

BÀI TẬP LỚN

Học phần: IT3160 - Lớp 147728

Nhập môn Trí tuệ nhân tạo



Đề tài: Khôi phục ảnh sử dụng mô hình SRCNN

Nhóm 22

Hoàng Ngọc Ánh - 20214992

Nguyễn Hữu Đức - 20210192

Nguyễn Hoàng Long - 20215081

Liễu Nhật Minh - 20215089

Nguyễn Hoàng Việt - 20215168

GVHD: Ph.D. Trần Thế Hùng

Trường Công nghệ thông tin và Truyền thông
Đại học Bách khoa Hà Nội

Tháng Sáu, 2024

Bảng 1: Nhiệm vụ thành viên & Đánh giá

Sinh viên	Nhiệm vụ	Đánh giá
Hoàng Ngọc Ánh - 20214992	Viết mã lọc nhiễu sử dụng NLM	Hoàn thành tốt (10)
Nguyễn Hữu Đức - 20210192	Quản lý tiến độ Chuẩn bị dữ liệu Viết báo cáo	Hoàn thành tốt (10)
Nguyễn Hoàng Long - 20215081	Đọc hiểu tài liệu tham khảo Xây dựng mô hình	Hoàn thành tốt (10)
Liễu Nhật Minh - 20215089	Đọc hiểu tài liệu tham khảo Huấn luyện mô hình	Hoàn thành tốt (10)
Nguyễn Hoàng Việt - 20215168	Viết mã dự đoán ảnh, đánh giá hiệu quả của mô hình	Hoàn thành tốt (10)

Tóm tắt: Báo cáo này trình bày kết quả áp dụng mô hình mạng nơ-ron tích chập siêu phân giải (SRCNN) cho phục hồi hình ảnh nhiễu. SRCNN là một phương pháp hiệu quả cho việc nâng cao chất lượng hình ảnh bị giảm độ phân giải hoặc nhiễu. Báo cáo mô tả quá trình xây dựng, huấn luyện và đánh giá hiệu quả của mô hình SRCNN trên tập dữ liệu LSDIR. Kết quả thu được cho thấy SRCNN có khả năng phục hồi hình ảnh nhiễu với độ chính xác cao và vượt trội so với các phương pháp truyền thống.

Mục lục

1 Giới thiệu	iii
1.1 Mô tả bài toán	iii
1.2 Các hàm đánh giá hiệu quả khôi phục ảnh	iii
1.3 Một số phương pháp thường dùng	iii
2 Cơ chế hoạt động của SRCNN [1]	iv
3 Xây dựng mô hình SRCNN	v
3.1 Triển khai và huấn luyện mô hình	v
3.2 Kết hợp mô hình CNN với bộ lọc NLM	vi
3.3 Kết quả thử nghiệm	vi
4 Mã nguồn	x
5 Kết luận	x

1 Giới thiệu

1.1 Mô tả bài toán

Cho ảnh X là ảnh có độ phân giải cao và Y là một phiên bản độ phân giải thấp của X . Gọi hàm được dùng để khôi phục chất lượng ảnh là F . Bài toán khôi phục ảnh tập trung vào việc tìm F sao cho $F(Y) \approx X$ hay ảnh khôi phục từ Y gần giống X nhất có thể.

1.2 Các hàm đánh giá hiệu quả khôi phục ảnh

Các hàm đánh giá hiệu quả khôi phục ảnh đóng vai trò quan trọng trong việc đánh giá chất lượng ảnh sau khi khôi phục. Một số hàm đánh giá phổ biến bao gồm SSIM (Structural Similarity Index Measure), MSE (Mean Squared Error) và PSNR (Peak Signal-to-Noise Ratio).

