N_R đại diện cho quan hệ nguyên tử và N_C đại diện cho các nguyên tắc nguyên tử. Các biểu thức khái niệm EL_{\perp} (Kriegel et al., 2020) được xây dựng theo ngữ pháp sau đây là trọng tâm nghiên cứu trong bài viết này.

$$C ::= T \mid \perp \mid N_C \mid C \sqcap D \mid \exists r.C \text{ với } r \in N_R.$$

Cho $C, D \in N_C, a, b \in N_I$ và $r \in N_R$. Hộp thuật ngữ hoặc TBox (ký hiệu là T) và Hộp nguyên tử hoặc ABox (ký hiệu là A) là hai nửa của bản thể luận EL, còn được gọi là cơ sở tri thức. TBox được tạo thành từ một tập hợp các tiên đề được gọi là Bao gồm khái niệm chung (GCI) của $C \sqsubseteq D$, biểu thị rằng C chi tiết hơn D hay nói cách khác C chi được hấp thụ bởi D và $C \sqcap D \sqsubseteq \perp$, chứng tỏ C và D là những khái niệm rời rạc. Các tuyên bố liên quan đến các đối tượng cụ thể ở dạng $C(a)$ hoặc $r(a, b)$ bao

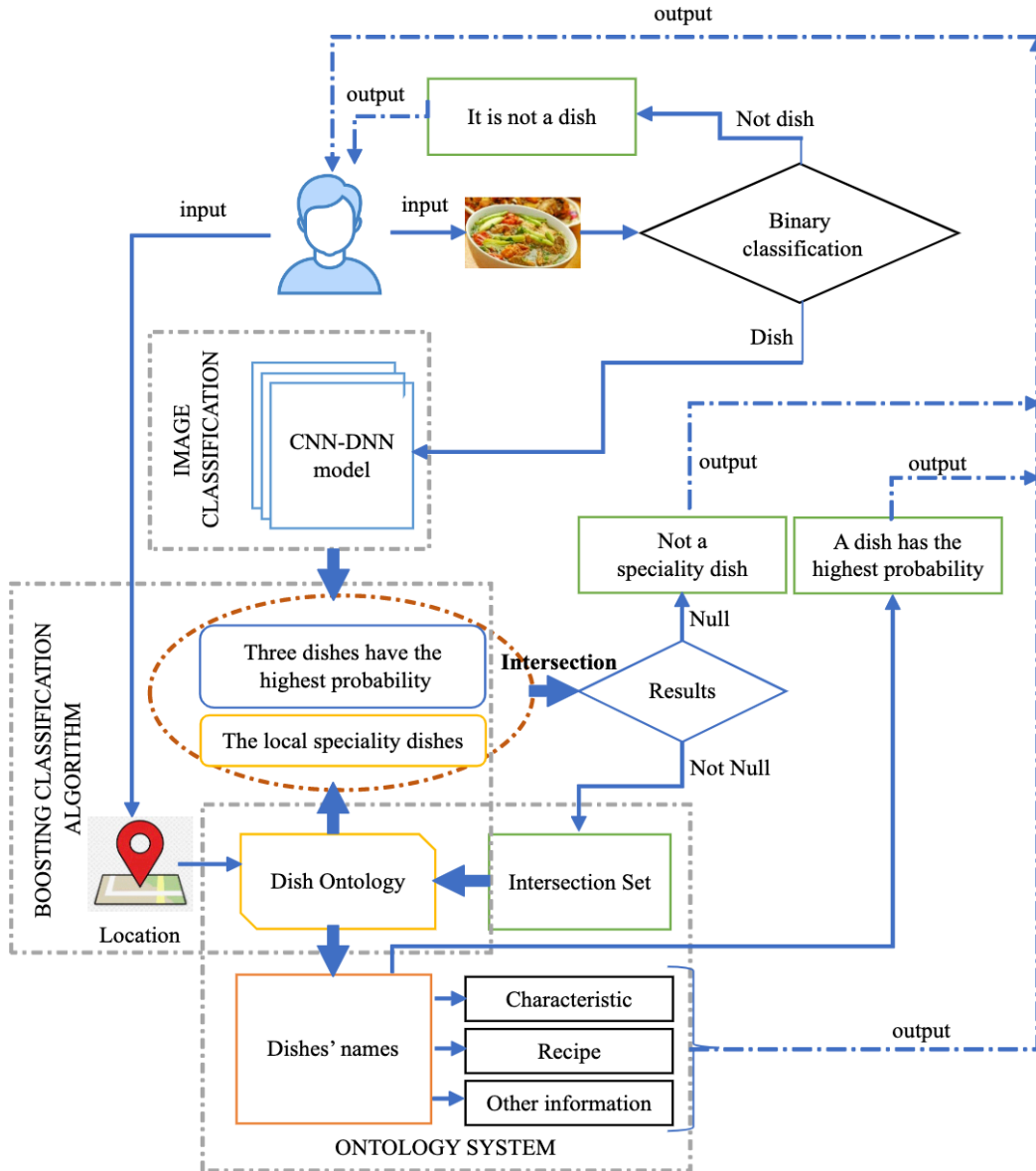
gồm ABox, là một tập hợp hữu hạn. Ngữ nghĩa được thể hiện dưới dạng $I = (\Delta^I, \cdot^I)$, bao gồm một miền diễn giải không rỗng Δ^I và một hàm diễn giải \cdot^I ánh xạ một cá thể $a \in N_I$ thành một phần tử $a^I \in \Delta^I$, một khái niệm $A \in N_C$ thành tập con $A^I \subseteq \Delta^I$ và vai trò $r \in N_R$ thành tập con $r^I \subseteq \Delta^I \times \Delta^I$.

