

**Ứng dụng mạng sâu và phép biến đổi Fourier  
trong khử nhiễu ảnh**  
**Application of deep learning and Fourier transformer  
to image denoising**  
Ngô Đức Lưu\*

\* Khoa Công nghệ Thông tin, Trường Đại học Bạc Liêu  
ndluu@blu.edu.vn

Ngày nhận bài:

02/7/2023

Ngày chấp nhận đăng:

15/8/2023

**Keywords:** Machine Learning, Computer vision, Image Denoising, Fourier Transformation, Residual Block, MIMO-Unet, Image noise.

**ABSTRACT**

*Image Denoising has become a topic of great interest in various computer vision fields thanks to its practical applications. Images are often affected by noise due to various factors such as environment, lighting conditions, and camera quality. Denoising helps to increase the accuracy and reliability of these images, thereby improving their analysis, recognition, and classification. Some previous image denoising methods focus on extracting features directly from the image in the spatial domain. However, restoring a clearly image from a noisy image captured by a camera requires more information about the noise frequency. This can be achieved by utilizing the rich representation of noise images using Fourier Transformation, which transforms the image from spatial domain to frequency domain. This study proposes using DeepRFT (Deep Learning Residual Fourier Transformation) to build an image denoising model specifically for Fujifilm camera noise. This solution is based on the use of an improved Fourier Transformation for the Residual block of the MIMO-UNet network. The experimental results show that our denoising model competes well with other denoising approaches in terms of evaluation metrics PSNR in both the self-collected dataset and the SIDD benchmark dataset.*

**TÓM TẮT**

*Khử nhiễu trở thành một trong những chủ đề được quan tâm của lĩnh vực thị giác máy tính nhờ vào khả năng áp dụng vào thực tế, những ảnh thường bị nhiễu do nhiều yếu tố khác nhau như môi trường, điều kiện ánh sáng, chất lượng máy ảnh, các yếu tố khác. Việc khử nhiễu giúp tăng độ chính xác và độ tin cậy của những ảnh này, từ đó cải thiện khả năng phân tích, nhận dạng và phân loại. Một số phương pháp khử nhiễu hình ảnh trước đây đều tập trung rút trích đặc trưng trực tiếp trên ảnh ở miền không gian. Để tái tạo lại hình ảnh sạch từ ảnh nhiễu chụp bởi máy ảnh đòi hỏi cần nhiều thông tin hơn nữa về nhiễu. Tận dụng khả năng biểu diễn phong phú ảnh nhiễu bởi phép biến đổi Fourier từ miền không gian sang miền tần số. Trong nghiên cứu này, chúng tôi đề xuất sử dụng DeepRFT để xây dựng mô hình khử nhiễu, cụ thể ảnh nhiễu được chụp từ máy ảnh Fujifilm. Giải pháp này dựa trên sử dụng phép biến đổi Fourier cải tiến cho khối dư của mạng MIMO-UNet. Kết quả thực nghiệm cho thấy mô hình khử nhiễu sử dụng mạng DeepRFT đạt kết quả cạnh tranh với các cách tiếp cận khử nhiễu khác như Autoencoder, RidNet, DnCNN, MIMO-UNet về chỉ số đánh giá PSNR trên cả tập dữ liệu tự thu thập và tập dữ liệu chuẩn SIDD.*

**Từ khóa :** Máy học, Thị giác máy tính, Khử nhiễu, Phép biến đổi Fourier, Khối dư, Mạng học sâu MIMO-Unet, Ảnh nhiễu.

## 1. Giới thiệu

Trong những năm gần đây, thị giác máy tính (Computer Vision) đã trở thành một trong những chủ đề được nhiều lĩnh vực quan tâm nhờ vào khả năng ứng dụng của chúng vào các bài toán thực tế. Một trong những ứng dụng này đó là khử nhiễu ảnh (Imaging Denoising). Ứng dụng này không chỉ được sử dụng cho việc phục hồi hình ảnh và video trong thực tế mà nó còn liên quan đến việc cải thiện chất lượng các mô hình thị giác máy tính nhờ vào khả năng xử lý mạnh mẽ của chúng. Do đó, có thể nói rằng chủ đề khử nhiễu hình ảnh đã được nghiên cứu rộng rãi và có nhiều cách tiếp cận khác nhau. Mỗi cách tiếp cận đều có lợi ích và bất lợi riêng của nó. Các phương pháp điển hình như lọc miền không gian (Transformation Domain Filtering) và lọc miền chuyển đổi (Spatial Domain Filtering) thì đơn giản hơn, nhanh hơn và dễ sử dụng hơn so với các phương pháp học sâu (Deep Learning), hay một số phương pháp khác. Tuy nhiên, lọc miền không gian chỉ hoạt động hiệu quả khi xử lý nhiễu và lọc miền chuyển đổi có thể gây ra biến dạng điểm ảnh. Mặc dù, phương pháp tiếp cận học sâu phức tạp hơn so với các phương pháp khử nhiễu khác, nhưng phương pháp này đã đạt được kết quả tốt hơn và đang được sử dụng cùng với các phương pháp khác để bù đắp cho những sai sót của chúng.

Gần đây, mô hình sử dụng tập ảnh với nhiều kích thước khác nhau vào một bộ giải mã kết hợp với phương pháp học sâu bắt đầu xuất hiện với hiệu suất ngày càng nhiều. Phương pháp này đã được chứng minh là hiệu quả trên các bài toán khử mờ, khử sương mù hay cả khử nhiễu đều đã đạt được kết quả tốt về độ đo chất lượng ảnh PSNR (Peak Signal-To-Noise Ratio) và SSIM (Structural Similarity Index Measure) khi đầu vào là ảnh đa kích thước (Călin-Adrian Popa và Cosmin Cernazanu, 2018), (Hanssie, 2018), (Mansurov N.,2012), (Healey G.E. và Kondepudy R.,1994). Điều này đã chứng minh tiềm năng của phương pháp học sâu trong thời gian qua. Trong

nghiên cứu này, chúng tôi sử dụng mạng học sâu MIMO-UNet (Hui R, 2020, trang 125 – 154) là mô hình cải tiến của mạng học sâu UNet (Ronneberger O và cộng sự, 2015) và phép biến đổi Fourier để xây dựng mô hình mạng DeepRFT cho việc xử lý nhiễu ảnh ở tập ảnh được lấy từ máy ảnh Fujifilm.

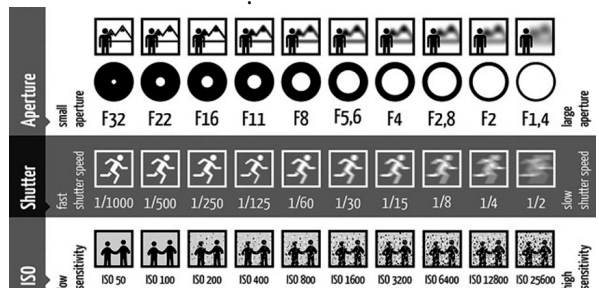
Bài báo này gồm 5 phần. Phần đầu tiên giới thiệu chung về bài báo. Phần thứ hai trình bày những nghiên cứu có liên quan. Phần thứ ba trình bày phương pháp nghiên cứu và mô hình đề xuất. Nội dung tiếp theo trình bày kết quả thực nghiệm. Cuối cùng là phần nhận xét, đánh giá kết luận của bài viết.

## 2. Nghiên cứu liên quan

### 2.1. Ảnh nhiễu

Do các hạn chế vật lý vốn có, cách cài đặt thông số của các thiết bị ghi khác nhau, các yếu tố khách quan như môi trường, ánh sáng, hình ảnh có xu hướng bị nhiễu ngẫu nhiên trong quá trình thu nhận hình ảnh. Nhiễu có thể hiểu là hiện tượng méo tín hiệu cơ bản gây cản trở quá trình quan sát hình ảnh, trích xuất thông tin ảnh và là một trong những nguyên nhân ảnh hưởng đến độ chính xác của máy học. Với sự gia tăng mạnh mẽ công nghệ trong việc tạo ra hình ảnh kỹ thuật số thường được chụp trong điều kiện không khí/ánh sáng kém, các phương pháp khôi phục hình ảnh đã trở thành công cụ không thể thiếu trong kỷ nguyên phân tích có sự hỗ trợ của máy tính. Trong thực tiễn, nhiếp ảnh gia thường sẽ điều chỉnh các thông số tạo nên sự cân bằng sáng của máy ảnh là khẩu độ (Aperture), tốc độ màn trập (Shutter) và ISO (International Organisation for Standardisation) như Hình 1.