- **SSIM (Structural Similarity Index Measure)** Đo lường mức độ giống nhau về cấu trúc giữa hai ảnh, bao gồm độ sáng, độ tương phản và cấu trúc. Giá trị SSIM dao động từ 0 đến 1, với giá trị càng cao cho thấy mức độ giống nhau về cấu trúc càng lớn giữa hình ảnh gốc và hình ảnh được xử lý.
- **MSE (Mean Squared Error)** Đo lường mức độ sai số trung bình giữa các pixel tương ứng trong hai ảnh. Giá trị MSE càng nhỏ, ảnh sau khi khôi phục càng giống với ảnh gốc. Tuy nhiên, việc ảnh khôi phục giống ảnh gốc không đồng nghĩa với việc ảnh khôi phục đẹp trong nhận thức thị giác của con người.
- **PSNR (Peak Signal-to-Noise Ratio)** Đo lường tỷ số tín hiệu và nhiễu, được tính bằng logarit cơ số 10 của tỷ số giữa giá trị tín hiệu tối đa và sai số trung bình bình phương. Giá trị PSNR càng cao, ảnh sau khi khôi phục càng có chất lượng tốt. Mặc dù PSNR dễ sử dụng nhưng cũng có những hạn chế tương tự như MSE, đó là chỉ tập trung vào sự khác biệt về điểm ảnh (pixel) mà không cân nhắc đến nhận thức thị giác của con người.

1.3 Một số phương pháp thường dùng

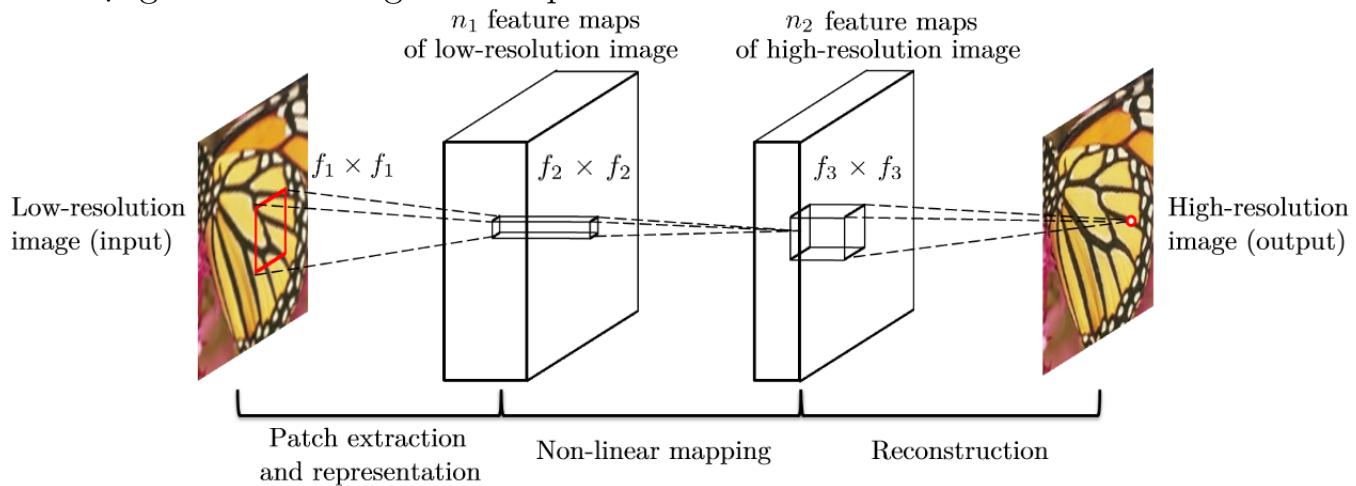
Hiện nay, có nhiều phương pháp khác nhau để khôi phục ảnh, bao gồm:

- **Phương pháp nội suy** Sử dụng thông tin từ các pixel lân cận để dự đoán giá trị của các pixel bị thiếu. Tuy nhiên, các phương pháp nội suy thường không thể tái tạo các chi tiết trong ảnh một cách chính xác.
- **Phương pháp lọc** Sử dụng các bộ lọc để loại bỏ nhiễu và tăng cường các chi tiết trong ảnh. Một số phương pháp lọc phổ biến bao gồm lọc trung bình, lọc Gaussian, lọc median,...

