

SIÊU PHÂN GIẢI DÙNG TRÍ TUỆ NHÂN TẠO NÂNG CAO CHẤT LƯỢNG ẢNH VỔNG MẠC MẮT NGƯỜI

SUPER RESOLUTION USING ARTIFICIAL INTELLIGENCE FOR RETINA IMAGE ENHANCEMENT

Doãn Thanh Bình

Trường Đại học Điện lực

Ngày nhận bài: 28/07/2020, Ngày chấp nhận đăng: 28/12/2020, Phản biện: TS. Hoàng Thị Phương Thảo

Tóm tắt:

Hình ảnh võng mạc là một lĩnh vực quan trọng trong y tế phục vụ mục tiêu điều trị các bệnh lý. Quan sát những thay đổi của các đường mạch máu ở võng mạc giúp các bác sỹ chẩn đoán nhiều bệnh, thu thập, phân tích các triệu chứng và phát triển các phương pháp điều trị liên quan. Do vậy, nâng cao chất lượng hình ảnh võng mạc là bước tiền xử lý quan trọng. Nhằm cải thiện chất lượng hình ảnh võng mạc một số kỹ thuật đã được đề xuất như biến đổi wavelet rời rạc (DWT) [1,2,3], siêu phân giải rất sâu (VDSR) [4], mạng nơon tích chập siêu phân giải (SRCNN) [5]... Tuy nhiên, vẫn chưa thể đem lại hiệu quả cao do vẫn tồn tại nhiễu cao, cho kết quả hình ảnh không tốt, chưa tối ưu về độ phức tạp tính toán, mức tiêu thụ bộ nhớ và tốc độ xử lý. Trong khuôn khổ bài báo này, tác giả đề xuất phương pháp siêu phân giải dùng trí tuệ nhân tạo nâng cao chất lượng ảnh võng mạc mắt người, nhằm tái tạo trực tiếp hình ảnh có độ phân giải cao từ hình ảnh có độ phân giải thấp ban đầu. Bằng các phân tích và kết quả tính toán các tham số chất lượng ảnh xử lý qua thực nghiệm, bài báo chứng minh rằng phương pháp được đề xuất là vượt trội so với các phương pháp đã thực hiện, đặc biệt là về thời gian thực thi.

Từ khóa:

nâng cao chất lượng ảnh Retina, mạng nơon tích chập sâu (CNN), siêu phân giải đơn ảnh.

Abstract:

The retina image is an important area for medical treatment of the disease. By observing the changes in the blood vessels in the retina lines, doctors can diagnose diseases, to collect and analyze the symptoms and the development of related treatments. Consequently, improving retinal image quality is an important preprocessing step. In order to improve retinal image quality, several techniques have been proposed such as wavelet transform [1,2,3], very-deep-super-resolution (VDSR) [4], super-resolution-convolutional neural network (SRCNN) [5]... but still can not provide high efficiency by persistent high noise, poor image results, not optimal for computational complexity and memory consumption. Therefore, in this paper, we propose a particular method of retinal images quality enhancement via super resolution using artificial intelligence to directly reconstruct the high resolution image from the original low resolution image. By the analysis and calculated results in picture quality parameters through experimental treatment, we will demonstrate that the proposed method is superior to the state-of-the-art methods, especially in terms of time performance.

Keywords:

retinal image enhancement, convolutional neural network, single image super resolution.

1. GIỚI THIỆU**1.1. Ảnh võng mạc mắt người**

Võng mạc là một cấu trúc nhiều lớp với nhiều lớp tế bào thần kinh kết nối với nhau bằng các khớp thần kinh. Các tế bào thần kinh chỉ nhạy cảm với ánh sáng trực tiếp là các tế bào tiếp nhận ánh sáng. Đối với tầm nhìn, đây là hai loại: các que và hình nón. Thanh chức năng chủ yếu trong ánh sáng mờ và cung cấp tầm nhìn màu đen và trắng, trong khi tế bào hình nón hỗ trợ nhận thức về màu sắc. Loại thứ ba của tiếp nhận ánh sáng là sử dụng tế bào hạch quang có khả năng cảm biến với cường độ của ánh sáng.

Đánh giá hình ảnh võng mạc là điều cần thiết để chăm sóc mắt hiện đại. Với sự ra đời của các thiết bị xử lý hình ảnh, ghi âm kỹ thuật số và xử lý hình ảnh võng mạc được bắt đầu nghiên cứu phát triển.

Bài báo này đề cập đến ảnh võng mạc dựa trên trí tuệ nhân tạo để nâng cao, cải thiện chất lượng ảnh, tái tạo trực tiếp hình ảnh có độ phân giải cao từ hình ảnh có độ phân giải thấp nhằm phục vụ cho các bước chẩn đoán lâm sàng về các bệnh lý liên quan đến nhãn khoa.