**Hình 1.** Sự ảnh hưởng của các thông số đối với ảnh thu được



ISO là một trong 3 thành tố quan trọng nhất trong nhiếp ảnh, và nó cũng có những ảnh hưởng đáng kể tới bức ảnh thu được, đơn giản ISO là một cài đặt thông số của máy ảnh nó có tác dụng lớn nhất là điều chỉnh độ nhạy sáng của cảm biến máy ảnh với ánh sáng. Mặc dù ISO ban đầu chỉ được đưa ra để đo độ nhạy của ảnh film trên các máy ảnh phim, nhưng sau này các nhà sản xuất máy ảnh kỹ thuật số đã quyết định dùng luôn từ này cho mục đích đo độ nhạy sáng trên cảm biến của máy ảnh kỹ thuật số như trên máy ảnh film. Khi tăng ISO thì cảm biến tăng thu nhận ánh sáng của môi trường vào cảm biến máy ảnh làm cho bức ảnh sáng hơn, khi giảm ISO thì bức ảnh tối hơn. Vì lý do này ISO sẽ giúp ảnh cho khi chụp trong điều kiện ánh sáng yếu, hoặc trong một số trường hợp đặc biệt mà không muốn ảnh hưởng đến tốc độ và khẩu độ máy ảnh. Tuy nhiên, mặt trái của việc làm dụng ISO để tăng sáng cho bức ảnh sẽ dẫn đến những hiệu ứng không hay cho bức ảnh. Một bức ảnh được chụp với mức ISO quá cao sẽ bị hạt hay còn gọi là bị nhiễu, trong một số trường hợp bức ảnh sẽ không dùng được do quá nhiễu ảnh hưởng đến thẩm mỹ như Hình 2.

**Hình 2.** Một bức ảnh được chụp với mức ISO quá cao sẽ bị nhiễu



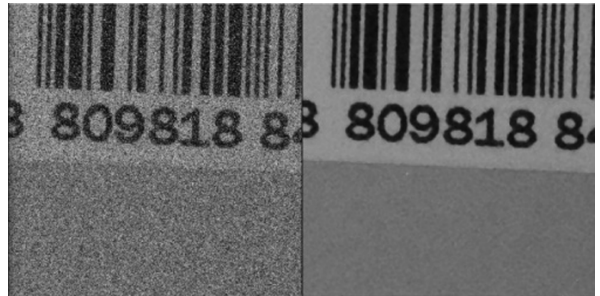
Có nhiều nguyên nhân gây ra các loại nhiễu trong thế giới thực, bao gồm nhiễu hạt photon (Photon Shot Noise), nhiễu có cấu trúc (El Gamal A.A. và Kim Y.-H, 2011), nhiễu dòng điện tối (Dark Current Noise) (Hui R, 2020, trang 125 – 154), nhiễu đọc (Readout Noise), nhiễu lượng tử (Quantization Noise),... Mỗi loại nhiễu có các thuộc tính riêng và có thể khử nhiễu bằng những phương pháp khác nhau. Tuy nhiên, trong bài viết này, chúng tôi chỉ nghiên cứu về nhiễu thực trong máy ảnh vì chúng thường gặp trong ảnh kỹ thuật số, đó là

vấn đề về độ nhạy của cảm biến máy ảnh hay còn gọi là ISO.

## 2.2. Khử nhiễu ảnh

Khử nhiễu ảnh (Image denoising) là một chủ đề được sự quan tâm lớn của nhà nghiên cứu do có rất nhiều ứng dụng, từ phân loại và phân vùng hình ảnh, đến khôi phục hình ảnh cho các phương tiện cũ và cải thiện hình ảnh.

**Hình 3.** Ảnh nhiễu (bên trái) và ảnh khử nhiễu (bên phải) bằng bộ lọc trung vị



Do mức độ chú ý mà nó nhận được, các mô hình và phương pháp đã và đang được phát triển để loại bỏ nhiễu trong hình ảnh. Dưới đây là một số phương pháp được sử dụng để khử nhiễu:

**Lọc không gian (Spatial Filtering):** Lọc không gian xoay quanh trực tiếp tương tác với chính các điểm ảnh của hình ảnh, bằng cách sử dụng mặt nạ, tương tự để lọc trong CNN, để áp dụng các mẫu cụ thể trên mỗi điểm ảnh trong hình ảnh. Ví dụ về lọc không gian là: bộ lọc tuyến tính như dưới dạng bộ lọc trung bình - sử dụng bộ lọc để tính trung bình giá trị của điểm ảnh với mỗi điểm ảnh liền kề (Jappreet Kaur và cộng sự, 2012) và bộ lọc phi tuyến tính như trung vị bộ lọc - tìm giá trị trung bình trong miền giá trị của bộ lọc và thay thế mỗi điểm ảnh trong miền giá trị với trung vị đó .

• **Lọc miền chuyển đổi (Transform Domain Filtering):** Lọc miền chuyển đổi hoạt động bằng cách biến đổi hình ảnh sang các miền khác (chẳng hạn như miền tần số) và lọc bỏ nhiễu trong miền đó . Có hai kỹ thuật đáng chú ý trong danh mục này: Lọc tần số không gian, lọc ra các tần số không gian nhất định bằng cách sử dụng Fast Fourier Transform và Wavelet Domain Filtering, chuyển đổi sang

miền wavelet bằng cách sử dụng Discrete Wavelet Transform.

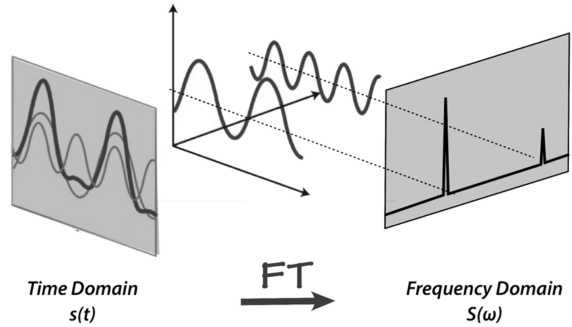
Gần đây, phương pháp học sâu đã được áp dụng để khử nhiễu ảnh bằng nhiều cách khác nhau: từ học quá trình lọc và thu nhỏ, đến sử dụng các mô hình CNN với bộ mã hóa tự động để chuyển đổi hình ảnh nhiễu thành hình ảnh bình thường.

### 2.3. Phép biến đổi Fourier (Fourier Transformer)

Biến đổi Fourier hay chuyển hóa Fourier, được đặt tên theo nhà toán học người Pháp Joseph Fourier, là phép biến đổi một hàm số hoặc một tín hiệu theo miền thời gian sang miền tần số. Phép biến đổi Fourier được sử dụng để biến đổi một ảnh ở miền không gian sang miền tần số, hay nói cách khác phân rã ảnh thành các thành phần sang tần số và

phổ của nó (Mansurov N, 2012).

**Hình 4.** Ảnh minh họa về tín hiệu qua biến đổi Fourier



Biến đổi tín hiệu từ miền thời gian (Time Domain) liên tục sang miền tần số (Frequency Domain) của biến đổi Fourier. Trong xử lý ảnh biến đổi Fourier 2 chiều (2DFT) với ảnh có kích thước  $M \times N$ , được định nghĩa là:

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-\frac{2i\pi(ux)}{M}} e^{-\frac{2i\pi(vy)}{N}} \quad (1)$$

Trong đó:

- $M$  và  $N$  là số điểm ảnh của một chiều  $x$  và  $y$  trong miền không gian.
- $f(x, y)$  là ảnh gốc đầu vào,  $F(u, v)$  là ảnh sau khi biến đổi fourier
- $e$  là cơ số logarit tự nhiên xử lý cho mỗi chiều của ảnh.
- $-i$  là phần ảo của số khi qua biến đổi Fourier
- Sử dụng trong công thức Euler:

$$e^{-i\omega} = \cos\omega - i\sin\omega \quad (2)$$

với  $e$  là cơ số Logarit tự nhiên,  $i$  là đơn vị của số phức, và  $\cos$  và  $\sin$  lần lượt là các hàm số lượng giác Cosin và Sin trong tần số.