- **Phương pháp biến đổi miền** Biến đổi ảnh sang một miền khác (như miền tần số) để xử lý nhiễu và sau đó biến đổi ảnh ngược lại miền không gian. Một số phương pháp biến đổi miền phổ biến bao gồm biến đổi Fourier, biến đổi wavelet,...
- **Phương pháp học máy** Sử dụng mạng nơ-ron nhân tạo để học cách khôi phục ảnh từ các tập dữ liệu hình ảnh. Các phương pháp học máy có thể tái tạo các chi tiết trong ảnh một cách chính xác hơn so với các phương pháp truyền thống. Trong báo cáo này, nhóm chúng tôi thử nghiệm bộ lọc Non-local Means (NLM) và mô hình SRCNN.
- **Phương pháp SRCNN (Super-Resolution Convolutional Neural Network)** Đây là một phương pháp học máy hiệu quả cho việc khôi phục ảnh, sử dụng mạng nơ-ron tích chập để học cách khôi phục ảnh từ các tập dữ liệu hình ảnh. SRCNN có thể tái tạo các chi tiết trong ảnh một cách chính xác và có hiệu quả tính toán cao.
Trong báo cáo này, chúng tôi thử nghiệm khôi phục chất lượng ảnh bằng phương pháp SRCNN.

2 Cơ chế hoạt động của SRCNN [1]

Mạng SRCNN bao gồm ba lớp chính:



- Lớp trích xuất và biểu diễn mảng
 - Lớp này chia ảnh đầu vào (Y) thành các mảng con nhỏ (patch) có kích thước cố định.
 - Mỗi mảng con được biến đổi thành một vector đặc trưng bằng cách sử dụng bộ lọc tích chập và hàm kích hoạt phi tuyến tính (ví dụ: ReLU).
 - Mục đích của lớp này là trích xuất các đặc trưng quan trọng từ ảnh đầu vào và biểu diễn dưới dạng vector.
- Lớp ánh xạ phi tuyến tính

- Lớp này bao gồm nhiều lớp tích chập liên tiếp, mỗi lớp được theo sau bởi hàm kích hoạt phi tuyến tính (ReLU).
 - Mục đích của lớp này là học tập mối quan hệ phi tuyến tính giữa các đặc trưng được trích xuất từ lớp trước.
 - Lớp này đóng vai trò chính trong việc nâng cao chất lượng hình ảnh.
- Lớp tái tạo

- Lớp này sử dụng một bộ lọc tích chập đơn để biến đổi vector đặc trưng sau khi qua lớp ánh xạ phi tuyến tính thành ảnh siêu phân giải.
- Mục đích của lớp này là tái tạo ảnh có độ phân giải cao từ các đặc trưng được học tập bởi mạng SRCNN.

3 Xây dựng mô hình SRCNN

3.1 Triển khai và huấn luyện mô hình

1. Chuẩn bị dữ liệu: Chuẩn bị bộ dữ liệu ảnh chất lượng cao và phiên bản có độ phân giải thấp tương ứng. Bộ dữ liệu này được sử dụng để huấn luyện mô hình SRCNN.

Chúng tôi sử dụng dữ liệu từ LSDIR Datasets. Do thời gian và năng lực tính toán có hạn, chúng tôi chỉ sử dụng 50 ảnh.

2. Xây dựng kiến trúc mạng: Xây dựng mô hình SRCNN với ba lớp chính như đã mô tả ở phần trước. Số lượng lớp tích chập, kích thước kernel và các tham số khác có thể được điều chỉnh theo yêu cầu và tài nguyên tính toán.
3. Hàm mất mát (Loss function): Chọn hàm mất mát phù hợp để đo lường độ sai lệch giữa ảnh dự đoán bởi SRCNN và ảnh chất lượng cao thực tế. Hàm Mean Squared Error (MSE) được sử dụng trong trường hợp này.
4. Bộ tối ưu hóa (Optimizer): Chọn bộ tối ưu hóa thích hợp (VD: Adam) để cập nhật các tham số của mô hình trong quá trình huấn luyện.
5. Huấn luyện mô hình: Huấn luyện mô hình SRCNN trên bộ dữ liệu đã chuẩn bị. Quá trình huấn luyện sẽ lặp lại nhiều lần, cập nhật các tham số của mô hình để giảm thiểu hàm mất mát.
6. Đánh giá mô hình: Sau khi huấn luyện, đánh giá hiệu quả của mô hình trên một tập dữ liệu kiểm định (validation set) riêng biệt. Các hàm đánh giá SSIM, PSNR có thể được sử dụng để đánh giá chất lượng ảnh phục hồi.