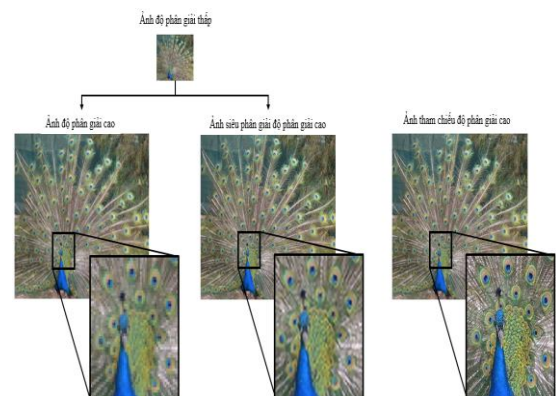
Bố cục của bài báo như sau: giới thiệu về ảnh võng mạc, cơ sở lý thuyết về siêu phân giải đơn giản, các phương pháp siêu phân giải dùng trí tuệ nhân tạo dựa trên mạng nơron tích chập, đưa ra mô hình cho phương pháp đề xuất. Phần cuối đưa ra kết quả thực nghiệm, và các tham số tính toán so sánh chất lượng xử lý giữa các

phương pháp hiện tại và phương pháp đề xuất.

1.2. Siêu phân giải đơn ảnh

Siêu phân giải là quá trình tạo hình ảnh có độ phân giải cao từ hình ảnh có độ phân giải thấp. Bài báo này xem xét siêu phân giải đơn ảnh (SISR), trong đó mục tiêu là khôi phục một hình ảnh có độ phân giải cao từ một hình ảnh có độ phân giải thấp. SISR là một thách thức vì nội dung hình ảnh tần số cao thường không thể được phục hồi từ hình ảnh độ phân giải thấp. Không có thông tin tần số cao, chất lượng của hình ảnh độ phân giải cao sẽ bị hạn chế.

Siêu phân giải đơn ảnh (SISR) là một bài toán cổ điển trong thị giác máy tính bậc thấp để tái tạo hình ảnh có độ phân giải cao (HR) từ hình ảnh có độ phân giải thấp (LR). Thực tế, một số lượng vô hạn các hình ảnh có độ phân giải cao có thể thu được một hình ảnh độ phân giải thấp giống nhau bằng phương pháp downsampling.



Hình 1. Ví dụ về siêu phân giải đơn ảnh (SISR)

Gần đây, nhờ mạng nơron tích chập (Convolutional Neural Network - CNN) phát triển hơn, nhiều phương pháp siêu phân giải dựa trên CNN đã xây dựng một mạng nhiều lớp để đạt được hiệu suất phục hồi tốt hơn. Kim cùng cộng sự đề xuất một mô hình CNN 20 lớp được gọi là siêu phân giải đơn ảnh nhiều lớp VDSR [4], áp dụng phương pháp mạng học nhiều lớp và cắt giảm độ dốc thích ứng để việc đào tạo trở nên dễ dàng hơn.

Mặc dù đạt được hiệu suất nổi bật, nhưng hầu hết các mạng nhiều lớp vẫn có một số nhược điểm. Thứ nhất, độ phức tạp cao về tính toán và tiêu thụ bộ nhớ lớn, ít được áp dụng trong thực tế, chẳng hạn như các ứng dụng thị giác di động và nhúng. Hơn nữa, các mạng tích chập truyền thống thường áp dụng các cấu trúc liên kết mạng nổi tầng, ví dụ: VDSR [4] và DRCN [6]. Bằng cách này, các ma trận lớp đầu vào (Feature map) của từng lớp được truyền tải đến lớp liên tiếp mà không có sự phân biệt.

Để giải quyết những điểm hạn chế này, tác giả đề xuất một mạng chất lọc thông tin mới (IDN) với các thông số bậc thấp và tính phức tạp tính toán như minh họa trong hình 2. Ở IDN được đề xuất, khối trích xuất đặc trưng (FBlock) đầu tiên sẽ trích xuất các đặc trưng từ hình ảnh độ phân giải thấp. Sau đó, các khối chất lọc thông tin (DBlocks) được xếp chồng lên nhau để chất lọc dần dần thông tin còn lại. Cuối cùng, khối tái cấu trúc (RBlock) sẽ tổng hợp các phép biểu diễn phần dư độ phân giải cao thu được để tạo ra dư ảnh. Để có được hình ảnh độ phân giải cao, tác

giả thực hiện thao tác bổ sung toán tử trên dư ảnh và hình ảnh phân giải thấp đã được upsampling.

Yếu tố chính của IDN là khối chất lọc thông tin, chứa đơn vị tăng cường và đơn vị nén. Đơn vị tăng cường chủ yếu gồm hai mạng tích chập như minh họa trong hình 2. Mỗi mạng là một môđun ba lớp. Các bản đồ đặc năng của môđun đầu tiên được trích xuất thông qua một đường dẫn ngắn (3 lớp). Qua đó, chúng có thể được coi là các đặc trưng đường ngắn cục bộ.