Trong Bảng 1, một bản tóm tắt về cú pháp và ngữ nghĩa của EL_{\perp} được trình bày. Khi yêu cầu khớp ở cột bên phải được đáp ứng, một mô hình của (hoặc để đáp ứng) một tiên đề gọi là Φ ở dạng cột bên trái trong bảng, được đại diện bởi $I \models \Phi$. Chẳng hạn, khi $C^I \subseteq D^I, I \models C \sqsubseteq D$. Tương tự, khi $a^I \in C^I$ (tương ứng $(a^I, b^I) \in r^I$), thì I thỏa mãn khẳng định khái niệm đã cho là $I \models C(a)$ (tương ứng $I \models r(a, b)$). Nếu một diễn giải I đáp ứng tất cả các tiên đề và khẳng định trong một ontology O thì nó là một mô hình của ontology đó. Nếu một ontology có một mô hình, nó được gọi là nhất quán. Nó không nhất quán nếu mô hình O không thỏa mãn tiên đề Φ . Nếu mỗi mô hình của O thỏa mãn tiên đề Φ thì bản thể luận đó được cho là kéo theo Φ , ký hiệu là $O \models \Phi$.

Bảng 1. Cú pháp và ngữ nghĩa của EL_{\perp}

Cú pháp	Ngữ nghĩa
$C \sqsubseteq D$	$C^I \subseteq D^I$
r	$r^I \subseteq \Delta^I \times \Delta^I$
a	$a^I \in \Delta^I$
$C \sqcap D$	$C^I \cap D^I$
T	Δ^I
\perp	\emptyset
$\exists r.C$	$\{x \in \Delta^I \mid \exists y \in \Delta^I \text{ s.t. } (x, y) \in r^I, y \in C^I\}$

