

Retrieval images with Convolutional Neural Networks

Nguyễn Tuấn Sinh, Nguyễn Tiên Sỹ và Dương Văn Sang

Industrial University of Ho Chi Minh City

Computer Vision Course
Ngày 28 tháng 10 năm 2022



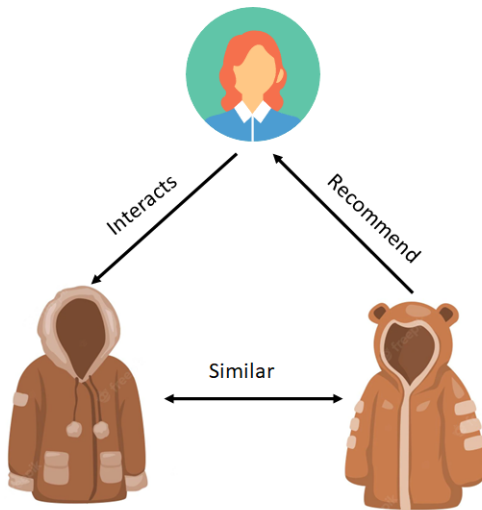
Nội dung

- 1 Giới thiệu
- 2 Phương pháp tiếp cận
 - Lớp CNN
 - Phân loại
 - SVM
 - AlexNet
 - VGG-16
 - Train mô hình
 - Test mô hình
 - Đánh giá mô hình
 - ResNet50
 - Đánh giá mô hình
 - Khuyến nghị
- 3 Dữ liệu và đặc trưng
- 4 Triển khai
 - Xử lý dữ liệu
 - Đánh giá
 - Phân lớp
- 5 Kết luận
- 6 Hướng nghiên cứu tương lai

Nội dung

- 1 Giới thiệu
- 2 Phương pháp tiếp cận
 - Lớp CNN
 - Phân loại
 - SVM
 - AlexNet
 - VGG-16
 - Train mô hình
 - Test mô hình
 - Đánh giá mô hình
 - ResNet50
 - Đánh giá mô hình
 - Khuyến nghị
- 3 Dữ liệu và đặc trưng
- 4 Triển khai
 - Xử lý dữ liệu
 - Đánh giá
 - Phân lớp
- 5 Kết luận
- 6 Hướng nghiên cứu tương lai

Giới thiệu

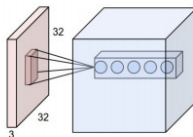


Hình: Khuyến nghị

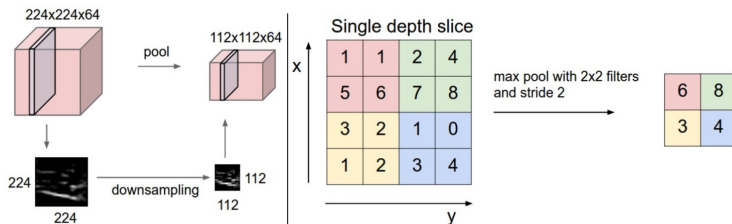
Nội dung

- 1 Giới thiệu
- 2 Phương pháp tiếp cận
 - Lớp CNN
 - Phân loại
 - SVM
 - AlexNet
 - VGG-16
 - Train mô hình
 - Test mô hình
 - Đánh giá mô hình
 - ResNet50
 - Đánh giá mô hình
 - Khuyến nghị
- 3 Dữ liệu và đặc trưng
- 4 Triển khai
 - Xử lý dữ liệu
 - Đánh giá
 - Phân lớp
- 5 Kết luận
- 6 Hướng nghiên cứu tương lai

