

NGHIÊN CỨU ĐÁNH GIÁ MÔ HÌNH NAFNET TRONG KHỬ MỜ VÀ KHỬ NHIỄU ẢNH

AN EVALUATION STUDY OF THE NAFNET MODEL FOR IMAGE DEBLURRING AND DENOISING

Lương Thị Thảo Hiểu, Dương Xuân Tùng

Khoa Công nghệ thông tin, trường Đại học Kinh tế - Kỹ thuật Công nghiệp

Đến tòa soạn ngày 20/01/2026, chấp nhận đăng ngày 05/3/2026

Tóm tắt: Khôi phục ảnh là một bài toán quan trọng trong lĩnh vực xử lý ảnh và thị giác máy tính, có nhiều ứng dụng trong giám sát an ninh, y tế, và các hệ thống thị giác thông minh. Gần đây, các mô hình học sâu đã đạt được những cải thiện đáng kể trong khôi phục chất lượng ảnh. Trong số đó, NAFNet là một kiến trúc mới đơn giản và hiệu quả với thiết kế loại bỏ hàm kích hoạt phi tuyến, nhằm giảm độ phức tạp tính toán và cải thiện khả năng hội tụ. Bài báo này tập trung đánh giá hiệu năng của NAFNet với một số mô hình khử mờ và khử nhiễu ảnh tiên tiến khác như DeepDeblur, SRNNNet, MPRNet, thực nghiệm trên bộ dữ liệu chuẩn với các chỉ số đánh giá PSNR và SSIM. Kết quả cho thấy NAFNet đạt hiệu năng cạnh tranh, bảo toàn chi tiết ảnh tốt và hội tụ nhanh trong quá trình huấn luyện.

Từ khóa: NAFNet, khôi phục ảnh, khử mờ, khử nhiễu, học sâu.

Abstract: Image restoration is a fundamental problem in image processing and computer vision, with applications in security surveillance, medical imaging, and intelligent vision systems. Recent deep learning-based approaches have achieved significant improvements. Among them, NAFNet is a recently proposed architecture that removes nonlinear activation functions to achieve a simple and efficient design. This paper evaluates the performance of NAFNet and compares it with several state-of-the-art image restoration models, including DeepDeblur, SRNNNet, and MPRNet, experimental results are obtained on benchmark datasets using PSNR and SSIM as evaluation metrics. The results indicate that NAFNet achieves superior restoration quality, preserves fine image details, converges faster during training, and demonstrates competitive overall performance compared to existing methods.

Keywords: NAFNet, Image Restoration, Image Deblurring, Image Denoising, Deep Learning.

1. GIỚI THIỆU

Các phương pháp khôi phục ảnh truyền thống như BM3D [1] thường phụ thuộc mạnh vào giả định thống kê, khó thích ứng trong điều kiện suy biến phức tạp. Sự phát triển của học sâu đã mở ra hướng tiếp cận mới, trong đó những mô hình dựa trên mạng nơ-ron tích chập (CNN) như DnCNN [2], UNet [3] đã đạt hiệu quả đáng kể trong các bài toán khử nhiễu và khử mờ ảnh. Gần đây, các kiến trúc tiên tiến

hơn dựa trên Transformer, tiêu biểu là Restormer [4], tiếp tục cải thiện chất lượng khôi phục ảnh trong các kịch bản suy giảm phức tạp. Tuy nhiên, các mô hình này có kiến trúc phức tạp, số lượng tham số lớn và đòi hỏi tài nguyên tính toán cao, điều này làm hạn chế khả năng ứng dụng trong thực tế.

Gần đây, NAFNet được đề xuất như một backbone khôi phục ảnh tối giản, loại bỏ các thành phần như normalization và hàm kích

hoạt phi tuyến, nhằm hướng tới thiết kế đơn giản nhưng vẫn đạt hiệu năng cạnh tranh [5]. Tuy nhiên, vẫn còn thiếu các nghiên cứu đánh giá độc lập về hiệu năng của NAFNet so với các mô hình học sâu phổ biến khác. Do đó, mục tiêu của bài báo này, thực hiện nghiên cứu đánh giá và so sánh NAFNet với các mô hình DeepDeblur, SRNNet và MPRNet, nhằm làm rõ ưu điểm và hạn chế của từng phương pháp trong bài toán khử mờ và khử nhiễu ảnh