Về mặt bản thể luận, C được gộp lại bởi D đối với một bản thể luận O iff $O \models C \sqsubseteq D$. Theo cách tương tự, chúng ta nói rằng a là một thể hiện của C đối với O nếu $O \models C(a)$. Khi mỗi tên khái niệm trong bản thể luận không trống trong I, nghĩa là với mỗi khái niệm $C_i \in N_C, \cdot^I(C_i) \neq \emptyset$, thì một diễn giải $I = (\Delta^I, \cdot^I)$ được cho là thỏa mãn. Phân lớp là hình thức lập luận cơ bản mà các bản thể học thuật ngữ. Nó đòi hỏi phải tính toán tất cả các giả định phụ được ngụ ý ($C \sqsubseteq D$) (và các giá trị tương đương ($C \equiv D$)) tồn tại giữa các khái niệm T và \perp , là những khối xây dựng cơ bản của các bản thể luận. Quá trình như vậy được trình bày chi tiết trong Baader et al., 2005 và trước tiên nó bao gồm việc sử dụng một loạt các quy tắc để chuyển đổi bản thể luận thành một dạng bình thường, sau đó sử dụng tập hợp các quy tắc suy luận (hoàn thành) để thực hiện quy trình lập luận phân lớp (Kriegel et al., 2020). Trong nghiên cứu này, giả định áp dụng các tiêu chí hoàn thành phân lớp cho các bản thể luận cung cấp ở dạng thông thường được xác định trước. Các ontology đầu vào được chuẩn hóa



Hình 1. Mô hình cho bảo tồn món ăn đặc trưng Việt Nam

3. PHƯƠNG PHÁP ĐỀ XUẤT

3.1. Mô hình bảo tồn món ăn Việt Nam

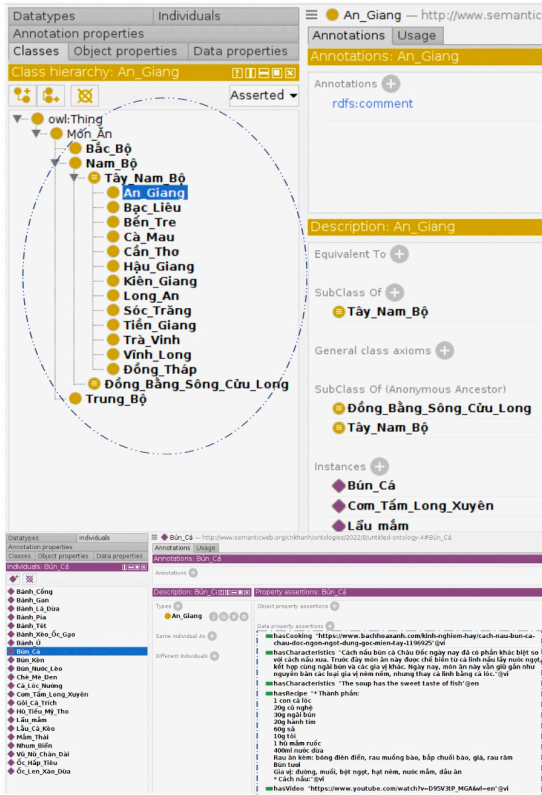
Một đề xuất về mô hình bảo tồn và phổ biến các món ăn Việt Nam được giới thiệu trong phần này (Hình 1). Nó hoạt động theo 3 bước: (1) phân lớp món ăn (phân lớp hình ảnh), (2) nhận dạng món ăn (thuật toán phân lớp tăng cường kết hợp với hệ thống ontology), (3) cung cấp thông tin liên quan (hệ thống ontology).

Lúc đầu, mô hình cần hai thông tin hình ảnh món ăn và vị trí làm đầu vào. Hình ảnh món ăn là ảnh chụp nhanh do người dùng chụp và vị trí được tự

động gửi đến hệ thống tại thời điểm chụp ảnh. Vai trò của bước đầu tiên là phân lớp hình ảnh món ăn (Hình 1, phân lớp hình ảnh). Nhưng một món ăn có thể nổi tiếng ở nhiều tỉnh và hương vị, cách chế biến... của một món ăn cũng khác nhau tùy theo tỉnh. Do đó, lựa chọn ba kết quả từ bước phân lớp, sau đó thực hiện kết hợp với thông tin vị trí để xác định tên món ăn (bước 2, Hình 1, thuật toán phân lớp tăng cường). Thông tin bổ sung (vị trí) giúp tăng độ chính xác của phân lớp. Cụ thể hơn, đây là hai bước đầu tiên: (1) trong suốt quá trình, CNN-DNN được huấn luyện ba món ăn có xác suất cao nhất được chọn; (2) vị trí hiện tại của người dùng và

những phát hiện này đóng vai trò là đầu vào hoặc bản thể luận món ăn Việt Nam được truy vấn và cung cấp tên món ăn làm đầu ra. Ở bước cuối cùng, với sự trợ giúp của hệ thống ontology, tên món ăn và thông tin liên quan được cung cấp cho người dùng dưới dạng đầu ra của khung (Hình 1, ontology).

