

HỆ THỐNG GỢI Ý SẢN PHẨM THỜI TRANG DỰA TRÊN HÌNH ẢNH TƯƠNG ĐỒNG FASHION PRODUCT RECOMMENDATION SYSTEM BASED ON SIMILAR IMAGES

NGUYỄN THÁI NGHE¹, NGUYỄN TRÚC NGÂN^{2,a}

¹Đại học Cần Thơ.

²Học viên cao học trường Đại học Sư phạm kỹ thuật Vĩnh Long.

^aTác giả liên hệ: nguyentrucngan142@gmail.com

Nhận bài (Received): 01/03/2024; Phản biện (Reviewed): 25/03/2024; Chấp nhận (Accepted): 01/04/2024

TÓM TẮT

Hiện nay, có rất nhiều công trình nghiên cứu về hệ thống gợi ý. Tuy nhiên, đa số chỉ tập trung vào nhóm các giải thuật gợi ý dựa vào nội dung văn bản thông qua các phương pháp lọc cộng tác, lọc nội dung mà chưa quan tâm nhiều đến việc gợi ý dựa vào hình ảnh. Trong nghiên cứu này, phương pháp được sử dụng là dựa trên hình ảnh tương đồng. Đề tài "Hệ thống gợi ý sản phẩm thời trang dựa trên hình ảnh tương đồng" đề xuất mô hình cụ thể để giải quyết bài toán đặt ra là dùng mạng nơ-ron tích chập để tìm kiếm những hình ảnh tương đồng. Tiếp theo là hệ thống được tích hợp vào website bán hàng thời trang để thực nghiệm, đưa ra những hình ảnh gợi ý một cách trực quan cho người dùng.

Từ khoá: Hệ thống gợi ý, hình ảnh tương đồng, mạng nơ-ron tích chập.

ABSTRACT

Currently, there are many research projects on recommendation systems. However, most only focus on the group of recommendation algorithms based on text content through collaborative filtering and content filtering methods without considering Pay more attention to image-based suggestions. In this study, the method used is based on image similarity. The topic "Fashion product recommendation system based on similar images" proposes a specific model to solve the problem of using convolutional neural networks to search for relevant images. Similar images. Next, the system is integrated into the fashion sales website for experimentation, providing visual suggestions for users.

Keywords: Recommendation system, similar images, convolutional neural networks

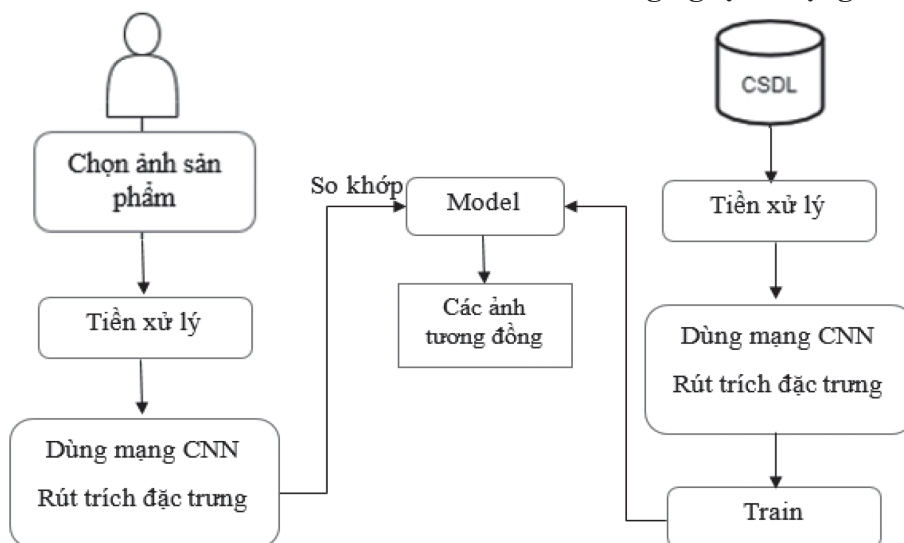
1. GIỚI THIỆU

Hệ thống gợi ý sản phẩm thời trang dựa trên hình ảnh tương đồng là một đề tài quan trọng và hữu ích trong lĩnh vực trí tuệ nhân tạo và công nghệ thông tin. Nó đóng vai trò trong việc cung cấp các gợi ý sản phẩm phù hợp với người dùng dựa trên hình ảnh mà họ cung cấp.