2. CÁC NGHIÊN CỨU LIÊN QUAN

Các nghiên cứu ban đầu về khôi phục ảnh dựa trên học sâu chủ yếu sử dụng các kiến trúc CNN truyền thống. ResNet được He *et al.* [6] giới thiệu với cơ chế kết nối (skip connection) tắt nhằm cải thiện khả năng huấn luyện mạng sâu và học ánh xạ dư (residual learning) giữa ảnh suy giảm và ảnh gốc. Trong khi đó, UNet với cấu trúc encoder-decoder và các kết nối skip đối xứng giúp bảo toàn thông tin không gian, ứng dụng rộng rãi trong các bài toán khôi phục ảnh. Tuy nhiên, các kiến trúc này còn hạn chế trong việc xử lý mờ không đồng nhất do thiếu khả năng khai thác ngữ cảnh đa tỉ lệ và dài hạn.

DeepDeblur [7] và SRNNet [8] đã chứng minh hiệu quả trong xử lý mờ chuyển động thông qua kiến trúc đa tỉ lệ và hồi quy. MPRNet [9] mở rộng hướng tiếp cận này bằng cách thực hiện khôi phục ảnh theo nhiều giai đoạn liên tiếp, giúp cải thiện chất lượng tái tạo nhưng đồng thời làm gia tăng độ phức tạp tính toán.

Song song đó, các mô hình dựa trên Transformer như SwinIR [10] và Restormer [4] khai thác cơ chế self-attention để học các phụ thuộc dài hạn trong ảnh, đạt hiệu năng cao trong nhiều bài toán khôi phục ảnh. Tuy nhiên, các mô hình này thường yêu cầu tài nguyên tính toán lớn, gây khó khăn cho việc

triển khai trong các hệ thống thực tế.

Khác với xu hướng gia tăng độ phức tạp kiến trúc, NAFNet được thiết kế theo hướng tối giản, tập trung vào các phép biến đổi tuyến tính và cơ chế học dư (residual learning). Bài báo này tập trung đánh giá hiệu quả của cách tiếp cận đơn giản này, thông qua so sánh thực nghiệm với các mô hình tiêu biểu trên.

3. MÔ HÌNH TOÁN HỌC CỦA NAFNET VÀ HÀM SAI SỐ

3.1. Mô hình suy biến ảnh

Trong nghiên cứu này, mục tiêu là ước lượng ảnh gốc không suy giảm $X \in \mathbb{R}^{H \times W \times C}$ từ ảnh quan sát bị suy giảm Y , trong đó Y bị ảnh hưởng bởi mờ chuyển động, nhiễu và các biến dạng phi tuyến. Quá trình suy giảm được mô hình hóa tổng quát như sau:

$Y = D(X) + N$. Trong đó $D(\cdot)$ là toán tử suy giảm và N biểu diễn nhiễu.

3.2. Mô hình toán học của NAFNet

NAFNet học một ánh xạ tham số hóa $F(\cdot; \theta)$ nhằm xấp xỉ ảnh gốc, $X = F(Y; \theta)$ với θ là tập tham số của mạng.

NAFNet sử dụng kiến trúc encoder-decoder đối xứng, bao gồm các tầng tích chập ban đầu, tiếp theo đó là các khối NAFBlock làm đơn vị cơ bản, đây là khối chịu trách nhiệm học ánh xạ đặc trưng: các phép biến đổi tuyến tính và chuẩn hóa được sử dụng thay cho các hàm kích hoạt phi tuyến truyền thống, các kết nối tắt, được sử dụng nhằm bảo toàn thông tin. Toàn bộ mạng hoạt động theo cơ chế học ánh xạ dư (residual learning): Trong đó $G(\cdot)$ biểu diễn phần dư mà mạng cần học, giúp cải thiện độ ổn định huấn luyện và khả năng hội tụ. Thiết kế này giúp mô hình tối ưu hơn và giảm chi phí tính toán.

Mỗi khối NAFBlock mô tả như sau:

Đầu vào của một NAFBlock là tensor đặc trưng $Z \in R^{H \times W \times C}$. Một NAFBlock bao gồm hai nhánh residual liên tiếp, mỗi nhánh được thiết kế theo nguyên tắc tuyến tính tối giản.