Những đóng góp chính trong nghiên cứu này được tóm tắt như sau:

- IDN sẽ thu được kết quả mang tính cạnh tranh dù sử dụng số lượng lớp tích chập ít hơn.
- Do cấu trúc ngắn gọn của IDN được đề xuất, nên nó có tốc độ nhanh hơn nhiều so với một số phương pháp siêu phân giải dựa trên CNN, ví dụ: VDSR [4], SRCNN [5], hay phương pháp DWT [1,2,3] truyền thống.

2. CÁC CÔNG TRÌNH NGHIÊN CỨU LIÊN QUAN

Siêu phân giải đơn ảnh đã được nghiên cứu rộng rãi trong những năm gần đây. Trong phần này, sẽ tập trung vào các phương pháp dựa trên mạng lưới nơron như VDSR [4], và SRCNN [5].

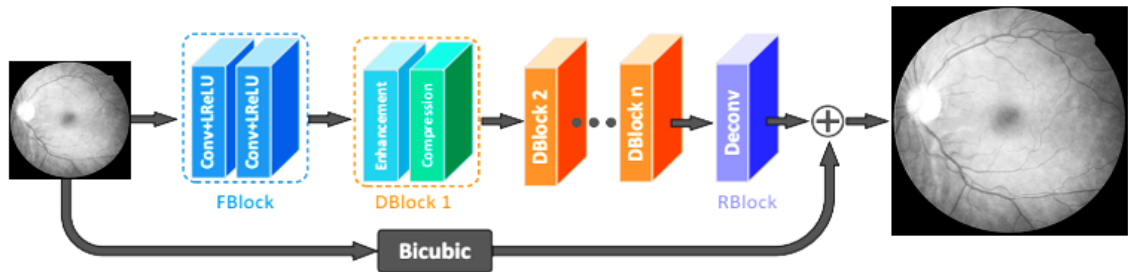
Dong cùng cộng sự [7, 8] trước tiên khai thác mạng nơron tích chập ba lớp, có tên là SRCNN, nhằm tối ưu hóa quá trình trích xuất đặc trưng, ánh xạ phi tuyến tính và tái tạo hình ảnh từ đầu đến cuối. Về sau, Shi cùng các cộng sự [9] đề xuất một mạng nơron tích chập điểm ảnh phụ

(subpixel) hiệu quả (ESPCN), trích xuất các ma trận lớp đầu vào trong không gian độ phân giải thấp và thay thế việc thực hiện phép toán nội suy (upsampling) song lập phương bằng một tích chập điểm ảnh phụ hiệu quả. Dong cùng cộng sự [5] sử dụng phương pháp giải chập để thúc đẩy SRCNN kết hợp với kích thước bộ lọc nhỏ hơn và nhiều lớp tích chập hơn. Kim cùng cộng sự [4] đề xuất một mô hình CNN nhiều lớp với kiến trúc hồi quy đa

biến nhằm thu được hiệu suất vượt trội, sử dụng thông tin theo ngữ cảnh trên các vùng hình ảnh lớn.

3. PHƯƠNG PHÁP ĐỀ XUẤT SIÊU PHÂN GIẢI DÙNG TRÍ TUỆ NHÂN TẠO

Ở phần này, đầu tiên tác giả mô tả kiến trúc mô hình được đề xuất. Tiếp theo, đề xuất đơn vị tăng cường và đơn vị nén, chúng là cốt lõi của phương pháp được đề xuất.



Hình 2. Cấu trúc của mạng đề xuất

3.1. Cấu trúc mạng

Mạng IDN được đề xuất như trong hình 2, gồm ba phần: khối trích xuất chức năng (FBlock), các khối chất lọc thông tin xếp nối (DBlocks) và khối tái tạo (RBlock).

Ở đây, x và y là đầu vào và đầu ra của mạng IDN.

Đối với FBlock, hai lớp chập 3×3 được sử dụng để trích xuất các ma trận lớp đầu vào từ hình ảnh độ phân giải thấp gốc. Quy trình này có thể được trình bày là:

$$B_0 = f(x) \quad (1)$$

Trong đó, f đại diện cho hàm trích xuất đặc trưng và B_0 biểu thị cho các đặc trưng được trích xuất và sử dụng làm đầu vào cho giai đoạn tiếp theo.

Phần tiếp theo bao gồm các khối chất lọc thông tin, sử dụng chế độ móc nối. Mỗi

khối chứa một đơn vị tăng cường và một đơn vị nén theo kiểu xếp chồng. Quá trình này có thể được xây dựng theo công thức sau:

$$B_k = F_k(B_{k-1}), k = 1, 2, \dots, n \quad (2)$$

Trong đó, F_k biểu thị hàm DBlock thứ k , B_{k-1} và B_k là đầu vào và đầu ra của DBlock thứ k tương ứng. Cuối cùng, chúng ta có một tích chập chuyển vị mà không có hàm kích hoạt như RBlock. Do đó, mạng IDN có thể được trình bày như sau:

$$y = R(F_n(B_{n-1})) + U(x) \quad (3)$$

Trong đó R , U lần lượt biểu thị RBlock và phép nội suy hai chiều.