3.2 Kết hợp mô hình CNN với bộ lọc NLM

Bên cạnh việc sử dụng riêng lẻ, SRCNN có thể được kết hợp với các bộ lọc khác để cải thiện hiệu quả khôi phục ảnh. Bộ lọc Non-local Means (NLM) là một bộ lọc lọc hiệu quả có thể loại bỏ nhiễu và bảo vệ các chi tiết trong ảnh. Việc kết hợp SRCNN với NLM có thể tận dụng được ưu điểm của cả hai phương pháp:

- Bộ lọc NLM có khả năng loại bỏ nhiễu hiệu quả.
- SRCNN có khả năng học các đặc trưng ảnh cấp cao và tái tạo các chi tiết.

3.3 Kết quả thử nghiệm



Hình 1: Osaka (Degraded, SRCNN, Origin)

Degraded Image

PSNR: 25.737766636795655

MSE: 520.5050553523914

SSIM: 0.8184708507528594

Reconstructed Image

PSNR: 26.49846456795781

MSE: 436.8729536193095

SSIM: 0.8600054492704533



Hình 2: Degraded



Hình 3: SRCNN



Hình 4: Origin
Comparison of Osaka (Ikayaki) Images



Hình 5: Handsome Man (Degraded, SRCNN, Origin)

Degraded Image

PSNR: 24.040401742421864

MSE: 769.4164041659573

SSIM: 0.8299910354306684

Reconstructed Image

PSNR: 26.527879247619456

MSE: 433.9240195243771

SSIM: 0.9040014645241033



Hình 6: Degraded



Hình 7: SRCNN



Hình 8: Origin
Comparison of Handsome Man (Face) Images

4 Mã nguồn

Toàn bộ chương trình, bao gồm mã nguồn, bộ dữ liệu và pre-trained model (.h5 format) được lưu trong Github repository cuisinecometwot/IT3160-Project-G22.

5 Kết luận

Mô hình SRCNN hoặc bộ lọc NLM kết hợp SRCNN đã đạt được kết quả tương đối khả quan trong việc phục hồi ảnh có độ phân giải thấp. Tuy nhiên, vẫn còn một số hạn chế cần được khắc phục:

- Nhiều ảnh chưa được xử lý triệt để: Một số nhiều ảnh vẫn còn sót lại trong ảnh phục hồi, ảnh hưởng đến chất lượng ảnh.
- Mất mát thông tin: Ở một số vị trí trong ảnh phục hồi, thông tin chi tiết bị mất mát, dẫn đến ảnh bị mờ hoặc thiếu nét.

Những hạn chế này có thể xuất phát từ một số nguyên nhân:

- Tham số bộ lọc chưa được tối ưu: Các tham số của bộ lọc NLM chưa được điều chỉnh phù hợp với từng trường hợp cụ thể, dẫn đến hiệu quả xử lý nhiều chưa tối ưu.
- Dữ liệu huấn luyện hạn chế: Mô hình SRCNN này được huấn luyện trên tập dữ liệu có kích thước nhỏ (50 ảnh), dẫn đến khả năng khai hóa thấp và hiệu quả phục hồi chưa cao.
- Sức mạnh tính toán hạn chế: Nguồn lực tính toán hạn chế dẫn đến khó khăn trong quá trình đào tạo lượng dữ liệu lớn.

Để cải thiện hiệu quả phục hồi ảnh có độ phân giải thấp, một số hướng giải quyết sau có thể được cân nhắc:

- Mở rộng tập dữ liệu huấn luyện: Thu thập thêm dữ liệu ảnh chất lượng cao và ảnh chất lượng thấp với đa dạng loại lỗi (nhiều, mờ, ...) để huấn luyện mô hình, giúp mô hình học được nhiều đặc điểm ảnh hơn và nâng cao khả năng khai hóa.
- Nghiên cứu các mô hình mới
- Thủ kết hợp mô hình SRCNN với các phương pháp khác

Tài liệu

- [1] Dong, Chao, et al. Image Super-Resolution Using Deep Convolutional Networks. 2015.