Khi thực hiện biến đổi Fourier trên ảnh có kích thước  $M \times N$  với  $M=4$  và  $N=4$ . Thu kết quả thu được một ma trận số phức  $F$ , với ss

$$F = U(x, u) * f(x, y) * V(y, v) \quad (3)$$

Ma trận biến đổi tính với công thức tính

$$(x, u) = e^{-\frac{2i\pi(ux)}{M}} \text{ với } x, u = 0:M-1$$

$$f(x, y) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}; U(x, u) = \begin{bmatrix} e^{-\frac{2i\pi \times 0 \times 0}{4}} & e^{-\frac{2i\pi \times 0 \times 1}{M}} & e^{-\frac{2i\pi \times 0 \times 2}{4}} & e^{-\frac{2i\pi \times 1 \times 3}{4}} \\ e^{-\frac{2i\pi \times 1 \times 0}{4}} & e^{-\frac{2i\pi \times 1 \times 1}{4}} & e^{-\frac{2i\pi \times 1 \times 2}{4}} & e^{-\frac{2i\pi \times 1 \times 3}{4}} \\ e^{-\frac{2i\pi \times 2 \times 0}{4}} & e^{-\frac{2i\pi \times 2 \times 1}{4}} & e^{-\frac{2i\pi \times 2 \times 2}{4}} & e^{-\frac{2i\pi \times 2 \times 3}{4}} \\ e^{-\frac{2i\pi \times 3 \times 0}{4}} & e^{-\frac{2i\pi \times 3 \times 1}{4}} & e^{-\frac{2i\pi \times 3 \times 2}{4}} & e^{-\frac{2i\pi \times 3 \times 3}{4}} \end{bmatrix}$$

$$\rightarrow U(x, u) = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{bmatrix}$$

Tương tự

$$V(y, v) = e^{-\frac{2i\pi(yv)}{N}} \text{ với } y, v = 0: N-1$$

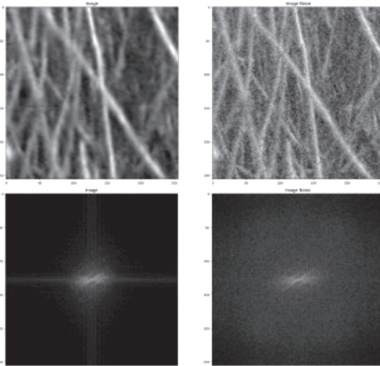
$$V(y, v) = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{bmatrix}$$

Thu được ma trận số phức của ảnh qua biến đổi với  $F = U(x, u) * f(x, y) * V(y, v)$ .

$$F = \begin{bmatrix} 4 & -2 - i2 & 0 & -2 + i2 \\ -2 - i2 & i2 & 0 & i \\ 1 & -1 & 0 & -1 \\ -2 + i2 & 2 & 0 & -i2 \end{bmatrix}$$

Biến đổi Fourier được sử dụng cho bài toán khử nhiễu:

**Hình 5.** Ảnh nhiễu (bên phải) và ảnh sạch (bên trái) sau khi áp dụng biến đổi Fourier



Các vùng biểu thị biên độ của tần số được biểu diễn phong phú cho phép phân tích nhiều hơn các thành phần tần số khác nhau của ảnh thuận lợi việc trích xuất đặc trưng trong xử lý ảnh. Để biến đổi ảnh trở về miền không gian từ miền Fourier hay miền tần số, sử dụng biến đổi Fourier ngược (IFT- Inverse Fourier Transform) của ảnh (u, v) bởi công thức:

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{i\frac{2\pi(ux)}{M}} e^{i\frac{2\pi(vy)}{N}} \quad (4)$$

Khi ảnh đầu vào lớn, tính toán biến đổi Fourier nhanh (FFT - Fast Fourier Transform) được thay thế sử dụng cho biến đổi Fourier. FFT chia nhỏ 2 khối 4 điểm thay vì tính toán cả khối 8 điểm như FT.

So sánh chi phí tính toán:

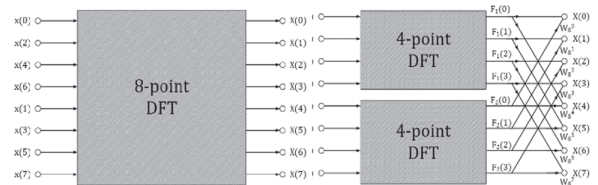
- 2DFT với  $N$  điểm:  $N^2$  phép nhân phức.

- FFT: 2 DFT với  $\frac{N}{2}$  điểm:  $\frac{N^2}{2} + \frac{N}{2}$  phép nhân phức.

Khi  $N$  càng lớn: độ lợi tính toán:

$\lim_{N \rightarrow \infty} \frac{\frac{N^2 + N}{2}}{N^2} = \frac{1}{2}$ . Vậy khi chia nhỏ khối DFT  $N$  điểm thành 2 khối DFT  $\frac{N}{2}$ , giảm được  $\frac{1}{2}$  chi phí tính toán. Càng chia nhỏ càng tiết kiệm được chi phí tính toán.

**Hình 6.** FT (bên trái) và FFT (bên phải) trong xử lý 8 điểm



Biến đổi Fourier là một công cụ quan trọng trong xử lý hình ảnh và nó có thể giúp chúng ta hiểu rõ hơn về cấu trúc và tính chất của ảnh khi được phổ lên miền tần số, cũng như thực hiện các thao tác xử lý khác nhau.

### 2.4. Phương pháp học sâu (Deep learning)

Học sâu (Deep Learning) là một lĩnh vực của trí tuệ nhân tạo và máy học. Phương pháp học sâu sử dụng mạng nơ-ron nhân tạo (Artificial Neural Networks) để giải quyết các bài toán mà tính toán thông thường chưa thể xử lý được như dữ liệu trừu tượng, dữ liệu mờ, dữ liệu không rõ ràng,... Với tính chất như vậy, học sâu thường được ứng dụng trong nhận diện hình ảnh, nhận diện giọng nói, xử lý ngôn ngữ tự nhiên hoặc dự đoán, dự báo, giảm nhiễu dữ liệu (Yu S., Ma J., và Wang W., 2018).

Đối với bài toán khử nhiễu ảnh, học sâu đã tiếp cận vấn đề này theo nhiều cách khác nhau: từ học quá trình lọc và thu nhỏ, đến sử dụng các mô hình CNN với bộ mã hóa tự động để chuyển đổi hình ảnh nhiễu thành hình ảnh

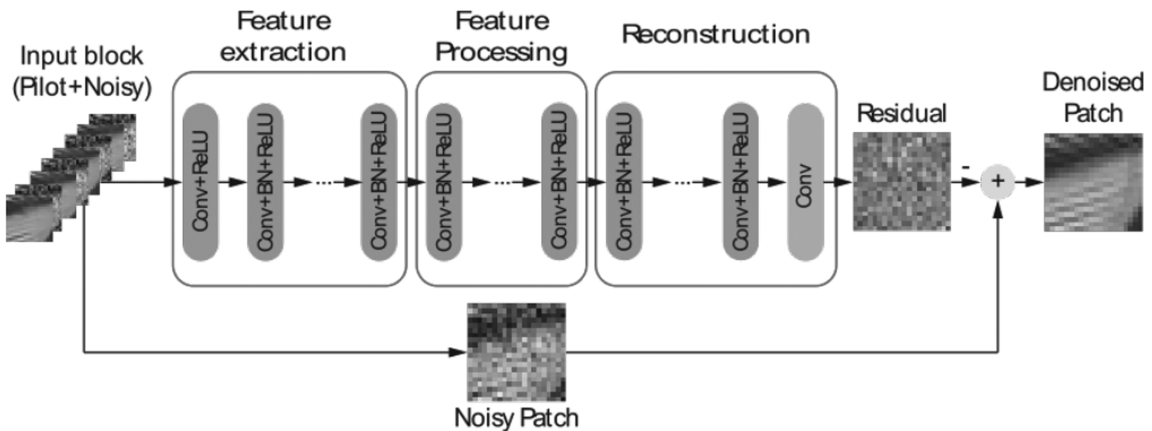
bình thường. Các kỹ thuật dựa trên phương pháp mạng tích chập học sâu có thể được phân thành nhiều loại nhiều: Mạng tích chập sâu cho ảnh nhiễu thêm vào (nhiều Gauss, Salt,...), mạng tích chập sâu cho ảnh nhiễu thực từ máy ảnh (nhiều ISO,...), và mạng tích chập sâu cho nhiễu khó thấy được nội dung ảnh (Blind Noise Image) hay nhiễu hỗn hợp (Hybrid Noisy Image) (Tian C. và cộng sự, 2020).