Hàm Loss

Xem xét hai hàm mất mát được sử dụng để đo

độ lệch giữa hình ảnh độ phân giải cao theo dự đoán \hat{I} và sự thật nền tảng I tương ứng. Hàm đầu tiên là sai số toàn phương trung bình (MSE), là hàm mất mát được sử dụng rộng rãi nhất để phục hồi hình ảnh chung như được định nghĩa dưới đây:

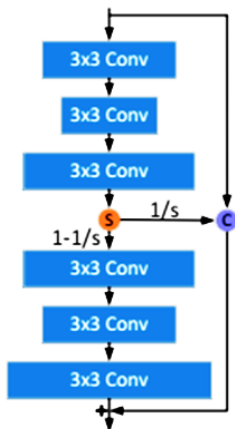
$$l_{MSE} = \frac{1}{N} \sum_{i=1}^N \|I_i - \hat{I}_i\|_2^2 \quad (4)$$

Tuy nhiên, qua thực nghiệm, Lim cùng cộng sự [16] chứng minh rằng đào tạo với hàm mất mát MSE không phải là một lựa chọn tốt. Hàm mất mát thứ hai có nghĩa là sai số tuyệt đối trung bình (MAE), được xây dựng như sau:

$$l_{MAE} = \frac{1}{N} \sum_{i=1}^N \|I_i - \hat{I}_i\|_1 \quad (5)$$

Bằng thực nghiệm, ta nhận thấy rằng mô hình với hàm Loss MSE có thể cải thiện hiệu suất của mạng đã đào tạo với hàm Loss MAE. Do đó, việc đầu tiên của phương pháp là đào tạo mạng với hàm Loss MAE và sau đó tinh chỉnh theo hàm Loss MSE.

3.2. Đơn vị tăng cường



Hình 3. Kiến trúc của đơn vị tăng cường trong mô hình đề xuất

Đơn vị tăng cường có thể được chia thành hai môđun, một là ba tích chập trên và hai là ba tích chập dưới. Môđun trên có ba tích chập 3×3 , mỗi tích chập được theo sau bởi hàm kích hoạt tinh chỉnh các đơn vị tuyến tính hở (LReLU), được bỏ qua ở đây. Hãy biểu thị các kích thước ma trận lớp đầu vào của lớp thứ i là D_i ($i = 1, \dots, 6$). Theo đó, mối quan hệ của các lớp chập có thể được biểu thị bằng:

$$D_3 - D_1 = D_1 - D_2 = d \quad (6)$$

trong đó d biểu thị độ chênh lệch giữa lớp thứ nhất và lớp thứ hai hoặc giữa lớp thứ nhất và lớp thứ ba. Tương tự, kích thước của các kênh trong môđun dưới cũng có mối quan hệ này và có thể được mô tả như sau:

$$D_6 - D_4 = D_4 - D_5 = d \quad (7)$$

trong đó $D_4 = D_3$. Môđun trên bao gồm ba lớp chập xếp tầng với các LReLU và đầu ra của lớp chập thứ ba được phân chia thành hai phân đoạn. Giả sử đầu vào của môđun này là B_{k-1} , chúng ta có:

$$P_l^k = C_a(B_{k-1}) \quad (8)$$

Trong đó B_{k-1} biểu thị đầu ra của khối trước đó và trong khi đó là đầu vào của khối hiện tại, C_a biểu thị hoạt động tích chập nối tiếp và P_l^k là đầu ra của môđun trên trong đơn vị tăng cường thứ k . Các ma trận lớp đầu vào với các kích thước $\frac{D_3}{s}$ và đầu vào của lớp chập đầu tiên được liên kết với nhau trong khuôn khổ kênh,

$$R^k = C\left(S\left(P_l^k, \frac{1}{s}\right), B_{k-1}\right) \quad (9)$$

Trong đó C, S đại diện cho hoạt động ghép nối và hoạt động phân chia. Cụ thể, chúng tôi biết kích thước của P_l^k là D_3 .

Do đó, $S\left(P_l^k, \frac{1}{s}\right)$ biểu thị rằng các đặc trưng kích thước $\frac{D_3}{s}$ được tìm nạp từ P_l^k .