Đề tài được quan tâm từ nhu cầu ngày càng tăng của người dùng trong việc tìm kiếm sản phẩm trực tuyến. Thay vì phải mô tả sản phẩm bằng văn bản hoặc sử dụng từ khóa, người dùng có thể chụp hoặc tải lên một hình ảnh để tìm kiếm các sản phẩm tương tự hoặc liên quan.

Bogdan Walek cùng cộng sự đề xuất

phương pháp xây dựng hệ thống gợi ý lưu trữ kho dựa trên nội dung về các sản phẩm người dùng quan tâm thu thập được theo vị trí. Nghiên cứu này tập trung vào việc phát triển một hệ thống gợi ý lưu trữ kho. Hệ thống này có thể được thiết kế để cung cấp các gợi ý sản phẩm cho người dùng dựa trên các yếu tố như sở thích, ưa chuộng, hoặc các yếu tố khác liên quan đến sản phẩm. Ưu điểm của phương pháp này là cải thiện trải nghiệm người dùng và tăng cường khả năng tương tác. Cũng có thể đề cập đến các ứng dụng cụ thể của hệ thống gợi ý trong các bối cảnh như thương mại điện tử, dịch vụ địa điểm hoặc các lĩnh vực khác.



Hình 1. Sơ đồ tổng quan hệ thống gợi ý

Hệ thống được chia ra làm hai phần chính:

- *Mô hình máy chủ phục vụ nhận dạng ảnh:*

Ảnh từ tập dữ liệu đưa vào được tiền xử lý và dùng mạng học sâu CNN để rút trích ra những đặc trưng của ảnh tạo thành một model nhận dạng.

- *Mô hình nhận dạng và gợi ý:*

Sau khi hệ thống nhận được ảnh của người dùng, thì tiến hành tiền xử lý, rút

Nhóm tác giả Huynh-Ly Thanh-Nhan, Huu-Hoa Nguyen, Nguyen Thai-Nghe đề xuất một số phương pháp sử dụng hệ thống gợi ý để xây dựng khóa học thích hợp cho sinh viên. Sử dụng hệ thống gợi ý trong lĩnh vực giáo dục có thể giúp tạo ra trải nghiệm học tập cá nhân hóa cho sinh viên. Hệ thống có thể dựa trên dữ liệu về hoạt động học tập trước đó, sở thích cá nhân, và kết quả đánh giá để đề xuất các khóa học phù hợp. Kết quả giá trị trung bình $AUC=74,2\%$

2. PHƯƠNG PHÁP ĐỀ XUẤT

2.1. Công nghệ sử dụng

trích đặc trưng của ảnh. Các đặc trưng này được đưa vào model nhận dạng, so khớp để tìm và gợi ý các hình ảnh có đặc trưng gần giống với ảnh của người dùng đưa vào.

Đề tài được thực nghiệm trên nền Google colab phiên bản free, kết hợp mã nguồn mở Tensorflow – Keras do Google phát triển.

2.2. Tập dữ liệu

Đề tài gợi ý sản phẩm sử dụng tập dữ liệu DeepFashion bao gồm 44.441 hình

ảnh hàng may mặc được sưu tầm từ kaggle để kiểm thử. Tập dữ liệu có dung lượng 593MB. Tập dữ liệu gồm một thư mục ảnh gồm 44000 sản phẩm, 1 thư mục file json (chứa thông tin của các hình ảnh tương ứng) và tệp styles.csv.