3.2. Món ăn Việt với ontology EL₁



Hình 2. Ontology cho những món ăn Việt Nam

Bảng 2. Dữ liệu những món ăn đặc sản của Việt Nam

STT	Tên món ăn	Tỉnh/Thành	Số lượng ảnh thu thập
1	Bánh cồng	Cần Thơ, Sóc Trăng	214
2	Bánh gan	Cần Thơ	192
3	Bánh lá dứa	Trà Vinh	137
4	Bánh ống lá dứa	Kiên Giang	143
5	Bánh pía	Sóc Trăng	455
6	Bánh tét	Trà Vinh	444
7	Bánh ú	Vĩnh Long, Trà Vinh	305
8	Bánh xèo	Bạc Liêu, Bến Tre, Cần Thơ, Hậu Giang, Sóc Trăng, Vĩnh Long, An Giang, Kiên Giang	400
9	Bún cá	An Giang, Kiên Giang	316
10	Bún kèn	Kiên Giang	174
11	Bún nước lèo	Sóc Trăng	333

Để lưu trữ thông tin về món ăn, một thiết kế ontology món ăn Việt Nam để cung cấp các thông tin liên quan về món ăn Việt Nam được cài đặt.

Về ẩm thực, mỗi vùng/khu vực có nét đặc trưng riêng. Do đó, công thức và hương vị của một món ăn sẽ khác nhau tùy theo vùng miền (Hình 2). Từ điểm này, ontology được xây dựng, trong đó các công thức nấu ăn tương ứng với từng địa phương. Hơn nữa, ontology sẽ được cấu trúc dựa trên đặc điểm mỗi địa phương thuộc về một vùng cụ thể.

4. KẾT QUẢ THỰC NGHIỆM

4.1. Thu thập và tiền xử lý bộ dữ liệu

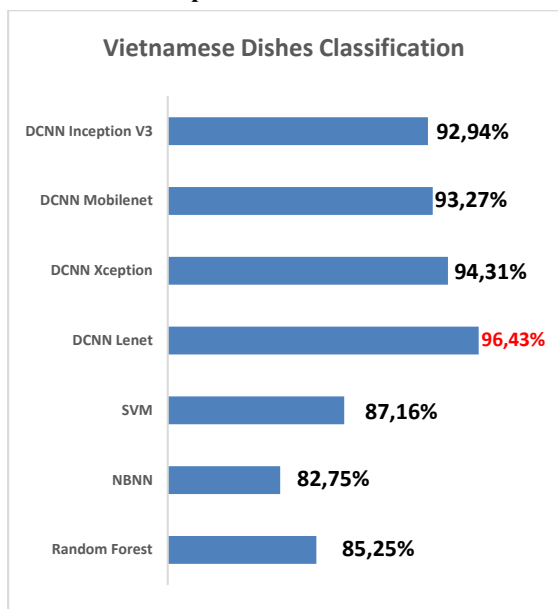
Bộ dữ liệu ảnh thu thập bao gồm 22 món ăn Việt Nam được biết đến ở các tỉnh Tây Nam Bộ (Bảng 2).

Các hình ảnh được thu thập từ internet bằng cách sử dụng cơ sở dữ liệu hình ảnh của Google trong bước đầu tiên. Sau đó, để loại bỏ lỗi trong ảnh thu thập được, một số bước sàng lọc đã được thực hiện, bao gồm chuẩn hóa ảnh và loại bỏ ảnh có lỗi. Hơn nữa, để xây dựng mô phân lớp, 5.331 hình ảnh của 22 món ăn ở các chế độ xem (nên) khác nhau được thu thập. Điều này đảm bảo rằng bộ sưu tập có thể bao phủ mọi góc nhìn.