Ngoài ra $S\left(P_l^k, \frac{1}{s}\right)$ ghép nối các đặc trưng với B_{k-1} trong khuôn khổ kênh. Mục đích là để kết hợp thông tin trước đó với một số thông tin hiện tại. Nó có thể được coi là thông tin đường dẫn cục bộ được giữ lại một phần. Sử dụng phần còn lại của thông tin đường dẫn ngăn cục bộ làm đầu vào cho môđun dưới, chủ yếu trích xuất thêm các ma trận lớp đầu vào đường dài,

$$P_2^k = C_b\left(S\left(P_l^k, 1 - \frac{1}{s}\right)\right) \quad (10)$$

Trong đó, P_2^k , C_b tương ứng là các hoạt động tích chập đầu ra và xếp nối của môđun dưới. Cuối cùng, như trong hình 3, thông tin đầu vào, thông tin đường dẫn cục bộ riêng và thông tin đường dài cục bộ được tổng hợp. Do đó, đơn vị tăng cường có thể được trình bày như sau:

$$P^k = P_2^k + R^k = C_b\left(S\left(C_a(B_{k-1}), \frac{1}{s}\right)\right) + C\left(S\left(C_a(B_{k-1}), \frac{1}{s}\right), B_{k-1}\right) \quad (11)$$

Trong đó P^k là đầu ra của đơn vị tăng cường. Tại thời điểm này, các đặc trưng đường dài cục bộ P_2^k và sự kết hợp của các đặc trưng đường ngăn cục bộ cùng các đặc trưng chưa qua xử lý R_k được sử dụng mà không có ngoại lệ bởi đơn vị nén.

3.3. Đơn vị nén

Cơ chế nén được hình thành bằng cách tận dụng lớp chập 1×1 . Cụ thể, các đầu ra của đơn vị tăng cường được chuyển tới lớp chập 1×1 , có vai trò hoạt động như

giảm kích thước hoặc chất lọc thông tin liên quan cho mạng sau này. Do đó, đơn vị nén có thể được xây dựng như sau:

$$B_k = f_F^k(P^k) = \alpha_F^k(W_F^k(P^k)) \quad (12)$$

Trong đó f_F^k biểu thị cho hàm lớp tích chập 1×1 (α_F^k biểu thị cho hàm kích hoạt và W_F^k là các thông số trọng số).

4. THỰC NGHIỆM VÀ ĐÁNH GIÁ KẾT QUẢ

Hình ảnh vồng mạc mắt người sử dụng cho thực nghiệm được lấy từ cơ sở dữ liệu DRIVE [19] công khai (nguồn ảnh vồng mạc kỹ thuật số). Hình ảnh có kích thước 565×584 pixel, 8 bit cho mỗi kênh màu sắc, định dạng nén *.TIFF. Hình ảnh ban đầu được bắt từ một nonmydriatic 3 thiết bị tích điện kép Canon CR5 (CCD) camera tại 45° trường nhìn (FOV). Theo [4, 11, 13, 14], tác giả sử dụng 91 hình ảnh từ Yang cùng cộng sự [13] và 200 hình ảnh từ Bộ Dữ liệu Phân đoạn Berkeley (BSD) [18] làm dữ liệu đào tạo. Ảnh thực nghiệm được giảm tỷ lệ bằng phép nội suy hai chiều để tạo các cặp hình ảnh độ phân giải thấp/độ phân giải cao cho cả tập dữ liệu đào tạo và thử nghiệm.

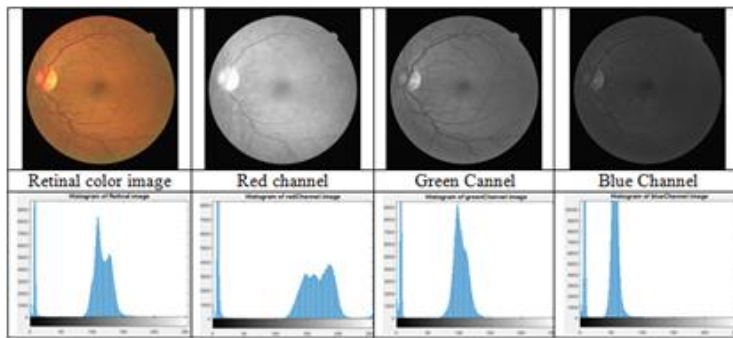
Mạng nơron trí tuệ nhân tạo phát hiện dư ảnh từ độ chói của hình ảnh màu. Kênh độ chói của hình ảnh - Y thể hiện độ sáng của từng pixel thông qua sự kết hợp tuyến tính của các giá trị pixel đỏ, lục và lam. Ngược lại, hai kênh sắc độ của một hình ảnh, C_b và C_r , là các kết hợp tuyến tính khác nhau của các giá trị pixel đỏ, lục và lam thể hiện thông tin sai lệch màu. Mạng nơron trí tuệ nhân tạo được đào tạo chỉ sử

dụng kênh độ chói vì nhận thức của mắt con người nhạy cảm hơn với những thay đổi về độ sáng so với thay đổi về màu sắc.

4.1. Histogram của ảnh võng mạc và các kênh màu

Biểu đồ này hiển thị số lượng pixel trong một hình ảnh ở từng giá trị cường độ khác nhau được tìm thấy trong hình ảnh đó, từ

biểu đồ này, có thể tìm hiểu mức độ phơi sáng hình ảnh tốt hơn nhiều so với việc nhìn vào hình ảnh này trên màn hình máy tính lớn. Nếu phơi sáng là không tối ưu, ngay lập tức thấy cách cải thiện nó từ biểu đồ hình ảnh. Mục đích để lựa chọn ảnh phù hợp cho thực nghiệm. Kênh độ chói Y chính là biểu đồ Red channel được thể hiện trên hình 4.



