

MỘT MẠNG NƠN TĂNG CƯỜNG MỨC TƯƠNG PHẢN CỦA ẢNH SỐ

A NEURAL NETWORK ENHANCES THE CONTRAST OF DIGITAL IMAGES

Phạm Tuấn Đạt

Khoa Công nghệ Thông tin, Trường Đại học Hàng hải Việt Nam

TÓM TẮT

Tăng cường mức tương phản của ảnh số có thể ứng dụng cho phục chế và lưu trữ cũng như phát hiện và phân đoạn đối tượng. Do đó, bài báo này trình bày một mô hình mạng huấn luyện nhằm cải thiện chất lượng hình ảnh tối màu có dải mức xám hẹp. Kết quả thử nghiệm trên tập kiểm thử cho tỷ lệ tín hiệu trên tạp âm là tương đối ấn tượng, xấp xỉ 32.5 dB. Mặt khác, mô hình mạng này có những ưu thế nhất định khi so sánh với phương pháp xử lý trên biểu đồ tần suất mức xám và một mô hình mạng học khác đã được công bố trước đó.

Từ khóa: Mã hóa tự động; Sai số tuyệt đối; Tỷ lệ tín hiệu trên tạp âm; Suy thoái.

ABSTRACT

Enhancing the contrast of digital images can be applied to restoration and storage, as well as object detection and segmentation. Therefore, this paper presents a training network model to improve the quality of dark images with a narrow gray level range. The test results on the test set show a relatively impressive signal-to-noise ratio, approximately 32.5 dB. On the other hand, this network model has certain advantages when compared with the gray-level frequency histogram processing method and another previously published network model.

Keywords: Auto-encoder; Absolute error; Signal-to-noise ratio; Degradation.

1. GIỚI THIỆU

Hình ảnh thu nhận trong cuộc sống không phải luôn cho chất lượng tốt. Những bức ảnh được chụp quét từ sách cũ cho tông màu nghèo nàn vì dải mức sáng hạn chế. Một trường hợp khác là khi người dùng muốn phục chế ảnh cũ sau nhiều năm hoặc rửa ảnh từ phim âm bản không được bảo quản trong điều kiện tiêu chuẩn. Vấn đề này nảy sinh ứng dụng thực

tế đó là cải thiện hình ảnh cho lưu trữ tái bản, hoặc tạo ra những bức ảnh như vừa mới chụp. Hơn nữa, trong lĩnh vực học sâu, cải thiện độ tương phản còn hỗ trợ cho bài toán phát hiện đối tượng và phân đoạn.

Kỹ thuật phổ biến [1] để tăng cường mức tương phản dựa trên các thao tác xử lý trên biểu đồ mức xám. Giải thuật cân bằng biểu đồ thích ứng giới hạn độ tương phản [2] cũng

là một đại diện điển hình. Ưu điểm của hướng tiếp cận truyền thống là không cần dữ liệu học và tính toán nhanh nhưng có thể cho ảnh đầu ra với chất lượng rất cao. Tuy vậy, các kỹ thuật trên không đem đến hiệu quả tương xứng nếu những hình ảnh đã có suy giảm đồng thời các đặc trưng cường độ và mức bão hòa.

Phương pháp huấn luyện dựa trên các mô hình học sâu đòi hỏi dữ liệu phù hợp và chi phí tính toán rất lớn. Để khắc phục vấn đề đó, gần đây có một hướng giải quyết [3] kết hợp mạng học với kỹ thuật xử lý ảnh truyền thống để tự huấn luyện tập ảnh đầu vào có mức xám thấp, nhờ đó giảm bớt một nửa khối lượng dữ liệu đào tạo.

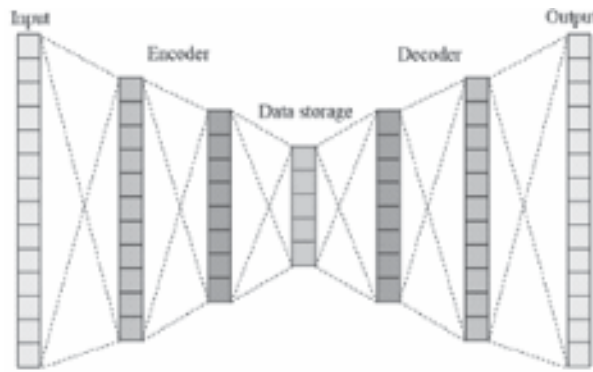
Bài báo này đề xuất một mô hình mã hóa tự động (auto-encoder [9]) đào tạo trên một tập dữ liệu chứa các cặp hình ảnh có mức tương phản thấp và cao. Những ảnh tối màu có dải mức xám hẹp được sản sinh dựa trên điều chỉnh giảm cường độ sáng và mức bão hòa. Mô hình được thử nghiệm và cho sai số tuyệt đối [11] nhỏ cũng như tỷ lệ tín hiệu trên tạp âm [10] tương đối cao, PSNR tổng thể cỡ 32.5 dB. Kết quả này tốt hơn nếu so sánh với việc xử lý trên biểu đồ tần suất mức xám và một mô hình mạng áp dụng giải pháp kết hợp nêu trên.