Bảng 2 cho thấy một món ăn đặc sản có thể thuộc về nhiều địa phương. Những món ăn đặc sản này và nguyên liệu chính của chúng tương tự nhau ở các vùng khác nhau, nhưng tùy theo đặc điểm của từng vùng mà người nấu có thể thêm bớt một chút gia vị đặc trưng, làm cho văn hóa ẩm thực của Việt Nam thêm đa dạng và phong phú.

STT	Tên món ăn	Tỉnh/Thành	Số lượng ảnh thu thập
12	Cá lóc nướng	Cà Mau, Cần Thơ, Đồng Tháp	180
13	Chè mè đen	Sóc Trăng, Trà Vinh	188
14	Cơm tấm Long Xuyên	An Giang	322
15	Gỏi cá trích	Kiên Giang	169
16	Hủ tiếu Mỹ Tho	Tiền Giang	199
17	Lẩu cá kèo	Cần Thơ, Kiên Giang	248
18	Mắm thái	An Giang	148
19	Nhum biển	Kiên Giang	170
20	Ốc hấp tiêu	Cần Thơ	214
21	Ốc len xào dừa	Cà Mau, Hậu Giang	180
22	Vũ nữ chân dài	An Giang	200
s			5.331

4.2. Phân lớp món ăn



Hình 3. Độ chính xác của giải thuật phân lớp

Khi nói đến phân lớp hình ảnh, nó có thể bị ảnh hưởng bởi các yếu tố bao gồm độ sáng, nền và chất lượng hình ảnh. Các phương pháp gần đây có thể giúp giải quyết vấn đề này bao gồm DNN, CNN Support vector machines (SVM) (Vapnik, 1995; Hearst, 1998), Naive Bayes nearest neighbor (NBNN) (Lu & Weng, 2007; Boiman et al., 2008).

Các thuật toán phân lớp được so sánh ở Hình 3. Các kiểm tra đã được thực hiện trên máy Linux, Ubuntu 22.04.01 LTS, CPU Intel(R) Core(TM) i7-5500U @ 2.40GHz, RAM 16GB. Cụ thể, cơ sở dữ liệu thu thập được chia thành 4 phần, 3 phần dành cho huấn luyện (3.998 ảnh) và 1 phần để kiểm tra (1.333 ảnh).

Trong Hình 3, mạng nơ ron tích chập sâu đã cho thấy hiệu quả vượt trội của nó, đặc biệt là mạng nơ ron tích chập sâu theo kiến trúc LeNet (Hình 3) để giải quyết vấn đề đã đề cập. Việc huấn luyện và phân lớp hình ảnh thực phẩm được triển khai trong thư viện Keras, một thư viện mạnh mẽ cho các thuật toán học sâu (Chollet, 2017). Một huấn luyện và thực nghiệm bộ tham số cụ thể epochs 25, 50, 125, 150, 175 và 200 được thực hiện. Độ chính xác và thời gian huấn luyện tăng theo số lượng epoches. Độ chính xác tăng nhẹ sau 150 epoches, vì vậy 200 epoches được chọn để huấn luyện mô hình phân lớp. Đối với các giải thuật máy học, một số bộ tham số đã sử dụng như sau: giải thuật SVM (Vapnik, 1995) sử dụng bộ tham số $c=100000$, $\gamma=0,01$ với nhân RBF; giải thuật rừng ngẫu nhiên (random forest) xây dựng với 200 cây quyết định. Mặt khác, các đặc trưng sử dụng cho (Vapnik, 1995), NBNN (Boiman et al., 2008; Pham, 2017), rừng ngẫu nhiên là đặc trưng SIFT (Lowe, 2006). Lưu ý rằng, kết quả phân lớp (Hình 3) được tính trung bình của 10 lần thực nghiệm cho mỗi mô hình máy học. Phương pháp k-fold được sử dụng để đánh giá các giải thuật. Hầu hết các giải thuật đều sử dụng thư viện Keras và Tensorflow để cài đặt.