Nhánh thứ nhất được mô tả như sau:

$$Z_1 = Z + F_1(Z) \quad (1)$$

trong đó $F_1(\cdot)$ là tổ hợp các phép biến đổi tuyến tính:

$$F_1(Z) = \text{Conv}_{1 \times 1}(\text{DWConv}_{3 \times 3}(\text{Conv}_{1 \times 1}(Z))) \quad (2)$$

Ở đây:

- Conv1×1 đóng vai trò trộn kênh
- DWConv3×3 là depth-wise convolution giúp giảm đáng kể số tham số

Nhánh residual thứ hai tiếp tục tinh chỉnh đặc trưng: $Z_2 = Z_1 + F_2(Z_1)$ (3)

Trong đó: $F_2(Z_1) = \text{Conv}_{1 \times 1}(\text{Conv}_{1 \times 1}(Z_1))$ (4)

Hai nhánh residual liên tiếp cho phép NAFBlock học các biến đổi phức tạp trong khi vẫn giữ cấu trúc tuyến tính đơn giản. Việc loại bỏ chuẩn hóa và kích hoạt phi tuyến giúp NAFNet tránh hiện tượng suy giảm biên độ đặc trưng, giảm chi phí tính toán và bộ nhớ, đảm bảo đạo hàm ổn định thông qua các kết nối dư.

3.3. Hàm sai số

3.3.1. Hàm sai số pixel L_2 (Mean square error)

Được tính là sai số bình phương giữa ma trận pixel của hình ảnh dự đoán và hình ảnh đích nhưng L_2 không thể biểu diễn hoàn toàn chất lượng của quá trình tái tạo [11]. Công thức như sau: $L_2(\hat{I}, I_{gt}) = \frac{1}{N} \sum_{i=1}^N (\hat{I}_i - I_{gt,i})^2$ (5)

Trong đó:

I_i : giá trị dự đoán;

$I_{gt,i}$: giá trị thực tế (Ground Truth- GT);

N : tổng số pixel.

3.3.2. Hàm sai số Gradient Loss (L_1) [12]

Đo sai số cấu trúc, tập trung vào biên (edges) và hướng chuyển động, giúp biên sắc nét hơn, giữ chi tiết mảnh, bằng cách so sánh đạo hàm (gradient) của ảnh, không so sánh trực tiếp pixel, nhằm khắc phục hiện tượng ảnh bị mờ bằng cách so sánh gradient giữa ảnh dự đoán và ground truth.

$$L_{grad} = \left\| \nabla \hat{I} - \nabla I_{gt} \right\|_1 \quad (6)$$

3.3.3. Hàm sai số Charbonnier Loss [13]

$$L_{char} = \frac{1}{N} \sum \sqrt{(\hat{I} - I_{gt})^2 + \epsilon^2} \quad (\epsilon = 10^{-3}) \quad (7)$$

Giữ chi tiết tốt hơn L2, đạo hàm ổn định gần 0, với ϵ rất nhỏ để đảm bảo khả vi tại 0

4. KIẾN TRÚC MẠNG NAFNET

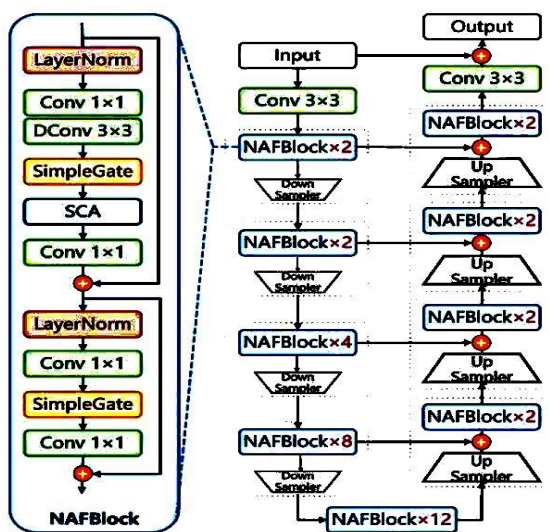
NAFNet được xây dựng theo kiến trúc encoder-decoder đối xứng, tương tự UNet, nhưng điểm khác biệt là toàn bộ khối tích chập truyền thống được thay thế bằng NAFBlock. Việc sử dụng NAFBlock giúp giảm độ phức tạp tính toán, tăng tốc độ huấn luyện, và vẫn đảm bảo chất lượng phục hồi ảnh cao.

Các thành phần chính:

- Lớp Convolution đầu tiên: Ảnh đầu vào I_{LQ} có kích thước $H \times W \times 3$ (Height $\times$ Width $\times$ Channels) được đưa qua một lớp Conv 3×3 để trích xuất đặc trưng cơ bản. Số kênh được mở rộng từ 3 lên C_0 có Width bằng 64 kênh. Đây là lớp nền tảng cho các tầng sâu hơn.
- Khối Encoder: Encoder bao gồm nhiều NAFBlock liên tiếp để khai thác đặc trưng.