2.3. Trích xuất đặc trưng

Giai đoạn 1: Huấn luyện mô hình mạng CNN được xây dựng từ mô hình mạng học chuyển giao (Transfer Learning Model) và tinh chỉnh (Fine-Tuning) mô hình huấn luyện. Quy trình huấn luyện mô hình sẽ thực hiện các bước sau: từ tập dữ liệu hình ảnh đầu vào, tiền xử lý tập hình ảnh, bộ đọc dữ liệu sẽ tiến hành đọc các tập hình ảnh để phân chia thành 2 tập dữ liệu phục vụ cho việc huấn luyện và đánh giá mô hình học sâu CNN là tập huấn luyện (train), tập kiểm thử (validation). Sau khi hoàn thành huấn luyện, mô hình sẽ được lưu tập tin ở định dạng H5 (một định dạng lưu trữ mảng hiệu quả).

Giai đoạn 2: Tất cả các ảnh trong tập hình ảnh dữ liệu được đưa qua mô hình mạng đã huấn luyện ở Giai đoạn 1 để trích xuất những đặc trưng véc-tơ của ảnh (những đặc trưng đó là hình dạng, màu sắc, kết cấu, ...). Các véc-tơ đại diện thu thập được tại lớp FC (fully connected) cuối cùng sẽ được lưu lại sử dụng trong quá trình tìm kiếm và gợi ý sau này.

2.4. Cấu hình các tham số

- Tốc độ học: trong khoảng [0.00001, 0.0002]
- Hàm tối ưu: Adam
- Hàm kích hoạt: Relu và Softmax
- Giá trị Drop-out: 0.1
- Hàm chi phí: cross-entropy
- Số vòng lặp: 20

Trong quá trình nghiên cứu, các tham số tốc độ học, hàm tối ưu, hàm kích hoạt, Drop out, hàm chi phí và số lần lặp đã được thay đổi. Tuy nhiên chúng tôi nhận thấy bộ tham số trên là cho kết quả cao nhất so với những bộ tham số còn lại. Cấu hình này được sử dụng để huấn luyện mạng nơ-ron, giải quyết một bài toán phân loại với các thiết lập nhất định để tối ưu hoá hiệu suất và tránh overfitting.

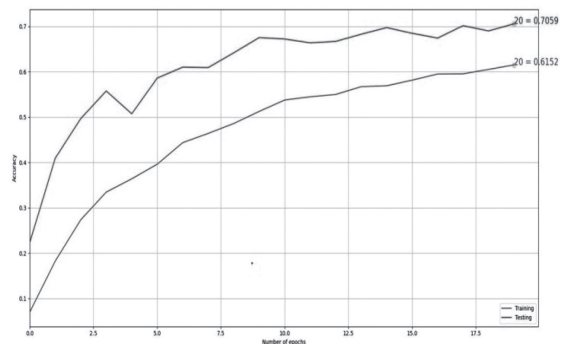
3. KẾT QUẢ THỰC NGHIỆM

3.1. Kết quả thực nghiệm với mô hình pre – trained VGG 16

Train accuracy	Val accuracy	Train loss	Val loss
0.6152	0.7059	1.9358	1.7253

Bảng 1 Kết quả huấn luyện pre-trained VGG16

Qua bảng kết quả cho thấy độ chính xác của mô hình còn thấp, chỉ đạt ở mức 71%



Biểu đồ 1 Biểu đồ trực quan huấn luyện pre-trained VGG16

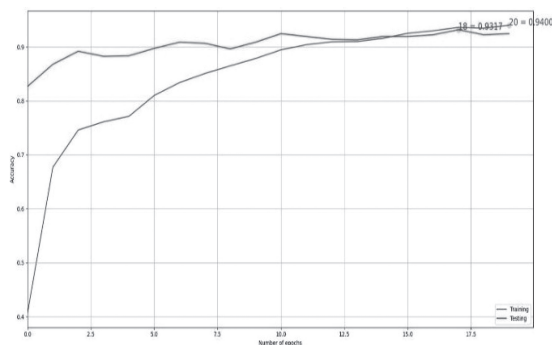
3.2. Kết quả thực nghiệm với mô hình pre – trained VGG 19

Train accuracy	Val accuracy	Train loss	Val loss
0.9400	0.9244	1.0392	1.0310

Bảng 1 Kết quả huấn luyện pre-trained VGG19

Qua bảng kết quả cho thấy độ chính

xác của mô hình tăng lên đáng kể so với mô hình VGG16. Giá trị Loss cũng giảm.