4.3. Xây dựng ontology món ăn

Thiết kế ontology món ăn bằng công cụ Protegé được thực hiện. Ontology nhẹ với nền tảng mô tả logic EL được xây dựng như Hình 2.

Ontology món ăn đã được xây dựng với 124 khái niệm và 565 tiên đề để lưu trữ thông tin món ăn bằng tiếng Việt và tiếng Anh. Ở đây, việc lưu trữ sử dụng “vi” để chỉ tiếng Việt và “en” để chỉ tiếng anh trong quá trình cài đặt cho các cá thể (individuals). Sau khi xác định tên món ăn được truy vấn, hệ thống sẽ hiển thị các thông tin liên quan như nguyên liệu chính, mùi vị và công thức nấu.

Nhìn chung, hệ thống hiệu quả với quá trình xử lý theo thời gian thực. Ngoài ra, độ chính xác của nhận dạng tên món ăn được tăng lên khi sử dụng thông tin bổ sung (vị trí).

5. THẢO LUẬN

Nghiên cứu này chủ yếu tập trung vào việc tìm kiếm thông tin về các món ăn đặc sản theo tỉnh và vùng miền ở Việt Nam. Kết quả nghiên cứu đạt hiệu quả như mong đợi với độ chính xác trên 95%. Hầu hết các kết quả của việc phân lớp và truy vấn thông tin món ăn đều phản ánh thông tin món ăn mà du khách mong muốn. Hơn nữa, một số giống cũng giống nhau về màu sắc, ví dụ: bánh đúc lá dứa và bánh lá dứa (cùng màu xanh), và hình dạng như bún cá và bún kèn, kết quả nhận dạng có thể gặp một số nhầm lẫn. Tuy nhiên, kiến thức bổ sung về lĩnh vực du khách cải thiện khả năng phân lớp món ăn và cung cấp kiến thức chính xác hơn. Mặt khác, số lượng hình ảnh của các loại thực phẩm rất sai lệch, tức là 143 ảnh bánh ống lá dứa và 444 ảnh bánh tét. Do đó, nó cũng ảnh hưởng đến kết quả phân lớp. Chúng ta có thể thu thập thêm hình ảnh cho một số loại thực phẩm để tìm sự cân bằng trong tương lai. Dự định trong tương lai là việc tích hợp một hệ thống thu thập hình ảnh để tích lũy hình ảnh từ khách du lịch. Lưu ý rằng hệ thống không tập trung chủ yếu vào kết quả phân lớp trong công việc này; hướng tiếp cận tập trung đưa sự chú ý đến khung (framework) để xây dựng hệ thống thực tế. Từ mô hình đã đề xuất, ta có thể thay đổi các mô hình huấn luyện để nâng cao độ chính xác. Tuy nhiên, việc cải thiện độ chính xác là công việc trong tương lai.

Nhìn chung, hiện nay có nhiều hệ thống nhận dạng khác nhau trong lĩnh vực này (Ung et al., 2020; Nguyen et al., 2021), tuy nhiên, hầu hết chúng không tích hợp với bản thể học để cải thiện kết quả phân lớp và khám phá thông tin rất cần thiết cho

TÀI LIỆU THAM KHẢO

An, H. W., & Moon, N. (2019). Design of recommendation system for tourist spot using sentiment analysis based on CNN-LSTM. *Journal of Ambient Intelligence and Humanized Computing*, 1-11. 10.1007/s12652-019-01521-w.

Kim, D. K., Shyn, S. K., Kim, D., Jang, S., & Kim, K. (2021, December). A daily tourism demand prediction framework based on multi-head attention CNN: The case of the foreign entrant in South Korea. *In 2021 IEEE SSCI* (pp. 01-10). IEEE.