▪ *Mạng nơ-ron tích chập*

Nhiều mô hình học sâu hay học nhiều tầng là một tập các thuật toán học máy với ý tưởng xây dựng mô hình dữ liệu mức độ trừu tượng cao dựa trên các dữ liệu có mức độ trừu tượng hóa thấp hơn, bằng cách phân lớp dữ liệu và các biến đổi phi tuyến (Olivas E.S. và cộng sự, 2010). Mạng nơ-ron tích chập (CNN) (Ahn B. và Cho N.I., 2017). Là một mô hình học sâu có thể xây dựng được các hệ thống khử nhiễu với độ chính xác cao. Ý tưởng của CNNs được lấy cảm hứng từ khả

năng nhận biết thị giác của bộ não con người. Để có thể nhận biết được các hình ảnh trong vỏ não người có hai loại tế bào đơn giản và tế bào phức tạp. Các tế bào đơn giản phản ứng với các hình dạng cơ bản các vùng kích thích thị giác và các tế bào phức tạp tổng hợp thông tin từ các tế bào đơn giản để xác định loại bỏ nhiễu khỏi các mẫu hình dạng cơ bản ở các vùng kích thích thị giác và các tế bào phức tạp hơn. Mạng nơ-ron được sử dụng rộng rãi trong xử lý hình ảnh và khử nhiễu bằng việc sử dụng các tích chập (Convolution) và pooling để tiến hành giảm chiều dữ liệu và tìm các đặc trưng của hình ảnh. Trong việc khử nhiễu hình ảnh, mạng CNN được sử dụng để loại bỏ nhiễu từ hình ảnh đầu vào bằng cách học các đặc trưng của hình ảnh chất lượng cao và sử dụng chúng để khắc phục nhiễu trong hình ảnh đầu vào. Một số kỹ thuật khử nhiễu CNN được sử dụng như Autoencoder, U-Net và Deep Denoising CNN,...vv.

**Hình 7.** Cấu trúc cơ bản của mạng nơ-ron tích chập trong mô hình khử nhiễu [11]



Cấu trúc cơ bản của mạng nơ-ron tích chập gồm các lớp tích chập (Convolution Layers), lớp phi tuyến (Nonlinear Layer) và chuẩn hóa dữ liệu (Batch Normalization) như hình trên, cấu trúc mạng nơ-ron tích chập trong mô hình khử nhiễu gồm:

- ✓ Khối đầu vào (Input Block).
- ✓ Trích xuất đặc trưng (Feature Extraction).
- ✓ Xử lý thông tin đặc trưng (Feature

Processing).

- ✓ Tái thiết lập cấu trúc (Reconstruction).
- ✓ Bộ lọc nhiễu đã được xây dựng từ các giai đoạn ở trên kết hợp với các bản vá nhiễu (Noisy Patch) đầu vào để khử nhiễu cho ảnh.

Các nghiên cứu về khử nhiễu đã và đang ngày càng phát triển cho ra đời một số kiến trúc mạng là nền tảng được sử dụng trong nhiều bài báo nghiên cứu và các dự án. Mỗi

kiến trúc mạng này mang lại kết quả tốt ở các bài toán khử nhiễu như Autoencoder (Mao X.-J. và cộng sự, 2016), Ridnet (Anwar S. và Barnes N., 2020), DnCNN (Zhang K. và cộng sự, 2017, trang 3142-3155) là các mô hình được dùng để so sánh trong nhiều bài báo trong những năm gần đây. Và cũng là những mô hình dùng để so sánh với mô hình MIMO-Unet, DeepRFT trong nghiên cứu này.

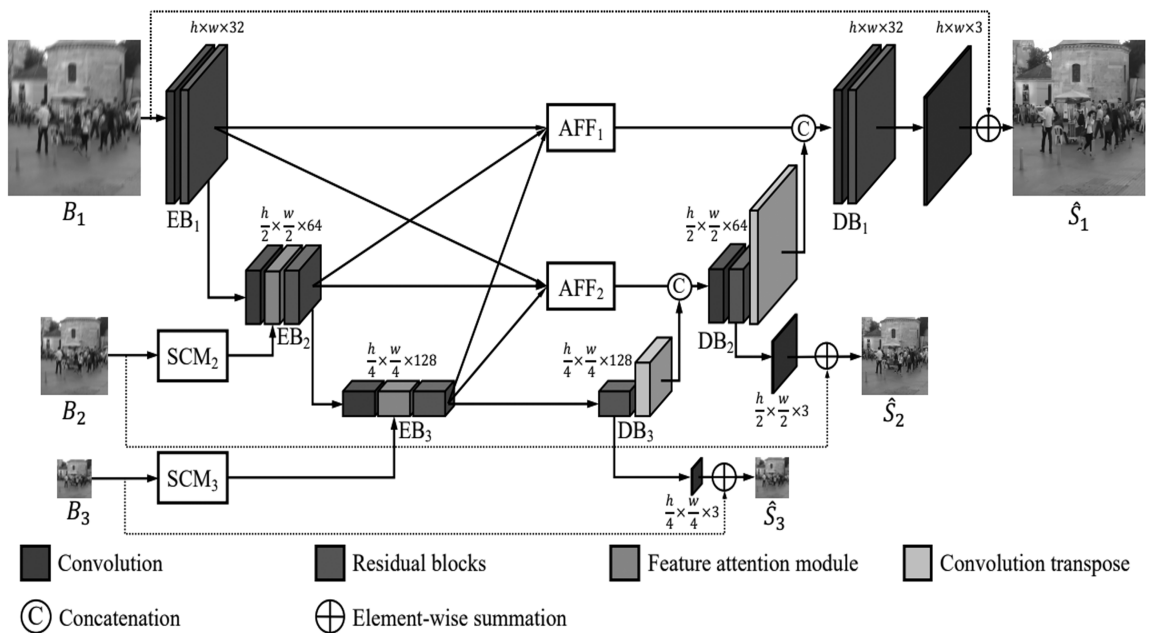
▪ *Mạng MIMO-UNet*

Trong mô hình khử nhiễu cũng như các mô hình xử lý hình ảnh để tăng độ hiệu quả của mô hình, các phương pháp truyền thống thường áp dụng mạng với hình ảnh đầu vào đa kích thước và nhiều độ nhiễu khác nhau trong cùng một mạng cho việc khử nhiễu nhưng mặt trái của việc đó dẫn đến lượng dữ liệu lớn, thời gian tính toán cao hơn rất nhiều (Anwar S. và Barnes N., 2020). Đã có một số bài báo nghiên cứu phương pháp sử dụng nhiều đầu vào và đầu ra cho mạng U-Net (Ronneberger O và cộng sự, 2015) ở các bài toán đo độ sâu siêu phân giải, nhận diện đối tượng, tái tạo ảnh, phân loại ảnh, giảm mờ của ảnh (Cho S.-J và cộng sự, 2021; Ahmed Raza S.E. và cộng sự, 2017; Fu Y. và cộng sự 2020; Pappan V.