Hình 4. Ảnh võng mạc và các kênh màu

4.2. Chi tiết thực nghiệm

Phần này sẽ đề cập các bước tiến hành thực nghiệm siêu phân giải dùng trí tuệ nhân tạo nâng cao chất lượng ảnh võng mạc mắt người.

Để thực hiện siêu phân giải hình ảnh đơn (SISR) bằng mạng đề xuất, các bước thực hiện như sau:

B1: Tạo một hình ảnh độ phân giải thấp mẫu từ một hình ảnh tham chiếu độ phân giải cao.

B2: Thực hiện SISR trên hình ảnh có độ phân giải thấp bằng cách sử dụng phép nội suy bicubic, một giải pháp xử lý hình ảnh truyền thống không dựa vào việc học sâu.

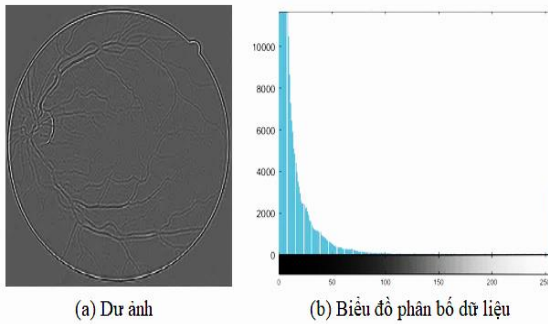
B3: Thực hiện SISR trên hình ảnh độ phân giải thấp bằng cách sử dụng mạng nơron đề xuất.

B4: Trực quan so sánh các hình ảnh độ phân giải cao được xây dựng lại bằng cách sử dụng phép nội suy bicubic, DWT [1,2,3], SRCNN [5], VDSR [4], và IDN.

B5: Đánh giá chất lượng của hình ảnh siêu phân giải bằng cách định lượng độ tương tự của hình ảnh với hình ảnh tham chiếu có độ phân giải cao qua các tham số tính toán chất lượng ảnh RMSE, PSNR, Entropi, SC, NIQE, SSIM. Cuối cùng, thực hiện đánh giá hiệu suất về thời gian thực thi giữa các phương pháp.

4.3. Phân tích và đánh giá kết quả thực nghiệm

Phần này sẽ so sánh trực quan và định lượng các kết quả thực nghiệm đã tiến hành với các phương giới thiệu và phương pháp đề xuất.



Hình 5. Dữ ảnh và sự phân bố dữ liệu của ảnh võng mạc

Các tham số tính toán chất lượng ảnh:

Tính toán Entropi:

$$H = -\sum_k p_k \log(p_k) \tag{13}$$

trong đó K là số lượng các mức xám và p_k là xác suất được kết hợp với mức xám k .

Tính toán RMSE:

$$RMSE = \sqrt{\sum \sum \frac{[R(i, j) - F(i, j)]^2}{MN}} \tag{14}$$

Trong đó i và j biểu thị vị trí không gian của pixel trong khi M và N là kích thước của ảnh.

Tính toán PSNR:

$$PSNR = 10 \log_{10} \left[\frac{(2^n - 1)^2}{MSE} \right] \tag{15}$$

$$MSE = \sum_{i=1}^M \sum_{j=1}^N \frac{[I(i, j) - F(i, j)]^2}{M \times N}$$

$I(i, j)$: ảnh gốc, $F(i, j)$: ảnh hợp nhất (fused image);

$M \times N$: kích thước ảnh I .

Tính toán SC:

$$SC = \frac{\sum_{i=1}^M \sum_{j=1}^N [f(i, j)]^2}{\sum_{i=1}^M \sum_{j=1}^N [f'(i, j)]^2} \tag{16}$$

$f(i, j)$: ảnh gốc; $f'(i, j)$: ảnh hợp nhất (fused image);

$M \times N$: kích thước ảnh f .

Tính toán NIQE [16]:

NIQE đo khoảng cách giữa các tính năng dựa trên NSS được tính toán từ hình ảnh A đến các tính năng thu được từ cơ sở dữ liệu hình ảnh được sử dụng để huấn luyện mô hình. Các tính năng được mô hình hóa như các bản phân phối Gaussian đa chiều.