Schroeder, V., Evans, E. D., Wu, Y. C. M., Voll, C. C. A., McDonald, B. R., Savagatrup, S., & Swager, T. M. (2019). Chemiresistive sensor

một món ăn, tức là nguyên liệu, video nấu ăn, nguồn gốc của món ăn... Thông thường, các tác giả tập trung phân lớp và nhận diện món ăn dựa trên đặc điểm trên ảnh thay vì khai thác các kiến thức liên quan về món ăn đó.

Một hạn chế của nghiên cứu này là danh sách các món ăn chỉ giới hạn ở các đặc sản của miền Tây Nam Bộ. Hơn nữa, khi số lượng món ăn tăng lên, độ chính xác của việc nhận dạng món ăn bị ảnh hưởng. Lý do là sự đa dạng của các món ăn cùng nhóm như bún cá, bún nước lèo. Ngoài ra, các món ăn có nguyên liệu chính giống nhau cũng là một nguyên nhân. Để khắc phục vấn đề này, cấu trúc ontology được vận dụng với nhiều thông tin được cung cấp hơn từ phía người dùng để tăng khả năng lọc từ các món ăn gọi ý sau khi phân lớp.

6. KẾT LUẬN

Đóng góp chính của nghiên cứu này là thiết lập một khung (framework), còn được gọi là mô hình, để bảo tồn và phổ biến các món ăn địa phương ở Tây Nam Bộ Việt Nam. Framework này sẽ giúp khách du lịch có thể tìm kiếm nhiều thông tin chi tiết của các món ăn. Cụ thể, bài báo đã giới thiệu bản thể luận của các món ăn Việt Nam ở khu vực Tây Nam Bộ. Ngoài ra, để phân lớp hình ảnh món ăn một cách hiệu quả, hệ thống đã được xây dựng và cài đặt một mô hình học sâu bằng cách sử dụng bộ dữ liệu hình ảnh món ăn Việt đã được thu thập; bên cạnh đó, cũng cung cấp các công thức nấu ăn của các món ăn liên quan. Hướng tiếp cận này sẽ là nền tảng cho việc phát triển các ứng dụng thực tế trên.

Nghiên cứu tiếp theo là sự mở rộng phạm vi điều tra đến khu vực đồng bằng sông Cửu Long ở Việt Nam; hơn nữa, sẽ phát triển một hệ thống toàn diện và đa ngôn ngữ với nhiều thông tin hơn về các món ăn truyền thống của Việt Nam.

array and machine learning classification of food. *ACS sensors*, 4(8), 2101-2108. DOI: 10.1021/acssensors.9b00825

Shen, Z., Shehzad, A., Chen, S., Sun, H., & Liu, J. (2020). Machine learning based approach on food recognition and nutrition estimation. *Procedia Computer Science*, 174, 448-453. DOI: 10.1016/j.procs.2020.06.113

Tai, T. T., Thanh, D. N. H., & Hung, N. Q. (2022). A dish recognition framework using transfer learning. *IEEE Access*, 10, 7793-7799. DOI: 10.1109/ACCESS.2022.3143119

He, Y., Xu, C., Khanna, N., Boushey, C. J., & Delp, E. J. (2013, July). Food image analysis:

- Segmentation, identification and weight estimation. In *2013 IEEE international conference on multimedia and expo* (pp. 1-6). IEEE. DOI: 10.1109/ICME.2013.6607548
- Wang, H., Sahoo, D., Liu, C., Shu, K., Achananuparp, P., Lim, E. P., & Hoi, S. C. (2021). Cross-modal food retrieval: learning a joint embedding of food images and recipes with semantic consistency and attention mechanism. *IEEE Transactions on Multimedia*, 24, 2515-2525. DOI: 10.1109/TMM.2021.3083109
- Giovany, S., Putra, A., Hariawan, A. S., & Wulandhari, L. A. (2017). Machine learning and SIFT approach for Indonesian food image recognition. *Procedia computer science*, 116, 612-620. DOI: 10.1016/j.procs.2017.10.020
- Tran, Q. L., Lam, G. H., Le, Q. N., Tran, T. H., & Do, T. H. (2021, July). A comparison of several approaches for image recognition used in food recommendation system. In *2021 COMNETSAT* (pp. 284-289). IEEE. DOI: 10.1109/COMNETSAT53002.2021.9530793
- Ung, H. T., Dang, T. X., Thai, P. V., Nguyen, T. T., & Nguyen, B. T. (2020). Vietnamese food recognition system using convolutional neural networks-based features. In *Computational Collective*, pp. 479-490. Springer International Publishing. DOI: 10.1007/978-3-030-63007-2_37
- Razali, M. N., Moug, E. G., Yahya, F., Hou, C. J., Hanapi, R., Mohamed, R., & Hashem, I. A. T. (2021). Indigenous food recognition model based on various convolutional neural network architectures for gastronomic tourism business analytics. *Information*, 12(8). DOI: 322. 10.3390/info12 080322
- Vajirakachorn, T., & Chongwatpol, J. (2017). Application of business intelligence in the tourism industry: A case study of a local food festival in Thailand. *Tourism Management Perspectives*, 23, 75-86.
- Jin, R., Dou, Y., Wang, Y., & Niu, X. (2017, August). Confusion Graph: Detecting Confusion Communities in Large Scale Image Classification. In *IJCAI* (pp. 1980-1986). DOI: 10.24963/ijcai.2017/275
- Xue, W., & Wang, W. (2020, April). One-shot image classification by learning to restore prototypes. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 34, No. 04, pp. 6558-6565). DOI: 10.1609/aaai.v34i04.6130
- Sakata, W., Shibata, T., Tanaka, R., & Kurohashi, S. (2019, July). FAQ retrieval using query-question similarity and BERT-based query-answer relevance. In *Proceedings of the 42nd International ACM SIGIR*. pp. 1113-1116. DOI: 10.1145/3331184.3331326
- Nguyen, D. Q., & Nguyen, A. T. (2020). PhoBERT: Pre-trained language models for Vietnamese. *arXiv preprint arXiv:2003.00744*.
- Peters, E. J. (2011). Appetites and aspirations in Vietnam: Food and drink in the long nineteenth century. *Rowman Altamira*.
- Heaton, J. (2020). Applications of deep neural networks. *arXiv preprint arXiv:2009.05673*.
- Baader, F., Brandt, S., & Lutz, C. (2005). Pushing the EL envelope (pp. 364-369). *Technische Universität Dresden*.
- Kriegel, F. (2020). Most specific consequences in the description logic EL. *Discrete Applied Mathematics*, 273, 172-204.
- Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, 28(5), 823-870. DOI: 10.1080/01431160600746456
- Chollet, F. (2021). *Deep learning with Python*. Simon and Schuster.
- Nguyen, T. T., Nguyen, T. Q., Vo, D., Nguyen, V., Ho, N., Vo, N. D., & Nguyen, K. (2021, August). VinaFood21: A novel dataset for evaluating Vietnamese food recognition. In *2021- RIVF* (pp. 1-6). IEEE.
- Lowe, G. (2004). Sift-the scale invariant feature transform. *Int. J*, 2(91-110), 2.
- Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., & Scholkopf, B. (1998). Support vector machines. *IEEE Intelligent Systems and their applications*, 13(4), 18-28.
- V. Vapnik.: *The Nature of Statistical Learning Theory*, Springer-Verlag, 1995.
- Boiman O, E. Shechtman and M. Irani. In Defense of Nearest-Neighbor Based Image Classification. *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition - CVPR*, 2008, pp. 1-8.
- Pham, N. K., Nguyen, M., & Do, T. N. (2017). *Điểm danh bằng mắt người với đặc trưng gist và máy học véc tơ hỗ trợ*. 10.15625/vap.2017.00019.
- Do, T. N., Pham, T. P., Pham, N. K., Nguyen, H. H., & Benferhat, S. (2019). Stacking of SVMs for classifying intangible cultural heritage images. *International Conference on Computer Science, Applied Mathematics and Applications*.
- Trần, Q. V., & Nguyễn, T. B. (2010). *Văn hoá ẩm thực Việt Nam nhìn từ lý luận và thực tiễn*. Từ điển Bách khoa & Viện văn hoá.