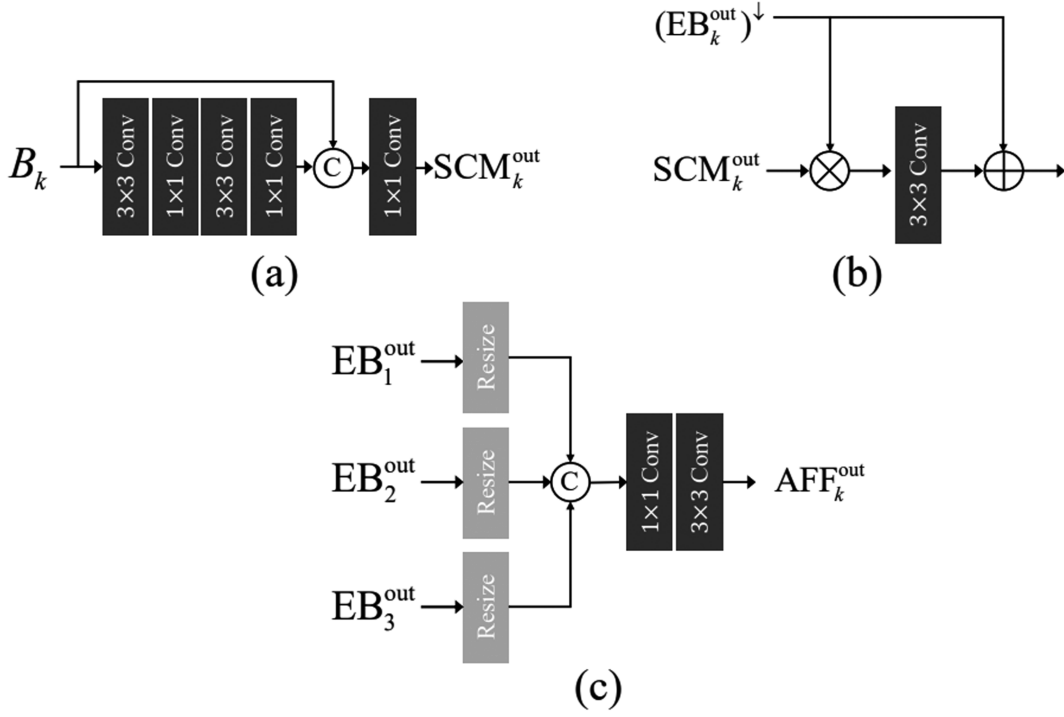
và Elad M., 2016; Xian C. và cộng sự 2020; Zamir S.W. và cộng sự, 2021) Mạng Unet được sử dụng trong khử nhiễu ảnh bằng cách huấn luyện mô hình trên các cặp ảnh nhiễu và ảnh gốc. Quá trình này giúp mô hình học cách loại bỏ các nhiễu trong ảnh và tái tạo lại ảnh gốc. Phần encoder của mạng Unet giúp rút trích các đặc trưng quan trọng từ ảnh, trong khi phần decoder giúp tạo ra ảnh tái tạo đã được khử nhiễu. Tuy nhiên, mạng Unet cũng có thể gặp một số vấn đề trong việc khử nhiễu ảnh, đặc biệt là khi nhiễu quá nhiều hoặc khi có nhiều loại nhiễu phức tạp trong ảnh.

MIMO-UNet (Multiple-Input Multiple-Output U-Net) là một kiến trúc mạng neural được phát triển bởi nhóm nghiên cứu ở đại học Hàn Quốc vào năm 2021 cho bài toán khử mờ ảnh hay tăng độ nét ảnh với đầu vào đa tỷ lệ. MIMO-UNet là một phiên bản nâng cao của mạng UNet bằng sự kết hợp giữa nhiều mạng UNet và kiến trúc MIMO (Multiple Input Multiple Output). MIMO-UNet được sử dụng để giải quyết các vấn đề liên quan đến việc phân tích và xử lý hình ảnh, ví dụ như việc phân tích từ nhiều nguồn đầu vào và tạo ra nhiều kết quả đầu ra.

Hình 8. Kiến trúc mạng MiMoUnet ở mô hình khử mờ [22]



Hình 9. Các khối trong MIMO-Unet (a) SCM (b) FAM (c)AFF



• **Khối mã hóa đơn đa đầu vào (Multi-input Single Encoder)**

Trong MIMO-UNet, các khối mã hóa (Encoder Block - EB) lấy ảnh đầu vào với kích thước khác nhau. Khi đó, ngoài việc trích xuất đặc trưng theo kích thước từ khối mã hóa phía trên truyền xuống, cũng trích xuất thông tin đặc trưng bậc thấp từ ảnh giảm kích thước từ khối SCM và sau đó kết hợp cả hai đặc trưng. Bằng cách tận dụng việc bổ sung thông tin từ đặc trưng giảm kích thước và đặc trưng có thể thu được từ hình ảnh giảm độ phân giải, khối mã hóa sẽ xử lý hình ảnh đa dạng hiệu quả hơn.

• **Khối giải mã đơn đa đầu ra (Multi-output single decoder)**

Các khối giải mã (Decoder Block-DB) khác nhau có bộ đặc trưng với các kích thước khác nhau, các bộ đặc trưng với kích thước khác nhau có thể sử dụng bất chước được các mạng con đa tầng (Multi-Stacked Sub-Networks). Không giống mạng giám sát (Intermediate Supervision), ở mạng MIMO-UNet giám sát trung gian được áp dụng cho từng khối giải mã. Việc tái tạo hình ảnh ở mỗi

cấp độ sử dụng công thức sau:

$$\hat{s} = \begin{cases} o(DB_n(AFF_n^{out}; DB_{n+1}^{out})) + B_n, & n = 1, 2, \\ o(DB_n(EB_n^{out})) + B_n, & n = 3 \end{cases} \quad (5)$$

Với  $AFF_n^{out}$ ,  $EB_n^{out}$  và  $DB_n^{out}$  là đầu ra của theo từng cấp độ của bộ mô đun trộn đặc trưng bất đối xứng (Asymmetric Feature Fusion -AFF), khối mã hóa EB, khối giải mã DB. Vì đầu ra của khối giải mã là bộ đặc trưng không phải là hình ảnh, nên cần một hàm ánh xạ  $o$  để tạo ra một đầu ra trung gian bằng sử dụng một lớp tích chập.

Mạng MIMO-UNet gồm 4 thành phần chính:

- SCM (Shallow Convolutional Module)
- Khối mã hóa EB (Encoder Block)
- Mô đun kết hợp đặc trưng bất đối xứng AFF (Asymmetric Feature Fusion)
- Khối giải mã DB (Decoder Block)
- **Shallow Convolutional Module**

Với các mô hình phổ biến hiện nay tiêu biểu là CNN cho bài toán khử nhiễu ảnh, ảnh đầu vào là toàn bộ ảnh với kích thước cố định.

Tuy nhiên MIMO-Unet xử lý ảnh bằng cách sử dụng hàm nội suy (interpolate) trong Pytorch để giảm kích thước ảnh đầu vào và đưa vào khối SCM (Shallow Convolutional Module) tạo ra các bộ đặc trưng với kích thước tương ứng.

Hàm nội suy được sử dụng để thực hiện việc thay đổi kích thước ảnh đầu vào để đảm bảo rằng các ảnh đầu vào có kích thước khác nhau vẫn có thể được xử lý bởi mô hình MIMO-Unet. Cụ thể, khi đưa vào một ảnh có kích thước khác với kích thước ảnh huấn luyện ban đầu, hàm interpolate sẽ thực hiện việc nội suy giá trị của các pixel trên ảnh đó để tạo ra một ảnh mới có kích thước mong muốn. Phương pháp nội suy có thể là 'Nearest', 'Linear', 'Bilinear', 'Bicubic', 'Trilinear', 'Area'.

Việc sử dụng hàm nội suy cho phép mạng xử lý được các ảnh có kích thước khác nhau mà không cần phải thực hiện một bước tiền xử lý để đưa tất cả các ảnh về cùng một kích thước, giúp tiết kiệm thời gian và tăng tính linh hoạt của mô hình.

Khối SCM gồm lớp tích chập  $3 \times 3$  và  $1 \times 1$  chồng lên nhau, ở lớp tích chập  $1 \times 1$  cuối kết hợp đặc trưng với bản vá đầu vào  $B_k$  bằng hàm concat, ở khối này trích xuất các đặc trưng bậc thấp như cạnh, đường viền, màu sắc. Đầu ra của khối SCM được đưa vào các khối mã hóa để tiếp tục bổ sung cho việc trích xuất đặc trưng bậc cao.