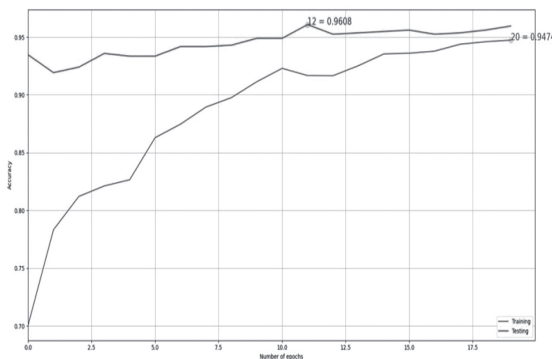
Biểu đồ 2 Biểu đồ trực quan huấn luyện pre-trained VGG19

3.3. Kết quả thực nghiệm với mô hình pre – trained Resnet

Train accuracy	Val accuracy	Train loss	Val loss
0.9474	0.9596	0.9743	0.8939

Bảng 2 Kết quả huấn luyện pre-trained VGG19

Qua bảng kết quả cho thấy độ chính xác của mô hình đạt 95%, cao hơn so với hai mô hình VGG16 và VGG19



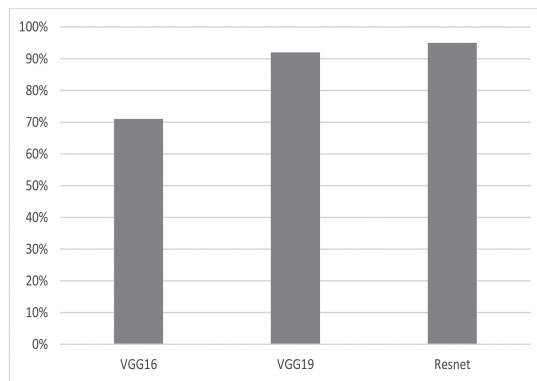
Biểu đồ 3 Biểu đồ trực quan huấn luyện pre-trained Resnet

3.4. So sánh kết quả của 3 mô hình

Mô hình	VGG16	VGG19	Resnet
Val accuracy	71%	92%	95%

Bảng 4: Bảng so sánh kết quả 3 mô hình

Qua bảng đánh giá các mô hình huấn luyện CNN cho thấy mô hình đề xuất Resnet đạt được Accuracy cao nhất và Loss thấp nhất so với các mô hình huấn luyện khác.



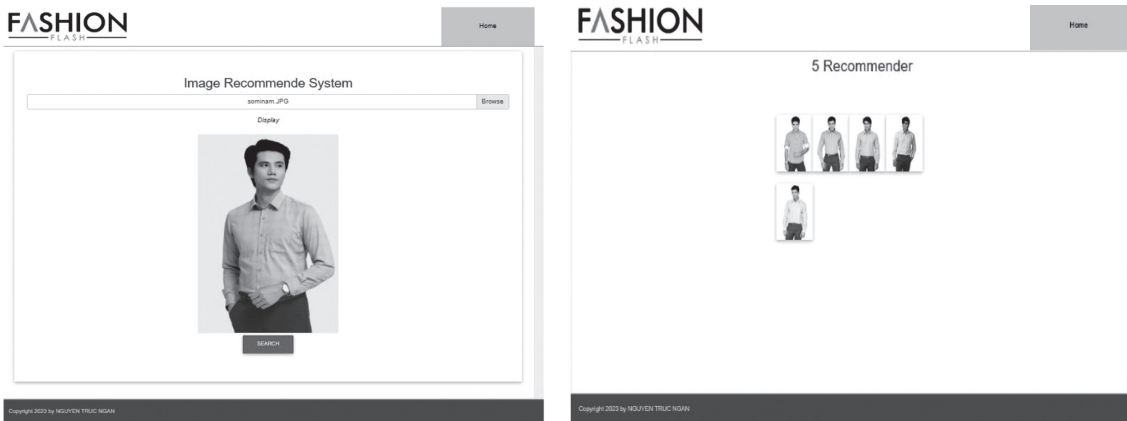
Biểu đồ 4 So sánh độ chính xác 3 mô hình

Từ kết quả so sánh độ chính xác của 3 mô hình trên, tôi quyết định chọn mô hình Resnet để xây dựng hệ thống gợi ý sản phẩm bằng hình ảnh.