2. MÔ HÌNH THỬ NGHIỆM VÀ SO SÁNH

2.1. Mô hình mạng auto-encoder

Tùy theo ứng dụng cụ thể mà kiến trúc của mô hình mạng huấn luyện sẽ tùy biến khác nhau, nhưng các mô hình học sâu luôn chứa các tầng cơ sở [5] của mạng nơ ron nhân chập. Trong khi mô hình tạo sinh [7, 8] chủ yếu dùng để sản sinh hình ảnh chứa nội dung mới trên cơ sở dữ liệu cho trước và từ khóa gợi ý, thì mô hình mã hóa tự động áp dụng cho các ứng dụng

phục chế hay cải thiện ảnh. Điểm tương đồng giữa các mô hình trên là đầu vào và đầu ra của mạng đều nhận và sản sinh dữ liệu ảnh.



Hình 1. Kiến trúc tổng quan mạng mã hóa tự động.

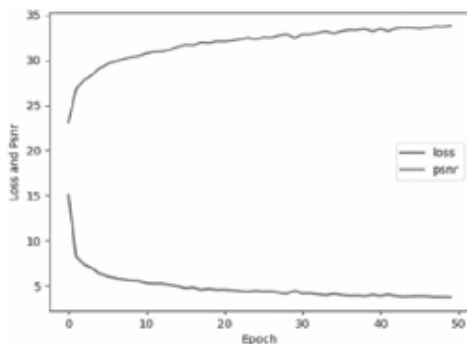
Hình ảnh trong hình 1 cho biểu diễn tổng quan nhưng không hoàn toàn giống cấu trúc của mô hình mạng được thử nghiệm. Thực tế, khối mã hóa chứa 3 tầng convolutional và 3 tầng spooling có trật tự xen kẽ, tương tự thế cho 3 tầng convolutional và 3 tầng upsampling của khối giải mã. Các tầng nhân chập trong mỗi khối này có từ 64 tới 256 ma trận lọc kích thước 3x3, khối lưu trữ dữ liệu thô (giữa khối mã hóa và giải mã) chỉ chứa 32 ma trận lọc. Tới nay, mặc dù đã có nhiều cải tiến cho mạng nhân chập nhưng quá trình huấn luyện của mạng nơ ron nhìn chung luôn đi kèm sai số tính toán. Nhằm loại trừ suy thoái huấn luyện, khối giải mã có thêm những kết nối cần thiết (skip connection [6]) của mạng Resnet. Mô hình cũng điều chỉnh khắc phục overfitting [12] cùng với giải thuật Adam [4] tối ưu độ chính xác.

2.2. Thử nghiệm

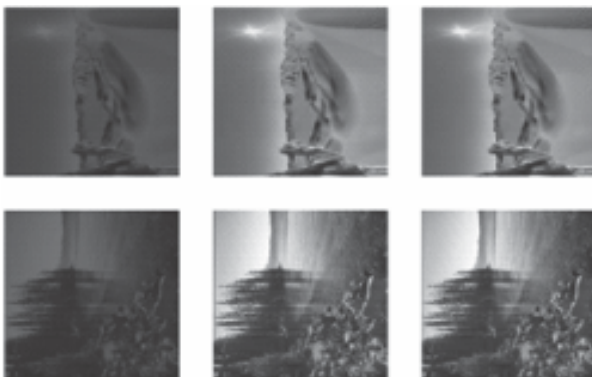
Tập dữ liệu thử nghiệm gồm 12000 cặp ảnh (24 bit màu theo phân giải 192x192, và được lưu trữ theo định dạng không mất dữ liệu) có tương phản thấp và cao tương ứng, có thể tải từ địa chỉ “<https://www.kaggle.com/datasets/tuandat/archive>”. Những hình ảnh có

mức tương phản thấp được sản sinh sau khi suy giảm đồng thời cường độ và mức bão hòa trên hệ màu HSV. Tập dữ liệu được chia thành tập huấn luyện và tập kiểm thử theo tỷ lệ 2:1. Trong đó, tập kiểm thử gồm có dữ liệu kiểm thử xác nhận (validation data) trong quá trình huấn luyện và dữ liệu kiểm thử độc lập (test data) sau huấn luyện, với tỷ lệ 1:3.