#### • Khối mã hóa

Khối mã hóa (Encoder Block-EB) gồm các lớp tích chập, khối dư (Residual Block) và mô đun sự chú ý đặc trưng (Feature Attention Module - FAM). Ảnh đầu vào sẽ được dùng hàm nội suy giảm kích thước ảnh xuống còn 1/2 và 1/4 để tạo ra 2 ảnh giảm kích thước mới, khối SCM sẽ trích các đặc trưng từ 2 ảnh giảm kích thước và tạo ra các bộ đặc trưng mới. Khối mã hóa trích xuất đặc trưng ảnh đầu vào và khối FAM kết hợp các bộ đặc trưng được giảm kích thước bởi hàm nội suy từ khối mã hóa EB và bộ đặc trưng giảm kích thước được tạo từ khối SCM để tập trung vào việc

tăng cường mối liên hệ các đặc trưng và điều chỉnh bằng cách giảm thiểu độ quan trọng của các đặc trưng không quan trọng và tăng cường độ quan trọng của các đặc trưng quan trọng hơn. Sau đó đưa vào mô đun trộn đặc trưng bất đối xứng.

#### • Mô-đun trộn đặc trưng bất đối xứng

Trong một số mạng khử nhiễu truyền thống, quy trình bắt đầu từ nền tới chi tiết để xử lý hình ảnh, các đặc trưng từ mạng con quy mô nhỏ hơn được sử dụng cho các mạng con quy mô lớn hơn, làm cho luồng thông tin không linh hoạt. Một phương pháp đặc biệt là xếp toàn bộ mạng theo hướng ngang hoặc dọc, cho phép thông tin lưu thông từ trên xuống dưới và từ dưới lên trên.

Mô-đun kết hợp đặc trưng bất đối xứng (AFF- Asymmetric Feature Fusion Module) được thiết kế để kết hợp các đặc trưng từ các nhánh với kích thước khác nhau trong mạng. Mỗi khối AFF lấy thông tin từ đầu ra của tất cả khối mã hóa và kết hợp các bộ đặc trưng đó tạo ra các bộ đặc trưng phức tạp hơn. Đầu ra của khối AFF được giao cho khối giải mã tương ứng.

$$AFF_1^{out} = AFF_1(EB_1^{out}, (EB_2^{out})^\uparrow, (EB_3^{out})^\uparrow)$$

$$AFF_2^{out} = AFF_2((EB_1^{out})^\downarrow, EB_2^{out}, (EB_3^{out})^\uparrow)$$

Trong đó,  $AFF_n^{out}$  là đại diện cho đầu ra của khối mô đun trộn đặc trưng bất đối xứng với  $n$  là thứ tự khối AFF. Với ( $\uparrow$ ) tăng kích thước (Upsample) và ( $\downarrow$ ) (Downsample) giảm kích thước được sử dụng cho hàm concat khi kết hợp các đặc trưng từ các kích cỡ khác nhau.

#### • Khối giải mã

Các bộ đặc trưng giảm kích thước 1/4 qua khối dư ở khối giải mã để học các đặc trưng quan trọng và tránh hiện tượng biến mất đạo hàm (Vanishing Gradient), các bộ đặc trưng mới tạo khi qua khối dư sẽ vừa thực hiện:

- Qua lớp tích chập cuối cùng tạo các bộ

đặc trưng đầu ra cuối cùng ở kích thước 1/4 so với ảnh đầu vào.

- Qua lớp tích chập chuyển vị (convolution Transpose) tăng kích thước sử dụng cho nối các bộ đặc trưng kích thước 1/2 tạo bởi khối  $AFF_2$ , sau đó đi qua lớp tích chập và khối dư để trích xuất thêm đặc trưng mới và thực hiện lại các bước trên.

### 3. Mô hình đề xuất

Tận dụng các ưu điểm kiến trúc mạng MIMO-UNet và phép biến đổi Fourier, chúng tôi xây dựng mạng DeepRFT cho việc xử lý nhiễu ảnh ở tập ảnh được lấy từ máy ảnh Fujifilm. DeepRFT được xây dựng dựa trên kiến trúc MIMO-UNet bằng cách sử dụng

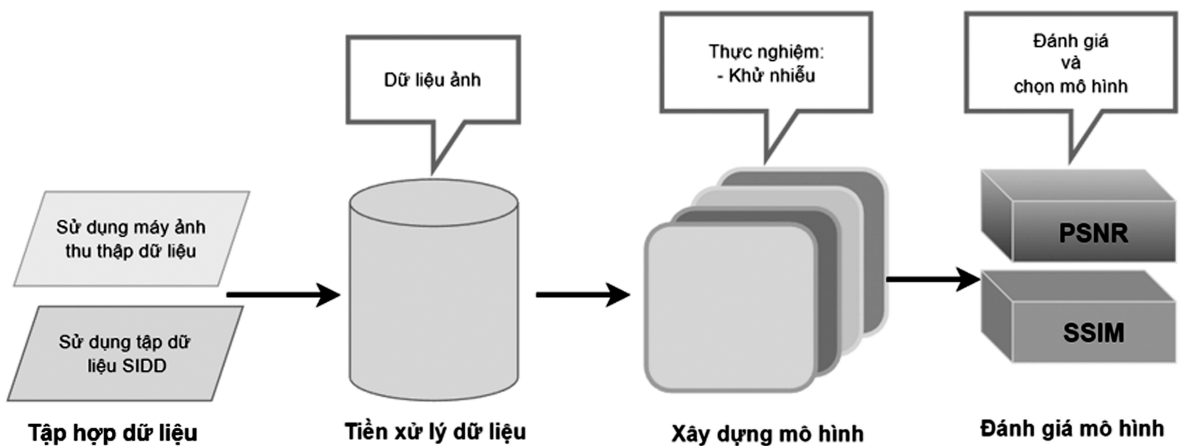
phép biến đổi Fourier tạo thành khối ResFFT-Conv Block thay vì Residual Block cải thiện hiệu suất và tốc độ tính toán trong quá trình rút trích đặc trưng ảnh nhiễu.

Phép biến đổi Fourier giúp tách thông tin tần số cao và thấp trong ảnh. Trong trường hợp ảnh bị nhiễu, thông tin tần số cao thường tương ứng với nhiễu và thông tin tần số thấp tương ứng với nội dung chính của ảnh. Bằng cách áp dụng biến đổi Fourier, chúng ta có thể lọc bỏ hoặc giảm thiểu nhiễu bằng cách loại bỏ thông tin tần số cao và chỉ giữ lại thông tin tần số thấp.

#### 3.1. Mô hình tổng quan của hệ thống

Mô hình tổng quan của hệ thống gồm có giai đoạn chính như hình sau:

Hình 10. Mô hình tổng quan của hệ thống



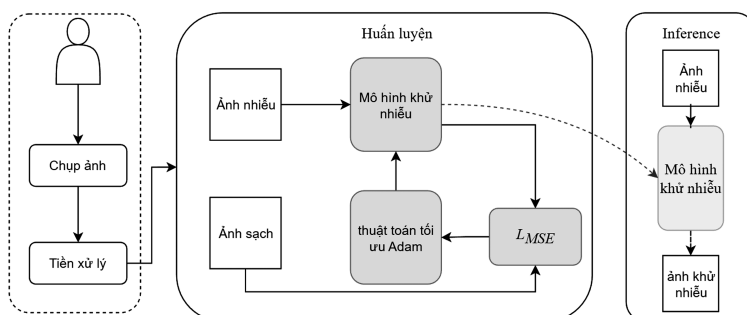
- Tập hợp dữ liệu: Dữ liệu ảnh tự thu thập và tập dữ liệu ảnh chuẩn SIDD có sẵn.
- Tiền xử lý dữ liệu để có tập dữ liệu ảnh chuẩn phục vụ cho mô hình khử nhiễu.
- Xây dựng mô hình khử nhiễu để khử nhiễu ảnh.
- Đánh giá mô hình khử nhiễu bằng các

độ đo PSNR và SSIM.

#### 3.2. Mô hình khử nhiễu

Trong thực nghiệm này, chúng tôi sử dụng các mô hình khử nhiễu có sẵn: Autoencoder, DnCNN, RidNet, MIMO-UNet để so sánh với mô hình đề xuất của chúng tôi DeepRFT trên cùng tập dữ liệu.