Tính toán SSIM [17]:

$$SSIM(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma \tag{17}$$

trong đó,

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}$$

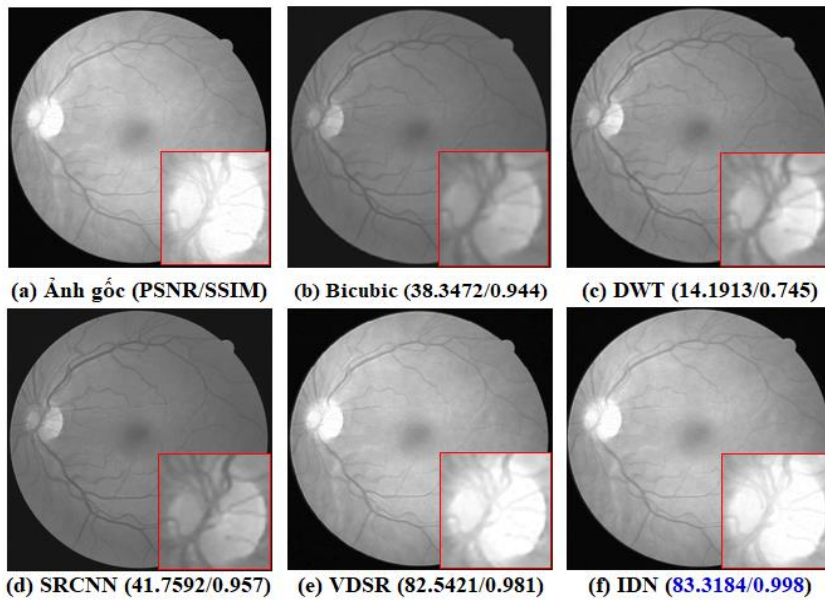
$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$$

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}$$

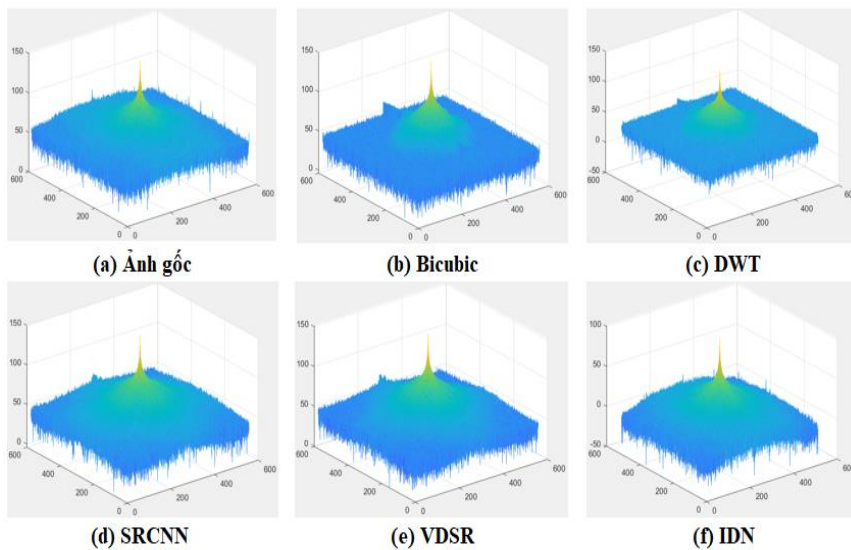
trong đó $\mu_{x,y}$, σ_x , σ_y và σ_{xy} là trung bình cục bộ, độ lệch chuẩn và hiệp phương sai cho hình ảnh x, y . Nếu $\alpha = \beta = \gamma = 1$, và $C_3 = C_2/2$ (lựa chọn mặc định của C_3) sự đơn giản hóa chỉ số đề:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_x\sigma_y + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{18}$$

So sánh trực quan ảnh được xử lý và biểu đồ mật độ phổ năng lượng:



Hình 6. So sánh trực quan ảnh võng mạc



Hình 7. Mật độ phổ năng lượng

Từ hình 6, chúng ta có thể thấy rằng kết quả trực quan theo phương pháp đề xuất thể hiện chất lượng hình ảnh tốt nhất. Ảnh khôi phục với phương pháp đề xuất cho độ nhạy sáng tốt hơn và khả năng quan sát rõ ràng hơn các chi tiết ảnh.

Theo kết quả mô phỏng trong hình 7, đối với mật độ phổ năng lượng, ảnh chất

lượng thấp sẽ cho một mật độ năng lượng quang phổ phẳng. Do đó, từ các kết quả PSD trên, dễ dàng thấy rằng phương pháp đề xuất cho chất lượng xử lý tốt nhất bởi vì PSD càng lớn cho thấy kết quả tăng cường ảnh càng tốt.

So sánh định lượng qua các tham số tính toán chất lượng ảnh.

Bảng 1. Kết quả định lượng chất lượng ảnh võng mạc

Phương pháp	Bicubic	DWT	SRCNN	VDSR	IDN
RMSE	3.074901084	49.7709469	2.082487134	0.019029733	0.006824656
PSNR	38.37418062	14.19128554	41.75915707	82.54214967	83.31838486
Entropi	5.341462778	4.435734623	5.299262245	6.519411707	7.272545443
SC	0.998209422	0.249589283	1.000384552	0.994228197	0.199988018
NIQE	5.004336182	5.313397536	5.069672023	5.229128662	4.163979766
SSIM	0.944457499	0.745228683	0.957719795	0.981807407	0.998874777
Time	-	0.1597907	22.0877573	0.0333412	0.0218909

Chúng ta thấy, kết quả tính toán định lượng chất lượng ảnh võng mạc được xử lý với màu xanh chỉ ra mức hiệu suất tốt nhất và màu đỏ chỉ ra mức hiệu suất tốt thứ hai.