Lớp CNN



Hình: Conv layer



Hình: Pooling layer

Phân loại

- SVM
- AlexNet
- VGG-16

SVM

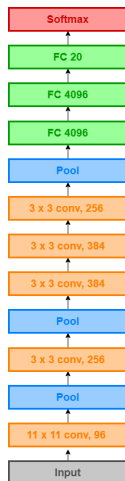
Phép biến đổi tuyến tính:

$$s = Wx_i + b$$

Hàm loss L_2 :

$$L_{SVM}(W, b; x_i) = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)$$

AlexNet



Hình: Mô hình AlexNet

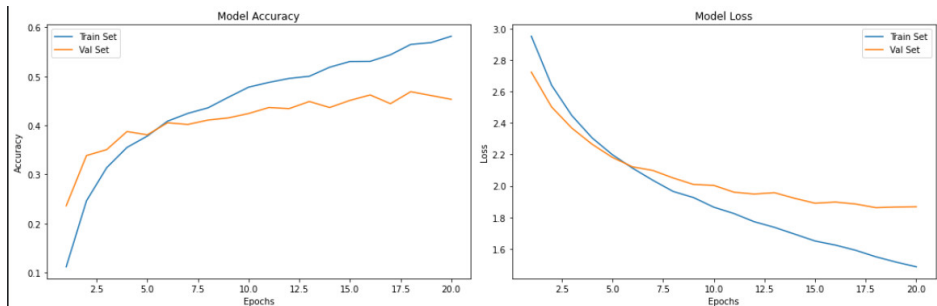
VGG-16



Hình: Mô hình VGG-16

Train mô hình

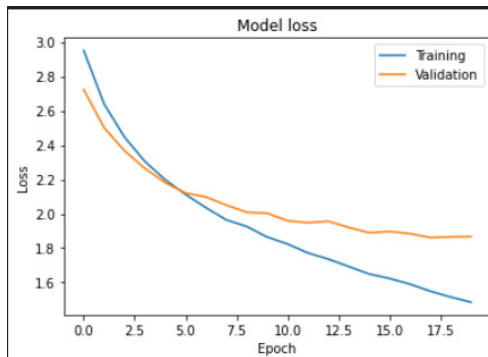
mô hình train VGG16



Hình: train mô hình phân loại VGG16

Test mô hình

mô hình test VGG16



Hình: test mô hình phân loại VGG16

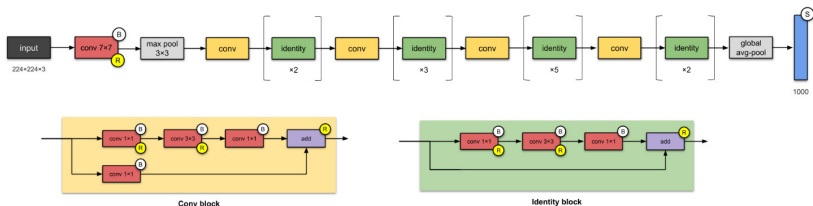
Đánh giá mô hình

Đánh giá mô hình phân loại VGG16

	precision	recall	f1-score	support
Pet Supplies	0.35	0.21	0.26	34
Grocery & Gourmet Food	0.50	0.64	0.56	28
Patio, Lawn & Garden	0.10	0.04	0.06	25
Cell Phones & Accessories	0.49	0.83	0.61	23
Beauty	0.10	0.09	0.09	23
Health & Personal Care	0.25	0.05	0.08	20
Clothing	0.96	0.93	0.94	27
Arts, Crafts & Sewing	0.76	0.76	0.76	25
Baby	0.53	0.31	0.39	26
Musical Instruments	0.45	0.68	0.54	19
Automotive	0.58	0.42	0.49	26
Jewelry	0.67	0.82	0.74	17
Home Improvement	0.20	0.23	0.22	30
Kitchen & Dining	0.44	0.77	0.56	22
Watches	0.42	0.45	0.43	22
Shoes	0.57	0.38	0.45	32
Home & Kitchen	0.51	0.74	0.60	34
Camera & Photo	0.35	0.21	0.26	29
Electronics	0.35	0.47	0.40	19
Computers & Accessories	0.26	0.33	0.29	18
accuracy			0.46	499
macro avg	0.44	0.47	0.44	499
weighted avg	0.45	0.46	0.44	499