Hình 11. Mô hình khử nhiễu



Quy trình huấn luyện một mô hình khử nhiễu có thể được thực hiện theo các bước sau:

- Chuẩn bị dữ liệu: Thu thập và chuẩn bị các tập dữ liệu huấn luyện, đánh giá và kiểm tra.
- Thiết lập mô hình: thiết lập các thông số của mô hình Autoencoder, DnCNN, RidNet, MIMO-UNet, DeepRFT như tỉ lệ học, hàm kích hoạt, v.v.
- Huấn luyện mô hình: Sử dụng tập dữ liệu huấn luyện để huấn luyện mô hình. Trong quá trình huấn luyện, mô hình trích xuất đặc trưng từ ảnh nhiễu học cách khử nhiễu cho các hình ảnh và được cập nhật thông số mô hình để cải thiện hiệu suất của nó. Để tránh hiện tượng b overfitting, tôi sử dụng các kỹ thuật như early stopping, batch norm.
- Đánh giá mô hình: Sử dụng tập dữ liệu đánh giá để đánh giá hiệu suất của mô hình. Các độ đo thông thường được sử dụng để đánh giá mô hình khử nhiễu là chỉ số đánh giá PSNR và SSIM.
- Kiểm tra mô hình: Sử dụng tập dữ liệu kiểm tra để đánh giá khả năng tổng quát hóa của mô hình, tức là khả năng áp dụng mô hình cho các hình ảnh mới, chưa được sử dụng trong quá trình huấn luyện.
- Tinh chỉnh mô hình: Dựa trên kết quả đánh giá trên tập đánh giá và tập kiểm tra, ta có thể điều chỉnh các thông số của mô hình

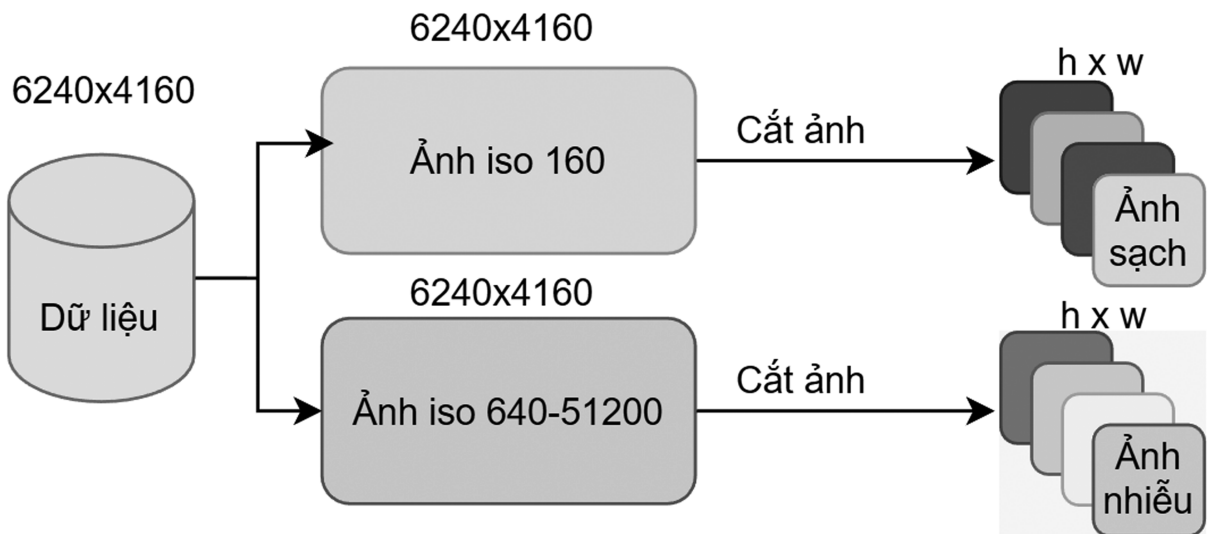
hoặc thiết lập lại các tham số huấn luyện để cải thiện hiệu suất của mô hình.

- Sử dụng mô hình (Inference): Sau khi huấn luyện và đánh giá, mô hình đã được tối ưu để khử nhiễu cho các hình ảnh mới. Ta có thể sử dụng mô hình này để khử nhiễu cho các hình ảnh trong thực tế.

### 3.3. Dữ liệu thực nghiệm

Trong thực nghiệm này, chúng tôi sử dụng hai tập dữ liệu: Tập dữ liệu FND (FujiNoiseDataset): Tất cả các ảnh nhiễu ở tập dữ liệu được thực hiện bởi nhóm thông qua thu tự thập bằng máy ảnh Fujifilm XT4, cảm biến ASP-C, Độ phân giải: 26.1MP, Dải ISO: 80 – 51200 với ống kính XF60mm Macro, các vật được chụp là vật tĩnh với nhiều thông số ISO khác nhau, do thay đổi ISO nên ảnh sẽ được máy đo độ sáng EV và tự động thay đổi tốc độ chụp bù sáng cho ảnh, khẩu ống kính máy ảnh được giữ nguyên, máy ảnh được gắn cố định trên tripod và chụp từ xa bằng ứng dụng Fujifilm Control (giảm thiểu được việc tiếp xúc với máy khi tăng dải ISO bằng ứng dụng điện thoại); Tập dữ liệu gồm 1100 ảnh của 100 cảnh và mỗi cảnh có 11 ảnh ở các ISO khác nhau với kích thước 6240x4120. Chúng tôi sử dụng dải ISO từ 160 đến 51200 để xem từng mức độ nhiễu của cảm biến máy ảnh, các ảnh được đặt tên dưới định dạng [cảnh]\_[ISO]. Ảnh thu thập được tiền xử lý như hình sau:

Hình 12. Quy trình tiền xử lý tập ảnh FND



Tập dữ liệu thứ 2 sử dụng cho mô hình khử nhiễu là tập dữ liệu SIDD (Smartphone Image Denoising Dataset). Tập dữ liệu gồm 30,000 hình ảnh, từ 10 cảnh trong các điều kiện ánh sáng khác nhau được thu thập bằng camera của 5 điện thoại thông minh (Google Pixel, iPhone 7, Samsung Galaxy S6 Edge, Motorola Nexus 6, LG G4), tập dữ liệu có 3 phiên bản bao gồm phiên bản nhỏ (small), trung bình (medium) và bản đầy đủ (full). Ảnh sẽ có 2 định dạng là MAT dành cho ảnh raw, PNG cho ảnh sRGB có độ phân giải khác nhau “2988x5312, 3000x5328, 3120x4208,..” được tổ chức theo cách nơi chúng được sắp xếp trong các thư mục riêng biệt có tên phù hợp với nội dung. Chúng tôi sử dụng phiên bản nhỏ của tập dữ liệu và tiền xử lý bằng cách cắt thành các bản vá nhỏ để kiểm tra độ hiệu quả mô hình khử nhiễu.

#### 4. Kết quả thực nghiệm

Trong quá trình thực nghiệm chúng tôi

đã có điều chỉnh tham số đầu vào như tốc độ học, kích thước ảnh, cửa sổ trượt của các mạng sao cho ra kết quả phù hợp nhất đối với từng cấu trúc. Các mô hình sử dụng kỹ thuật dừng sớm như: mô hình Autoencoder đạt kết quả cao nhất ở epoch 21; DnCNN ở epoch 32; RIDNet ở epoch 26; MIMO-UNet ở epoch 65 và DeepRFT ở epoch 71 ở mỗi tập dữ liệu đầu với kích thước ảnh 512x512. Tối ưu được độ chính xác khi bắt đầu đánh giá mô hình đã được huấn luyện, sau đó các trọng số của các mô hình đã huấn luyện được kiểm thử bằng tập kiểm thử của tập dữ liệu tự thu thập.

Đánh giá các mô hình khử nhiễu dựa trên các chỉ số PSNR và SSIM thì các mô hình có kết quả cạnh tranh rõ để đưa ra nhận xét về dữ liệu và mô hình. Bảng 1 và Bảng 2 cho thấy kết quả đánh giá của 5 mô hình khử nhiễu đã huấn luyện trên tập dữ liệu tự thu thập FND.