Kết quả thử nghiệm cho thấy không nảy sinh vấn đề overfitting vì độ lệch PSNR tổng thể giữa tập huấn luyện và tập kiểm thử xác nhận ít hơn 2.5 dB. Bên cạnh đó, sai số tuyệt đối tổng thể (Loss) trên tập huấn luyện và tập kiểm thử độc lập xấp xỉ nhau. Mặt khác, tiến trình đào tạo cũng không bắt gặp tình huống suy thoái, độ chính xác tỷ lệ thuận với số lượng vòng học. Hình 2 biểu diễn xu hướng của Loss và PSNR trong 50 vòng lặp trên tập huấn luyện.



Hình 2. Xu hướng tiến trình huấn luyện.



Hình 3. Minh họa thử nghiệm dựa trên mạng mã hóa tự động.

Bảng 1. Loss và PSNR trên tập huấn luyện và kiểm thử độc lập.

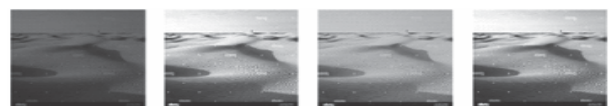
Tập huấn luyện		Tập kiểm thử độc lập	
Loss	PSNR	Loss	PSNR
4.3	33.0	4.4	32.5

Ở hình 3, theo các cột từ trái qua phải, lần lượt chứa ảnh màu tối, ảnh sau cải thiện và ảnh có dải mức sáng cao. Theo chỉ số PSNR trên tập này (Bảng 1) tuy không xuất sắc nhưng khá cao. Từ quan sát thực, ảnh được bảo tồn rất tốt các chi tiết và dải mức xám, gần giống ảnh chụp ở điều kiện lý tưởng.

2.3. So sánh

Để có đánh giá khách quan về hiệu quả của mạng auto-encoder này, thử nghiệm so sánh mô hình với một giải thuật tương phản tự động cài đặt bởi thư viện lập trình Pillow và một mạng học kết hợp, cùng tiến hành trên tập dữ liệu đã cho.

Giải thuật tương phản tự động thực hiện tính tần suất của các mức xám trong ảnh. Tiếp theo, nó ánh xạ các mức xám trở lại hình ảnh kết quả. Phép ánh xạ cho hình ảnh cần thỏa mãn mục tiêu: những điểm ảnh nổi bật sẽ sáng hơn, còn những điểm ảnh tối màu sẽ tối hơn. Nhằm bảo tồn các chi tiết trong bức ảnh thì giải thuật chuyển đổi từ ảnh màu 24 bit sang ảnh xám 8 bit trước khi tiến hành các biến đổi xử lý ảnh. Kỹ thuật này cho kết quả PSNR rất cao trên hình ảnh suy giảm cường độ sáng hoặc độ bão hòa nhưng lại không cho hiệu quả tối ưu trên tập dữ liệu thử nghiệm độc lập, PSNR xấp xỉ 25.0 dB.



Hình 4. Minh họa thử nghiệm dựa trên tương phản tự động và mô hình học kết hợp.

Bảng 2. PSNR của giải thuật và các mạng học.

Tương phản tự động	Mạng học kết hợp	Mạng học Auto-encoder
25.0	20.0	32.5

Hướng kết hợp giữa mô hình mạng học và các phép lọc số chỉ tiếp nhận dữ liệu ảnh màu tối có dải mức sáng hẹp. Không tham chiếu tới tập hình ảnh có mức sáng tiêu chuẩn, mạng thực hiện quá trình tự đào tạo để điều chỉnh độ lệch màu tiềm ẩn, độ lệch giữa cường độ xám trung bình của mỗi vùng cục bộ và ngưỡng chọn trước, bảo toàn sự trơn mịn cũng như sự tương phản màu. Do đó, ước lượng tổn thất không dựa theo sai số tuyệt đối. Tiến trình đào tạo trên dữ liệu huấn luyện đã dẫn tới suy thoái sau vài vòng lặp, tổn thất ước lượng không có điều chỉnh đáng kể, kết quả là giá trị PSNR trên tập thử nghiệm độc lập khoảng 20.0 dB.