Cứ sau một số tầng, mô hình sử dụng phép downsampling (giảm độ phân giải) bằng cách dùng Conv 4x4 với stride=2. Kích thước đặc trưng giảm một nửa ($H/2, W/2$) trong khi số kênh tăng gấp đôi ($C \rightarrow 2C$). Các đặc trưng ở từng mức của encoder sẽ được lưu để ghép nối với decoder theo dạng skip-connection.

- **Khối Bottleneck:** Đây là tầng có độ trừu tượng cao nhất, giữ đặc trưng toàn cục của ảnh. Nhiều NAFBlock được xếp chồng tại đây để xử lý cấu trúc ảnh phức tạp, loại bỏ nhiễu, và học biểu diễn quan hệ giữa các vùng ảnh.
- **Khối Decoder:** Sử dụng các phép upsampling (tăng độ phân giải) để khôi phục kích thước không gian. Mỗi tầng decoder sẽ cộng đặc trưng từ encoder cùng mức, khôi phục chi tiết bị mất. Điều này giúp ảnh sau tái tạo giữ được cả thông tin toàn cục lẫn chi tiết cục bộ.
- **Lớp Conv cuối:** Sau quá trình khôi phục, đặc trưng đưa qua Conv 3x3 để dự đoán phần dư ($I_{Residual}$). Ảnh đầu ra đã được khôi phục I_{SR} được tính bằng cách cộng phần dư này với ảnh đầu vào ($I_{SR} = I_{LQ} + I_{Residual}$). Phương pháp học phần dư giúp mô hình tập trung vào việc học sự khác biệt (nhiều/mờ) thay vì học lại toàn bộ bức ảnh, giúp ổn định và tăng tốc độ hội tụ.



Hình 1. Kiến trúc mạng NAFNet

5. THỰC NGHIỆM VÀ ĐÁNH GIÁ

5.1. Bộ dữ liệu

Các thí nghiệm được thực hiện trên các bộ dữ liệu chuẩn nhằm đánh giá hiệu quả của mô hình trong hai bài toán khôi phục ảnh: khử mờ và khử nhiễu. Đối với khử mờ, chúng tôi sử dụng bộ dữ liệu GoPro [14], gồm 3.214 cặp ảnh mờ - sắc nét được tổng hợp từ video thực tế, với độ phân giải 1280×720 , bao quát nhiều bối cảnh khác nhau: cảnh trong nhà, ngoài trời và các điều kiện chiếu sáng đa dạng. Trong mỗi cặp: ảnh sắc nét được xem như ground truth, ảnh mờ được tạo ra do chuyển động của camera hoặc đối tượng.

Đối với khử nhiễu, các thí nghiệm được tiến hành trên bộ dữ liệu SIDD [15] (Smartphone Image Denoising Dataset). Chúng tôi sử dụng phiên bản SIDD Medium, gồm 320 cặp ảnh nhiễu - sạch, không gian màu RGB, được chụp từ nhiều thiết bị smartphone khác nhau dưới các điều kiện chụp đa dạng. Ảnh ground truth được tạo bằng cách trung bình nhiều lần chụp liên tiếp. Hiệu năng khôi phục ảnh được đánh giá bằng các chỉ số PSNR và SSIM.

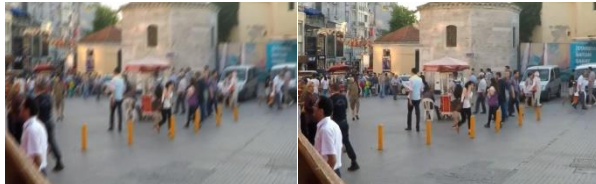
5.2. Tiền xử lý dữ liệu

Đối với cả hai bộ dữ liệu, tất cả các ảnh được chuyển sang không gian màu RGB và chuẩn hóa về khoảng $[0,1][0,1][0,1]$ bằng cách chia cho 255, phù hợp với thiết lập đầu vào và cơ chế Layer Normalization trong NAFNet.