**Bảng 1.** Kết quả đánh giá PSNR kiểm thử của mô hình trên tập dữ liệu thu thập

ISO	Ảnh nhiễu	Autoencoder	DnCNN	RidNet	MiMoUnet	DeepRFT
640	33.217	30.662	32.593	33.235	34.269	<b>35.224</b>
1000	32.081	29.609	30.711	30.870	32.968	<b>33.343</b>
1600	31.786	29.337	30.134	30.240	32.609	<b>32.892</b>
3200	31.463	29.467	30.12	30.246	32.683	<b>32.87</b>
6400	30.131	28.926	29.223	29.467	31.878	<b>32.083</b>
8000	29.891	28.772	28.996	29.225	31.692	<b>31.888</b>
10000	29.219	28.644	28.819	28.983	31.431	<b>31.695</b>
12800	28.426	28.521	28.215	28.405	30.974	<b>31.262</b>
25600	27.847	28.092	27.108	27.347	30.170	<b>30.502</b>
51200	26.150	26.723	25.562	25.795	28.599	<b>28.947</b>

**Bảng 2.** Kết quả đánh giá SSIM kiểm thử của mô hình khử nhiễu trên tập dữ liệu thu thập

ISO	Ảnh nhiễu	Autoencoder	DnCNN	RidNet	MiMoUnet	DeepRFT
640	0.8974	0.8964	0.9118	0.9131	0.9261	<b>0.9324</b>
1000	0.8652	0.8911	0.8739	0.8705	0.9104	<b>0.9133</b>
1600	0.8442	0.8847	0.8537	0.8490	0.8998	<b>0.9039</b>
3200	0.8377	0.8811	0.8453	0.8409	0.8951	<b>0.8973</b>
6400	0.8121	0.8739	0.8185	0.8135	0.8911	<b>0.8915</b>
8000	0.7972	0.8727	0.8054	0.7993	0.8915	<b>0.8924</b>
10000	0.7861	0.8706	0.7949	0.7882	0.8867	<b>0.8889</b>
12800	0.7677	0.8686	0.7676	0.7591	0.8853	<b>0.8871</b>
25600	0.6763	0.8568	0.6893	0.6770	0.8762	<b>0.8795</b>
51200	0.6363	0.8344	0.6495	0.6372	0.8694	<b>0.8711</b>

Bảng 1 và Bảng 2 thể hiện chỉ số đánh giá PSNR và SSIM của ảnh nhiễu ở mỗi cấp độ ISO được tính theo chất lượng ảnh ở mức độ ISO đó so với ảnh ở ISO = 160. Kết quả này cho thấy trong một số trường hợp, Autoencoder có thể khử nhiễu hiệu quả hơn DnCNN và RIDNet. Tuy nhiên, điều này không có nghĩa là Autoencoder luôn tốt hơn DnCNN và RIDNet trong việc khử nhiễu ảnh. Kết quả đánh giá các mô hình đều giảm khi mức độ ISO càng tăng, khi tập dữ liệu nhiễu xuất hiện vấn đề. Kết quả các mạng MIMO-UNet và DeepRFT đạt được kết quả cạnh tranh với các mức nhiễu ISO, ở mạng MIMO-UNet

và DeepRFT khi sử dụng khối dư và khối dư biến đổi Fourier nhanh để khai thác đặc trưng nhiễu đã đạt hoạt động khá hiệu quả với nhiễu độ nhiễu.

Để xem xét lại độ hiệu quả của các mô hình khử nhiễu cũng như vấn đề về sai sót trong tập dữ liệu tự thu thập nên chúng tôi đã thực hiện trên tập dữ liệu SIDD, là tập dữ liệu tiêu chuẩn được sử dụng phổ biến đánh giá các mô hình khử nhiễu ảnh như Autoencoder, RIDNet, DnCNN (Anwar S. và Barnes N., 2020; Jang G. và cộng sự, 2022) Bảng 3 cho thấy kết quả khử nhiễu của mô hình DeepRFT trên tập dữ liệu SIDD hiệu quả hơn các mô hình còn lại.

**Bảng 3.** Kết quả đánh giá PSNR và SSIM kiểm thử của mô hình trên tập dữ liệu SIDD

Đánh giá	Autoencoder	DnCNN	RidNet	MIMO-UNet	DeepRFT
PSNR	36.379	36.37	38.83	39.315	<b>39.611</b>
SSIM	0.9418	0.936	0.948	0.953	<b>0.959</b>

## 5. Kết luận

Trong nghiên cứu này, chúng tôi đã nghiên cứu và đề xuất phương pháp sử dụng phép biến đổi Fourier và mô hình mạng học sâu với cách tiếp cận đa kích thước đầu vào để trích xuất đặt trung tần số nhiễu cho hệ thống khử nhiễu. Chúng tôi đã xây dựng thành công mô hình thực nghiệm để so sánh các phương pháp cổ điển như phương pháp lọc trung bình, lọc trung vị và 5 mô hình khử nhiễu Autoencoder, DnCNN, RIDNet,

MIMO-UNet, DeepRFT đã được huấn luyện trên tập dữ liệu thu thập và tập dữ liệu SIDD. Kết quả thực nghiệm cho thấy phương pháp đề xuất của chúng tôi DeepRFT khử nhiễu tốt hơn trên từng mức độ nhiễu khác nhau. Trong tương lai, chúng tôi sẽ tiếp tục mở rộng nghiên cứu để cải thiện mô hình hơn nữa khi sử dụng phương pháp đầu vào ảnh đa kích thước và khối biến đổi fourier cho những bài toán khác nhau trong lĩnh vực thị giác máy tính để khẳng định tính hiệu quả của nó.

## TÀI LIỆU THAM KHẢO

- Călin-Adrian Popa and Cosmin Cernazanu (2018), *Fourier transform-based image classification using complex-valued convolutional neural networks*.
- El Gamal A.A. và Kim Y.-H. (2011), *Network information theory*, Cambridge University Press, Cambridge ; New York.
- Hanssie (2018). ISO, Aperture & Shutter Speed | A Cheat Sheet For Beginners. SLR Lounge.
- Hui R. (2020). Chapter 4 - Photodetectors. *Introduction to Fiber-Optic Communications*. Academic Press, 125–154.
- Jappreet Kaur , Manpreet Kaur , Poonamdeep Kaur , Manpreet Kaur (2012). *Comparative Analysis of Image Denoising Techniques*.

- Olivas E.S., Guerrero J.D.M., Martinez-Sober M. và cộng sự., btv. (2010), Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques, IGI Global.
- Ronneberger O, Fischer P, Brox T (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation.
- Yu S., Ma J., và Wang W. (2018), Deep learning tutorial for denoising.
- Ahmed Raza S.E., Cheung L., Epstein D. và cộng sự. (2017). MIMO-Net: A multi-input multi-output convolutional neural network for cell segmentation in fluorescence microscopy images. 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), Melbourne, VIC, IEEE, 337–340.
- Fu Y., Xu J., Xiang Y. và cộng sự. (2020). Multi-scale patches based image denoising using weighted nuclear norm minimisation. IET Image Processing, 14.
- Healey G.E. và Kondepudy R. (1994). Radiometric CCD camera calibration and noise estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **16(3)**, 267–276.
- Papayan V. và Elad M. (2016). Multi-Scale Patch-Based Image Restoration. *IEEE Trans on Image Process*, 25(1), 249–261.
- Zhang K., Zuo W., Chen Y. và cộng sự. (2017). Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *IEEE Trans on Image Process*, 26(7), 3142–3155.
- Ahn B. và Cho N.I. (2017). Block-Matching Convolutional Neural Network for Image Denoising, accessed: 25/01/2023.
- Anwar S. và Barnes N. (2020). Real Image Denoising with Feature Attention, accessed: 26/02/2023.
- Cho S.-J., Ji S.-W., Hong J.-P. và cộng sự. (2021). Rethinking Coarse-to-Fine Approach in Single Image Deblurring., accessed: 21/12/2022.
- Jang G., Lee W., Son S. và cộng sự. (2022). C2N: Practical Generative Noise Modeling for Real-World Denoising, accessed: 31/03/2023.
- Mansurov N. (2012). Why DX has no Future. *Photography Life*, <<https://photographylife.com/why-dx-has-no-future>>, accessed: 20/03/2023.
- Mao X.-J., Shen C., và Yang Y.-B. (2016). Image Restoration Using Convolutional Auto-encoders with Symmetric Skip Connections., accessed: 27/12/2022.
- Tian C., Fei L., Zheng W. và cộng sự. (2020). Deep Learning on Image Denoising: An overview, accessed: 04/09/2022.
- Xian C., Qian K., Zhang Z. và cộng sự. (2020). Multi-Scale Progressive Fusion Learning for Depth Map Super-Resolution. <<http://arxiv.org/abs/2011.11865>>, accessed: 31/12/2022.
- Zamir S.W., Arora A., Khan S. và cộng sự. (2021). Multi-Stage Progressive Image Restoration, accessed: 06/02/2023.