Trong hình 4, các hình ảnh theo thứ tự gồm có: ảnh màu tối, ảnh cải thiện với tương phản tự động, ảnh cải thiện bằng mạng học kết hợp, và ảnh có tương phản cao. Có thể thấy hình ảnh được tăng cường theo những kỹ thuật này vẫn giữ được chi tiết biên nhưng độ sáng cũng như màu sắc của ảnh cải thiện không giống được ảnh ban đầu có tương phản cao. Cũng từ những thử nghiệm liên quan, nhận thấy mô hình mạng huấn luyện kết hợp cần có phân giải hình ảnh phù hợp và tập tham số như hàm kích hoạt hoặc ma trận lọc số tương ứng với mỗi tập dữ liệu cụ thể.

3. KẾT LUẬN

Bài báo đã nêu một mô hình mã hóa tự động và thử nghiệm kết quả tăng cường trên một tập dữ liệu chứa hình ảnh đã suy giảm cường độ và mức bão hòa. Mạng huấn luyện của mô hình cho PSNR không thật xuất sắc nhưng rất ấn tượng và hình ảnh tăng cường được cải thiện rõ rệt. So với kỹ thuật xử lý ảnh số truyền thống

cũng như mạng học kết hợp, mô hình cần nhiều dữ liệu huấn luyện và thời gian tính toán nhưng không đòi hỏi hiểu biết sâu đối với đặc trưng màu sắc và có thể thích nghi với nhiều tập dữ liệu đào tạo cải thiện hoặc phục chế ảnh. Hơn nữa, mô hình này còn hỗ trợ được cho các bài toán phát hiện và phân đoạn đối tượng.

Lời cảm ơn:

Tác giả gửi lời cảm ơn tới những đồng nghiệp tại Khoa Công nghệ Thông tin và Trường Đại học Hàng hải đã giúp đỡ, góp ý cho nội dung bài báo. ♦

Ngày nhận bài: 25/10/2024

Ngày phản biện: 22/11/2024

Tài liệu tham khảo:

- [1]. Fredrik Lundh, "Secret Labs AB". Python Imaging Library, 2024.
- [2]. Karel Zuiderveld, "Contrast limited adaptive histogram equalization". Computer Vision Research Group, Utrecht University, 1994.
- [3]. Chunle Guo, Chongyi Li, Jichang Guo, Chen Change Loy, Junhui Hou, Sam Kwong, và Runmin Cong, "Zero-reference deep curve estimation for low-light image enhancement". Computer Vision Foundation Open Access, 2020.
- [4]. Diederik P. Kingma và Jimmy Lei Ba, "Adam: a method for stochastic optimization". International Conference on Learning Representations (ICLR), 2015.
- [5]. Jiuxiang Gao, Zhenhua Wang, Jason Kuen, Lianyang Ma, Amir Shahroudy, Bing Shuai, Ting Liu, Xingxing Wang, Li Wang, Gang Wang, Jianfei Cai và Tsuhan Chen, "Recent advances in convolutional neural networks". Nanyang University, 2017.
- [6]. Kaiming He, Xiangyu Zhang, Shaoqing Ren và Jian Sun, "Identity mappings in deep residual networks". European Conference on Computer Vision, 2016.
- [7]. Mehdi Mirza và Simon Osindero, "Conditional generative adversarial nets". <https://arxiv.org/pdf/1411.1784>, 2014.
- [8]. Alec Radford, Luke Metz, và Soumith Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks". ICLR, 2016.
- [9]. Umberto Michelucci, "An introduction to autoencoders". <https://arxiv.org/pdf/2201.03898>, 2022.
- [10]. Umme Sara, Morium Akter, và Mohammad Shorif Uddin, "Image quality assessment through FSIM, SSIM, MSE and PSNR - A comparative study". Journal of Computer and Communications, Vol. 7, No. 3, 2019.
- [11]. Jun Qi, Jun Du, Sabato Marco Siniscalchi, Xiaoli Ma, và Chin-Hui Lee, "On mean absolute error for deep neural network based vector-to-vector regression". IEEE Signal Processing Letters, Vol. 27, pp. 1485-1489, 2020.
- [12]. Xue Ying, "An Overview of Overfitting and its Solutions". Journal of Physics: Conference Series, 2019.