Quy trình thực hiện gợi ý sản phẩm bằng hình ảnh được tiến hành như sau:

- Giai đoạn 1: Sử dụng thư viện Annoy (phương pháp Nearest Neighbor Search) để đánh chỉ mục (index) các véc-tơ đặc trưng ảnh trong cơ sở dữ liệu hình ảnh đã được trích xuất từ mô hình CNN trước đó.

- Giai đoạn 2: Khi người dùng chọn hình ảnh truy vấn đến hệ thống, hệ thống sẽ trích xuất véc-tơ đặc trưng từ ảnh truy vấn này sau đó so sánh độ tương đồng với các véc-tơ đặc trưng hình ảnh trong cơ sở dữ liệu hình ảnh bằng phép đo khoảng cách cosin. Tập 5 hình ảnh gần giống nhất với ảnh truy vấn được trả về theo khoảng cách được tính, trong đó hình ảnh gần giống nhất sẽ là ảnh có khoảng cách ngắn nhất đến ảnh truy vấn.



Hình 2. Hình ảnh demo tìm kiếm áo sơ mi

4. KẾT LUẬN

Trong nghiên cứu này, tôi đã trình bày phương pháp sử dụng một mạng nơ ron tích chập đã huấn luyện từ trước (pre-trained model) và tối ưu lại các tham số huấn luyện trên tập dữ liệu DeepFashion thu thập trên trang web Kaggle với tổng cộng 44.000 hình ảnh để trích xuất véc-tơ đặc trưng và so sánh độ tương đồng cho tìm kiếm ảnh truy vấn. Tôi áp dụng thuật toán CNN phân loại ảnh để trích xuất được đặc trưng của ảnh làm đại diện tốt nhất. Đồng thời áp dụng phương pháp “láng giềng gần nhất” (nearest neighbors) nhằm cải thiện tốc độ trả về kết quả tìm kiếm với thời gian truy vấn chấp nhận được. Tìm kiếm hình ảnh bằng phương pháp học sâu có ưu điểm là tận dụng được khả năng xử lý của mạng nơ ron tích chập cho cả thao tác phân loại và tính toán véc-tơ đại diện cho các ảnh trong tập ảnh tìm kiếm. Nghiên cứu này đã cho thấy việc áp dụng mạng nơ ron tích chập có kết quả tốt để góp phần nâng cao hiệu quả cho các hệ thống gợi ý sản phẩm bằng hình ảnh.

Việc huấn luyện trên mạng nơ ron tích

chập là một hoạt động tiêu tốn tài nguyên và thời gian, nghiên cứu này chưa thực hiện tối ưu một cách triệt để các tham số của mô hình. Do đó, các tham số của mạng nơ ron tích chập cũng cần được bổ sung và mở rộng khoảng giá trị khi thực hiện tối ưu nhằm tìm ra bộ tham số tốt nhất.

Ngoài ra, để có thể áp dụng cho một bài toán gợi ý sản phẩm theo ảnh có độ chính xác cao thì không thể chỉ áp dụng mạng CNN để trích xuất đặc trưng của hình ảnh với đầu ra của mạng là một vector đặc trưng vì những sản phẩm có kích thước khá giống nhau dẫn tới việc mô hình sẽ cho ra kết quả tìm kiếm dễ bị nhầm lẫn.