Hình: Classification Report Metrics Interpretation

ResNet50



Hình: Mô hình Resnet

Đánh giá mô hình

Đánh giá mô hình phân loại ResNet50

	precision	recall	f1-score	support
Pet Supplies	0.43	0.22	0.29	27
Grocery & Gourmet Food	0.70	0.55	0.62	29
Patio, Lawn & Garden	0.46	0.19	0.27	32
Cell Phones & Accessories	0.71	0.83	0.76	29
Beauty	0.10	0.25	0.15	16
Health & Personal Care	0.26	0.28	0.27	25
Clothing	0.95	0.95	0.95	22
Arts, Crafts & Sewing	0.93	0.93	0.93	28
Baby	0.24	0.43	0.31	21
Musical Instruments	0.79	0.68	0.73	22
Automotive	0.68	0.65	0.67	23
Jewelry	0.85	0.63	0.72	27
Home_Improvement	0.50	0.33	0.40	30
Kitchen & Dining	0.65	0.69	0.67	16
Watches	0.65	0.38	0.48	29
Shoes	0.29	0.50	0.36	16
Home & Kitchen	0.67	0.74	0.70	27
Camera & Photo	0.24	0.60	0.34	20
Electronics	0.60	0.54	0.57	28
Computers & Accessories	0.25	0.07	0.11	28
accuracy			0.52	495
macro avg	0.55	0.52	0.51	495
weighted avg	0.56	0.52	0.52	495

Khuyến nghị

Các bước:

- Trích xuất đặc trưng.
- Đưa vào mô hình khuyến nghị.
- Tính toán độ tương đồng:

Cosine:

$$s_{\text{cosine}} = \frac{v_i^T v_j}{\|v_i\| \|v_j\|}$$



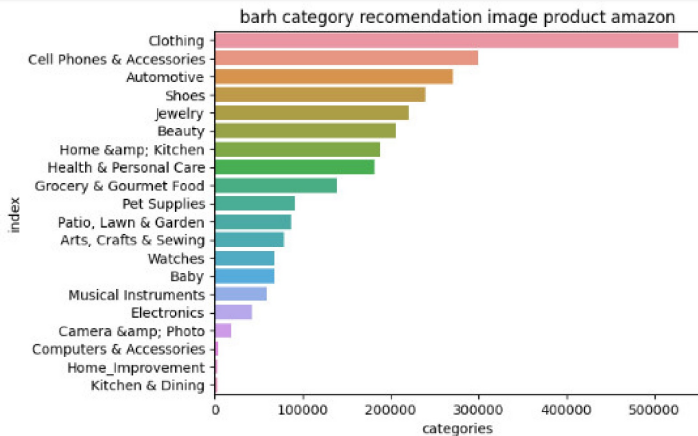
Hình: Ví dụ về kết quả hệ thống đề xuất của chúng tôi

Nội dung

- 1 Giới thiệu
- 2 Phương pháp tiếp cận
 - Lớp CNN
 - Phân loại
 - SVM
 - AlexNet
 - VGG-16
 - Train mô hình
 - Test mô hình
 - Đánh giá mô hình
 - ResNet50
 - Đánh giá mô hình
 - Khuyến nghị
- 3 **Dữ liệu và đặc trưng**
- 4 Triển khai
 - Xử lý dữ liệu
 - Đánh giá
 - Phân lớp
- 5 Kết luận
- 6 Hướng nghiên cứu tương lai

Dữ liệu và đặc trưng

Sử dụng bộ dữ liệu sản phẩm của Amazon



Hình: Phân phối nhãn của tập dữ liệu

Dữ liệu và đặc trưng

Bộ dữ liệu bao gồm:

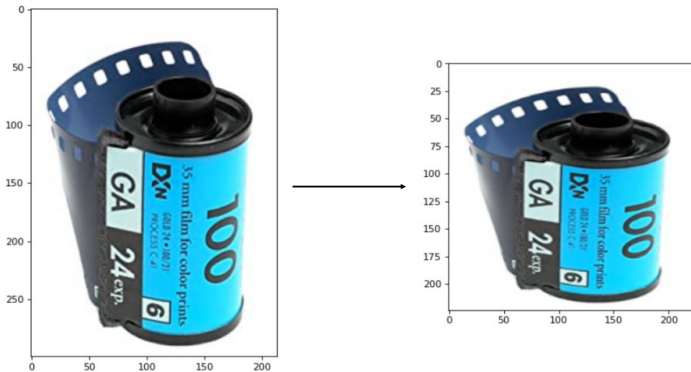
- **asin** - ID của sản phẩm
- **categories** - danh mục của sản phẩm
- **imUrl** - url hình ảnh sản phẩm
- **title** - tên sản phẩm