- Bài toán khử mờ trên bộ dữ liệu GoPro, các ảnh có độ phân giải 1280×720 được cắt thành các patch kích thước 256×256 nhằm giảm chi phí bộ nhớ và tăng số lượng mẫu huấn luyện. Các cặp patch mờ - sắc nét được căn chỉnh chính xác để đảm bảo ánh xạ đầu vào - đầu ra trong quá trình học.
- Bài toán khử nhiễu trên bộ dữ liệu SIDD, các cặp ảnh nhiễu - sạch được xử lý chuẩn hóa

và cắt patch, đồng thời áp dụng các phép data augmentation chuẩn của NAFNet, bao gồm lật và xoay ảnh theo các góc cố định. Các patch sau tiền xử lý được tổ chức thành mini-batch với kích thước từ 16 đến 32.

5.3. Thiết lập huấn luyện



a) Ảnh mờ

b) Ảnh sau khử mờ



a) Ảnh nhiễu

b) Ảnh sau khử nhiễu

Hình 2. Kết quả huấn luyện ảnh

Huấn luyện NAFNet gồm các bước từ chuẩn bị dữ liệu, khởi tạo tham số, tính toán hàm mất mát cho đến cập nhật trọng số. Mục tiêu là khôi phục ảnh I_{SR} gần giống với ảnh gốc I_{GT} .

Mã giả: Quá trình huấn luyện

Đầu vào: Ảnh gốc - GT, và ảnh mờ hoặc ảnh nhiễu.

Đầu ra: Mô hình khử mờ hoặc khử nhiễu được huấn luyện, và ảnh đã được xử lý khử mờ hoặc khử nhiễu.

Các bước chi tiết:

a. Khởi tạo mô hình: Mạng NAFNet được khởi tạo với tham số ngẫu nhiên θ . Siêu tham số cần thiết bao gồm batch size, số epoch, learning rate, hệ số λ . Bộ tối ưu sử dụng là AdamW thay Adam truyền thống. AdamW giúp đảm bảo quá trình huấn luyện ổn định

hơn và đạt hiệu năng cao hơn cho các tác vụ khôi phục ảnh

b. Vòng lặp huấn luyện:

For epoch = 1 **to** epoch_max **do**:

For mỗi batch dữ liệu (I_{GT} , I_{LQ}):

Bước 1: Sinh ảnh dự đoán

$I_{pred} = \text{NAFNet}(I_{LQ}; \theta)$

Bước 2: Tính toán hàm mất mát

$$L_{char} = \frac{1}{N} \sum \sqrt{(\hat{I} - I_{g,t})^2 + \varepsilon^2} (\varepsilon = 10^{-3})$$

$$Loss_{total} = 0.1L_{char}$$

Bước 3: Cập nhật tham số bằng Gradient Descent

$\theta_{new} \leftarrow \text{Update}(\theta_{old}, \nabla Loss_{total}, m_t, v_t, \lambda)$

End

Bước 4: Đánh giá mô hình

Tính PSNR, SSIM trên tập kiểm tra, lưu lại trọng số tốt nhất

End

c. Sau mỗi epoch, chỉ số PSNR, SSIM được tính trên tập kiểm tra. Mô hình đạt giá trị cao nhất sẽ được lưu lại để sử dụng.

6. KẾT QUẢ VÀ THẢO LUẬN

Thực nghiệm huấn luyện NAFNet trên hai bộ dữ liệu chuẩn GoPro và SIDD, số epoch 52 (tương đương với 450.000 iteration), batch size 4, learning rate = $2e-4$. Hiệu năng của các mô hình được đánh giá bằng hai chỉ số phổ biến: PSNR (Peak Signal-to-Noise Ratio), SSIM (Structural Similarity Index), phản ánh lần lượt mức độ phục hồi tín hiệu và khả năng bảo toàn cấu trúc ảnh.

▪ PSNR: Đo lường độ tương phản giữa tín hiệu gốc và tín hiệu bị suy giảm, thường sử dụng để đánh giá chất lượng ảnh. Giá trị PSNR cao đồng nghĩa với việc ảnh tái tạo có độ méo thấp so với ảnh gốc. Theo dõi PSNR theo thời gian giúp đánh giá quá trình học của

mô hình, nếu PSNR tăng, điều đó có nghĩa mô hình đang dần cải thiện chất lượng ảnh đầu ra. Ngược lại, có thể mô hình đang gặp vấn đề trong việc tái tạo chi tiết ảnh.