Hướng phát triển sắp tới, ngoài việc dùng vector embedding từ mô hình thì tôi sẽ nghiên cứu kết hợp thêm việc trích xuất đặc trưng về màu sắc trên từng khu vực của hình ảnh. Chúng tôi sẽ thu thập thêm dữ liệu huấn luyện mô hình với nhiều hình ảnh hơn để tăng độ chính xác của mô hình mạng nơ-ron học sâu. Và tiếp tục nghiên cứu thêm nhiều phương pháp học sâu khác để tìm mô hình có độ chính xác cao hơn.

TÀI LIỆU THAM KHẢO

- [1] Nguyễn Hùng Dũng, Nguyễn Thái Nghe. 2014. Hệ thống gợi ý sản phẩm trong bán hàng trực tuyến sử dụng kỹ thuật lọc cộng tác. Tạp chí Khoa học Trường Đại học Cần Thơ, số 31a (2014), trang 36-51. ISSN: 1859-2333.

- [2] Huỳnh Lý Thanh Nhân, Nguyễn Thái Nghe. 2013. Hệ thống dự đoán kết quả học tập và gợi ý lựa chọn môn học. Kỷ yếu hội thảo quốc gia lần thứ XVI: Một số vấn đề chọn lọc của CNTT&TT (@2013), trang 110-118. Nhà xuất bản Khoa học kỹ thuật. ISBN: 987- 604-67-0251-1.
- [3] Nguyễn Tấn Phong, Nguyễn Thái Nghe. 2014. Một giải pháp trong xây dựng Hệ thống gợi ý bài hát. Trang 149-154, kỷ yếu hội thảo quốc gia lần thứ XVII: Một số vấn đề chọn lọc của CNTT&TT (@2014). Nhà xuất bản Khoa học và kỹ thuật. ISBN: 978-604-67-0426-3
- [4] Lư Chân Thiện và Nguyễn Thái Nghe. 2015. MỘT TIẾP CẬN TRONG XÂY DỰNG HỆ THỐNG GỢI Ý THEO NGŨ CẢNH. Kỷ yếu Hội nghị quốc gia lần thứ VIII “Nghiên cứu cơ bản và ứng dụng Công nghệ thông tin” FAIR 2015.
- [5] Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, and Domonkos Tikk. “Session-based Recommendations with Recurrent Neural Networks.” Proceedings of the 4th International Conference on Learning Representations (ICLR), 2016.
- [6] Ruining He, Wang-Cheng Kang, and Julian McAuley. “Sequence-Aware Recommendation through Personalized Heterogeneous Embedding.” Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI), 2017.
- [7] Ludovic Denoyer and Thanh-Nam Hoang. “Context-Aware Session-Based Recommendations with Recurrent Neural Networks.” Proceedings of the 25th ACM International Conference on Information and Knowledge Management (CIKM), 2016.
- [8] Nguyễn Thái Nghe. “Chương 2. Hệ thống gợi ý: Kỹ thuật và ứng dụng - PDF Free Download.” Accessed: Jan. 01, 2024. [Online]. Available: <https://docplayer.net/71285438-Chuong-2-he-thong-goi-y-ky-thuat-va-ung-dung.html>
- [9] Nicola Fronzetti, “Predictive Neural Network Applications for Insurance Processes”, September 2019.
- [10] Fukushima, K.: Neocognitron, “A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position”, Biological Cybernetics 36(4), 193–202,1980.
- [11] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner, “Gradient-Based Learning Applied to Document Recognition”, IEEE 1998
- [12] Alex Krizhevsky, Ilya Sutskever, Geoffrey Hinton, University of Toronto Canada, “ImageNet Classification with Deep Convolutional Neural Networks”, Neural Information Processing Systems (NIPS), 2012.
- [13] Matthew D Zeiler, Rob Fergus, “Visualizing and Understanding Convolutional Networks”, arXiv:1311.2901v3 [cs.CV] 28 Nov 2013.
- [14] Karen Simonyan, Andrew Zisserman, “Very Deep Convolutional Networks For Large-Scale Image Recognition”, ICLR 2015.
- [15] “Giới thiệu mạng ResNet.” Accessed: Jan. 02, 2024. [Online]. Available: <https://viblo.asia/p/gioi-thieu-mang-resnet-vyDZOa7R5wj>