5. KẾT LUẬN

Trong bài báo này, tác giả đề xuất một phương pháp mới dùng trí tuệ nhân tạo dựa trên mạng nơron, sử dụng các khối chất lọc để trích xuất dần nguồn đặc trưng phong phú và hiệu quả nhằm mục đích tái tạo hình ảnh độ phân giải cao. Phương

pháp đề xuất thu được kết quả có hiệu suất rất tốt về RMSE, PSNR, Entropi, SC, NIQE, SSIM và cho thời gian thực thi rất nhanh so với các phương pháp được giới thiệu. Mạng nhỏ gọn này sẽ được áp dụng rộng rãi hơn trong thực tế. Trong tương lai, phương pháp siêu phân giải hình ảnh này sẽ được nghiên cứu để hỗ trợ vấn đề về phục hồi hình ảnh khác như giảm nhiễu và giảm thiểu hiện tượng sai khác giữa ảnh gốc và ảnh sau xử lý (hiện tượng compression artifact).

TÀI LIỆU THAM KHẢO

- [1] Daubechies, Ten lectures on wavelets, CBMS-NSF conference series in applied mathematics. SIAM Ed, 1992.
- [2] Mallat, S. "A theory for multiresolution signal decomposition: the wavelet representation," IEEE Pattern Anal. and Machine Intell., vol. 11, no. 7, pp. 674–693, 1989.
- [3] Meyer. Y, Ondelettes et opérateurs, Tome 1, Hermann Ed, 1990 (English translation: Wavelets and operators, Cambridge Univ. Press. 1993).
- [4] J. Kim, J.K. Lee, and K.M. Lee. Accurate image super-resolution using very deep convolutional networks. In CVPR, pp. 1646–1654, 2016.
- [5] C. Dong, C.C. Loy, and X. Tang. Accelerating the super-resolution convolutional neural network. In ECCV, pp. 391–407, 2016.
- [6] J. Kim, J.K. Lee, and K.M. Lee. Deeply-recursive convolutional network for image super-resolution. In CVPR, pp. 1637–1645, 2016.

- [7] C. Dong, C.C. Loy, K. He, and X. Tang. Learning a deep convolutional network for image super-resolution. In ECCV, pp. 184–199, 2014.
- [8] C. Dong, C.C. Loy, K. He, and X. Tang. Image super-resolution using deep convolutional networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(2):295–307, 2016.
- [9] W. Shi, J. Caballero, F. Huszar, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In CVPR, pp. 1874–1883, 2016.
- [10] X.-J. Mao, C. Shen, and Y.-B. Yang. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. In NIPS, 2016.
- [11] Y. Tai, J. Yang, and X. Liu. Image super-resolution via deep recursive residual network. In CVPR, pp. 3147–3155, 2017.
- [12] M.S.M. Sajjadi, B. Scholkopf, and M. Hirsch. Enhancenet: Single image super-resolution through automated texture synthesis. In ICCV, pp. 4491–4500, 2017.
- [13] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang. Deep laplacian pyramid networks for fast and accurate super-resolution. In CVPR, pp. 624–632, 2017.
- [14] Y. Tai, J. Yang, X. Liu, and C. Xu. Memnet: A persistent memory network for image restoration. In ICCV, pp. 3147–3155, 2017.
- [15] J. Yang, J. Wright, T.S. Huang, and Y. Ma. Image super-resolution via sparse representation. IEEE Transactions on Image Processing, 19(11):2861–2873, 2010.
- [16] Mittal, A., R. Soundararajan, and A.C. Bovik. "Making a Completely Blind Image Quality Analyzer." IEEE Signal Processing Letters. Vol. 22, Number 3, pp. 209–212, Mar. 2013.
- [17] Zhou, W., A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. "Image Quality Assessment: From Error Visibility to Structural Similarity." IEEE Transactions on Image Processing. Vol. 13, Issue 4, pp. 600–612, Apr. 2004.
- [18] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In CVPR, pp. 416–423, 2001.
- [19] DRIVE database. Article (CrossRef Link).

Giới thiệu tác giả:



Tác giả Doãn Thanh Bình tốt nghiệp đại học ngành kỹ thuật điện tử năm 2008, nhận bằng Thạc sĩ năm 2010, nhận bằng Tiến sĩ ngành kỹ thuật điện tử năm 2018 tại Trường Đại học Bách khoa Hà Nội. Tác giả hiện đang công tác tại Phòng Khảo thí và Kiểm định chất lượng, Trường Đại học Điện lực.

Lĩnh vực nghiên cứu: lý thuyết nghịch đảo suy rộng (generalized inverses), GSVD nhằm mô hình hóa, đánh giá các hệ thống MIMO; nghiên cứu các hệ thống thông tin trải phổ đa sóng mang (multi-carrier) áp dụng cho các hệ thống thông tin thế hệ tiếp theo.