▪ SSIM: Đo lường độ tương đồng về mặt cấu trúc. Theo dõi SSIM theo thời gian giúp đánh giá chất lượng ảnh đầu ra của mô hình. Nếu SSIM tăng theo thời gian, điều đó cho thấy mô hình đang dần cải thiện khả năng tái tạo, hình ảnh gần giống với ảnh gốc hơn. Ngược lại, mô hình đang gặp khó khăn trong việc giữ lại các đặc trưng quan trọng của hình ảnh.

Bảng 1. So sánh PSNR/SSIM trên ảnh mờ

Mô hình	PSNR(dB)	SSIM
DeepDeblur	29.08	0.914
SRNNNet	30.26	0.934
MPRNet	32.66	0.959
NAFNet	30.65	0.938

Bảng 1 tổng hợp kết quả so sánh PSNR và SSIM của NAFNet với một số mô hình khử mờ tiêu biểu trên bộ dữ liệu GoPro.

Kết quả cho thấy NAFNet đạt 30.65 dB PSNR và 0.938 SSIM, vượt trội hơn các phương pháp tiền nhiệm như DeepDeblur (29.08 dB / 0.914) và SRNNNet (30.26 dB / 0.934). Điều này thể hiện kiến trúc phi tuyến đơn giản của NAFNet vẫn có khả năng khai thác hiệu quả đặc trưng không gian, cải thiện chất lượng ảnh so với các mô hình end-to-end và recurrent truyền thống.

So với MPRNet, mô hình đa tầng đạt kết quả cao nhất (32.66 dB / 0.959), NAFNet có hiệu năng thấp hơn về mặt định lượng. Tuy nhiên, cần lưu ý rằng MPRNet sử dụng kiến trúc phức tạp với chi phí tính toán lớn, trong khi NAFNet chỉ được huấn luyện và triển khai trên một GPU, cho thấy sự cân bằng hợp lý giữa hiệu năng và độ phức tạp mô hình.

Bảng 2. So sánh PSNR/SSIM trên ảnh nhiễu

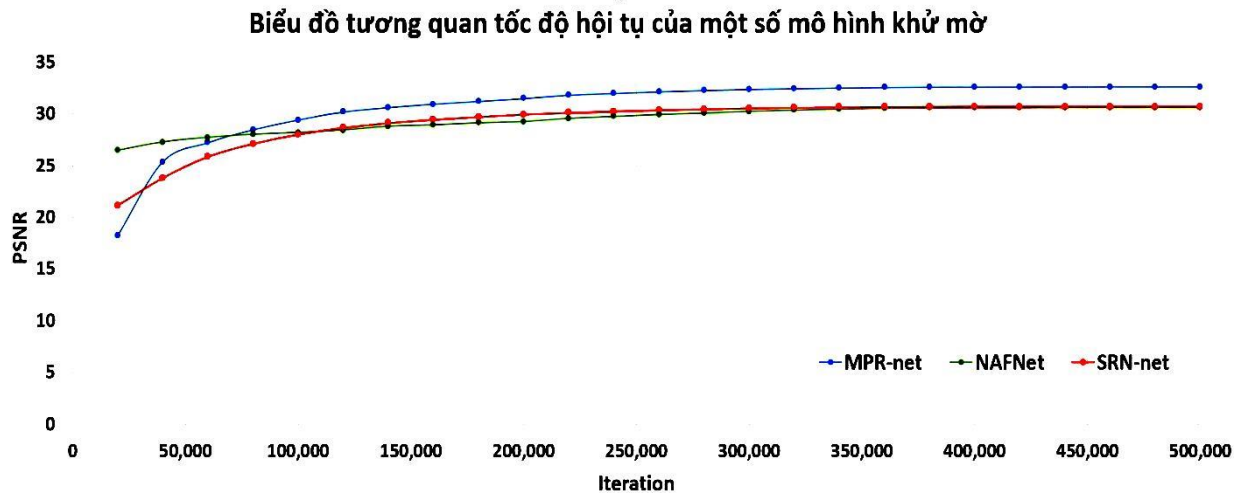
Mô hình	PSNR(dB)	SSIM
MPRNet	39.71	0.958
DnCNN	38.85	0.954
Restormer	40.02	0.96
NAFNet	39.64	0.9119

Bảng 2 tổng hợp kết quả khử nhiễu ảnh trên bộ dữ liệu SIDD.