Nội dung

- 1 Giới thiệu
- 2 Phương pháp tiếp cận
 - Lớp CNN
 - Phân loại
 - SVM
 - AlexNet
 - VGG-16
 - Train mô hình
 - Test mô hình
 - Đánh giá mô hình
 - ResNet50
 - Đánh giá mô hình
 - Khuyến nghị
- 3 Dữ liệu và đặc trưng
- 4 **Triển khai**
 - Xử lý dữ liệu
 - Đánh giá
 - Phân lớp
- 5 Kết luận
- 6 Hướng nghiên cứu tương lai

Xử lý dữ liệu

Điều chỉnh kích thước ảnh cho phù hợp với đầu vào của mô hình



Đánh giá

- Phân lớp:

$$\text{Accuracy} = \frac{\text{\#hình ảnh được phân loại chính xác}}{\text{\#hình ảnh trong tập dữ liệu xác thực}}$$

Phân lớp

Model	Training accuracy	Validation accuracy	Test accuracy
SVM (baseline)	0.25	0.23	0.20
AlexNet	0.5340	0.3541	0.3406
VGG	0.9164	0.4388	0.4108
VGG vs BN	0.5688	0.4610	0.4448
ResNet	0.9911	0.5250	0.5201

Bảng: Nhóm thực hiện

Model	Training accuracy	Validation accuracy	Test accuracy
SVM (baseline)	0.2610	0.1807	0.2679
AlexNet	0.6484	0.4064	0.3946
VGG	0.8769	0.5110	0.5010

Bảng: Nhóm tác giả

Nội dung

- 1 Giới thiệu
- 2 Phương pháp tiếp cận
 - Lớp CNN
 - Phân loại
 - SVM
 - AlexNet
 - VGG-16
 - Train mô hình
 - Test mô hình
 - Đánh giá mô hình
 - ResNet50
 - Đánh giá mô hình
 - Khuyến nghị
- 3 Dữ liệu và đặc trưng
- 4 Triển khai
 - Xử lý dữ liệu
 - Đánh giá
 - Phân lớp
- 5 **Kết luận**
- 6 Hướng nghiên cứu tương lai

Kết luận

Trong dự án này, nhóm đã xây dựng một đề xuất mua sắm thông minh để tìm kiếm hình ảnh. So sánh thử các mạng nơ-ron khác nhau cho mô hình phân loại ảnh và các cách khác nhau để đánh giá sự giống nhau giữa hai ảnh. Chúng ta có thể đạt được độ chính xác phân loại là 0.46 và khuyến nghị sản phẩm có điểm tương đồng cao hơn 0.46. Vấn đề quá phù hợp, đây có thể là một trong những việc cần làm sau này.

Nội dung

- 1 Giới thiệu
- 2 Phương pháp tiếp cận
 - Lớp CNN
 - Phân loại
 - SVM
 - AlexNet
 - VGG-16
 - Train mô hình
 - Test mô hình
 - Đánh giá mô hình
 - ResNet50
 - Đánh giá mô hình
 - Khuyến nghị
- 3 Dữ liệu và đặc trưng
- 4 Triển khai
 - Xử lý dữ liệu
 - Đánh giá
 - Phân lớp
- 5 Kết luận
- 6 Hướng nghiên cứu tương lai

Hướng nghiên cứu tương lai

Trong bước tiếp theo, nhóm có thể cố gắng đào tạo mô hình của mình về lượng dữ liệu lớn hơn. Điều này có thể tăng độ chính xác của mô hình. Hiện tại chúng tôi chỉ sử dụng 20 danh mục. Tuy nhiên, các sản phẩm trong danh mục khác nhau rất nhiều, điều này giải thích độ chính xác thấp trong phân loại. Nhóm sẽ cố gắng tìm một thông tin danh mục cụ thể hơn và đào tạo mô hình.

Thank you!