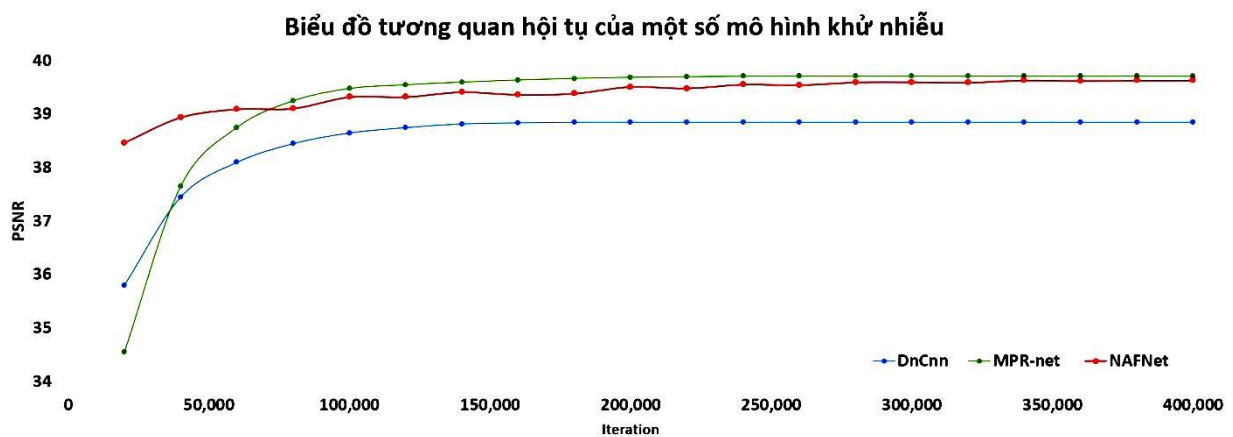
NAFNet đạt 39.64 dB PSNR, cho thấy hiệu năng cạnh tranh so với các mô hình học sâu chuyên biệt cho khử nhiễu như DnCNN (38.85 dB) và MPRNet (39.71 dB).

So với Restormer, mô hình Transformer đạt PSNR cao nhất (40.02 dB), NAFNet có kết quả thấp hơn một mức nhỏ. Tuy nhiên, chỉ số SSIM của NAFNet thấp hơn đáng kể, phản ánh sự khác biệt trong khả năng bảo toàn cấu trúc tinh vi khi xử lý nhiễu phức tạp. Điều này cho thấy NAFNet, với kiến trúc đơn giản và không sử dụng cơ chế attention toàn cục, vẫn đạt hiệu năng phục hồi tín hiệu tốt nhưng còn hạn chế trong việc mô hình hóa mối quan hệ dài hạn giữa các vùng ảnh trong bài toán khử nhiễu.

Quan sát Hình 3, ngay từ những vòng lặp đầu tiên, NAFNet có điểm xuất phát PSNR cao nhất (khoảng 26.5 dB) so với MPRNet và SRNNNet. Điều này cho thấy mô hình có khả năng học các đặc trưng hình ảnh rất nhanh ở giai đoạn đầu. Trong những vòng lặp tiếp theo, NAFNet duy trì đà tăng ổn định, không ghi nhận các dao động lớn, điều này chứng tỏ khả năng tối ưu hóa hiệu quả và nhất quán. Trong những vòng lặp cuối, mặc dù MPRNet vượt lên dẫn đầu sau khoảng 100.000 iterations và đạt mức cao nhất (xấp xỉ 33 dB), nhưng NAFNet vẫn giữ vị trí ổn định ở mức khoảng 30 dB, tương đương với SRNNNet.



Hình 3. Biểu đồ tương quan hội tụ giữa NAFNet và một số mô hình khử mờ



Hình 4. Biểu đồ tương quan hội tụ giữa NAFNet và một số mô hình khử nhiễu

Trong tác vụ khử nhiễu (hình 4), NAFNet thể hiện ưu thế tuyệt đối về chất lượng hình ảnh. NAFNet bắt đầu với mức PSNR cao vượt trội (khoảng 38.5 dB) so với hai mô hình còn lại. Sau đó, NAFNet duy trì vị trí dẫn đầu trong suốt quá trình huấn luyện. Trong khi MPRNet cần nhiều thời gian để bắt kịp và DnCNN, dừng lại ở mức PSNR thấp hơn (khoảng 38.8 dB), NAFNet tiệm cận mức 40 dB. NAFNet đạt đến điểm hội tụ (nơi đồ thị nằm ngang) rất nhanh, cho thấy hiệu quả tính toán cao và khả năng xử lý nhiễu tốt.

Như vậy qua 2 đồ thị khử mờ và khử nhiễu, thể hiện NAFNet hoạt động tốt trên cả hai tác vụ khôi phục ảnh (khử mờ và khử nhiễu).

Trong cả hai trường hợp, NAFNet thường bắt đầu với mức PSNR cao hơn các đối thủ, giúp giảm bớt thời gian huấn luyện cần thiết để đạt được kết quả chấp nhận được. NAFNet đặc biệt mạnh trong việc khử nhiễu, mang lại chỉ số PSNR cao nhất và độ ổn định vượt trội so với các kiến trúc truyền thống như DnCNN hay mô hình phức tạp như MPRNet.

7. KẾT LUẬN

Bài báo đã trình bày một nghiên cứu đánh giá thực nghiệm mô hình NAFNet trong bài toán khử mờ và khử nhiễu ảnh, dựa trên các bộ dữ liệu chuẩn và những chỉ số đánh giá được sử dụng phổ biến trong lĩnh vực khôi phục ảnh. Từ các kết quả thực nghiệm, có thể nhận thấy

rằng NAFNet đạt hiệu năng ổn định và cạnh tranh trên cả khử mờ và khử nhiễu ảnh. Mặc dù không luôn đạt kết quả cao nhất về mặt định lượng, nhưng mô hình thể hiện ưu thế rõ rệt về tốc độ hội tụ và tính ổn định trong suốt quá trình huấn luyện.

Thông qua so sánh với các mô hình học sâu tiêu biểu như DeepDeblur, SRNNNet, MPRNet, DnCNN và Restormer, kết quả thực nghiệm cho thấy NAFNet đạt hiệu năng cạnh tranh về cả PSNR và chất lượng thị giác, đồng thời duy trì được sự ổn định và hiệu quả nhờ kiến trúc

đơn giản. Nghiên cứu này làm rõ tiềm năng của NAFNet trong việc đóng vai trò như một mô hình nền tảng đáng tin cậy cho các bài toán khôi phục ảnh trong môi trường thực tế.

Trong tương lai, nghiên cứu sẽ được mở rộng theo hướng đánh giá NAFNet trên các kịch bản suy giảm ảnh phức tạp. Đồng thời nhóm sẽ khảo sát các biến thể tối ưu của NAFNet nhằm phục vụ các ứng dụng thực tiễn chuyên sâu như siêu phân giải ảnh và cải thiện ảnh trong điều kiện ánh sáng yếu.

TÀI LIỆU THAM KHẢO

- [1] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering", *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, 2007.
- [2] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising", *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, 2017.
- [3] O. Ronneberger, P. Fischer, and T. Brox, "UNet: Convolutional networks for biomedical image segmentation", in *Proc. MICCAI*, 2015, pp. 234–241.
- [4] S. Zamir et al., "Restormer: Efficient transformer for high-resolution image restoration", in *Proc. CVPR*, 2022, pp. 5728–5739.
- [5] L. Chen et al., "Simple baselines for image restoration", in *Proc. ECCV*, 2022, pp. 17–33.
- [6] K. He et al., "Deep Residual Learning for Image Recognition", *CVPR*, 2016.
- [7] S. Nah et al., "Deep Multi-scale CNN for Dynamic Scene Deblurring", *CVPR*, 2017.
- [8] X. Tao et al., "Scale-Recurrent Network for Deep Image Deblurring", *CVPR*, 2018.
- [9] S. W. Zamir et al., "Multi-Stage Progressive Image Restoration", *CVPR*, 2021.
- [10] J. Liang et al., "SwinIR: Image Restoration Using Swin Transformer", *ICCV*, 2021.
- [11] Wei-Sheng Lai et al., "Deep Laplacian pyramid networks for fast and accurate super-resolution", *CVPR*, 2017, pp. 624–632.
- [12] Mathieu et al., "Deep Multi-Scale Video Prediction Beyond Mean Square Error", *ICLR* 2016.
- [13] Johnson et al., "Perceptual Losses for Real-Time Style Transfer and Super-Resolution", 2016.
- [14] <https://seungjunna.github.io/Datasets/gopro.html>
- [15] <https://abdokamel.github.io/sidd/>

Thông tin liên hệ:

Lương Thị Thảo Hiếu

Điện thoại: 0942160880 - Email: ltthieu@uneti.edu.vn

Khoa Công nghệ thông tin, Trường Đại học Kinh tế - Kỹ thuật Công nghiệp.