

# KHỬ NHIỄU POISSON TRÊN ẢNH SỐ DỰA TRÊN HỌC TƯƠNG PHẢN TĂNG CƯỜNG

## POISSON IMAGE DENOISING BASED ON AUGMENTED CONTRASTIVE LEARNING

Phạm Công Thắng\*, Phan Minh Nhật

Trường Đại học Bách khoa - Đại học Đà Nẵng, Việt Nam<sup>1</sup>

\*Tác giả liên hệ / Corresponding author: pethang@dut.udn.vn

(Nhận bài / Received: 02/01/2025; Sửa bài / Revised: 28/02/2025; Chấp nhận đăng / Accepted: 01/3/2025)

DOI: 10.31130/ud-jst.2025.003

**Tóm tắt** - Khử nhiễu Poisson trên ảnh số là một thách thức quan trọng trong xử lý ảnh, đặc biệt khi nhiễu phụ thuộc tín hiệu và thay đổi theo cường độ tín hiệu. Các phương pháp khử nhiễu hiện nay chủ yếu dựa trên giả định nhiễu Gaussian, do đó không thể áp dụng hiệu quả với nhiễu Poisson. Nghiên cứu này giới thiệu phương pháp khử nhiễu ảnh thông qua học tương phản có giám sát (PDSCL) nhằm giải quyết vấn đề này. Phương pháp sử dụng mạng nơ-ron để học các biểu diễn thưa, mô hình hóa các vùng ảnh tương tự và lặp lại, từ đó tăng khả năng khử nhiễu. Đồng thời, học tương phản tăng cường được áp dụng bằng cách sử dụng các ảnh nhiễu hoặc đã khôi phục làm ví dụ tiêu cực trong không gian đồng thuận, cải thiện hiệu quả khử nhiễu. Để khắc phục sự mơ hồ trong quá trình học, nhóm tác giả đề xuất chiến lược học sắp xếp ví dụ tiêu cực theo mức độ khó, nhằm tối ưu hóa hiệu suất của giải pháp đề xuất.

**Từ khóa** - Khử nhiễu ảnh; nhiễu Poisson; mạng nơ-ron; học tương phản tăng cường; biểu diễn thưa; không gian đồng thuận

### 1. Đặt vấn đề

Khử nhiễu Poisson trong ảnh là một bài toán nghịch đảo đầy thách thức nhằm khôi phục hình ảnh từ hình ảnh nhiễu quan sát được. Các chiến lược khử nhiễu hiện tại giả định rằng hầu hết các hình ảnh nhiễu có thể được mô hình hóa bằng nhiễu Gaussian [1, 2]. Mặc dù, giả định này cho thấy hiệu suất hợp lý cho một số ứng dụng, nhưng nó không thực tế về mặt vật lý đối với các loại ảnh nhiễu Poisson chuyên dụng vì nhiễu thay đổi tỷ lệ thuận với cường độ tín hiệu và phụ thuộc vào tín hiệu. Do quá trình thu nhận hình ảnh có tính chất rời rạc, dữ liệu thu thập từ các cảm biến hình ảnh thường bị xuống cấp và được mô hình hóa như một quá trình Poisson [3]. Do đó, các phương pháp khử nhiễu Poisson trực tiếp có tầm quan trọng hàng đầu trong bài toán này.

Các phương pháp học có giám sát đã được chứng minh là hiệu quả trong các nhiệm vụ khử nhiễu ảnh [1, 2], và phần lớn được xây dựng dựa trên các mạng nơ-ron (neural networks, NN) phức tạp. Hiệu quả của các phương pháp này phụ thuộc đáng kể vào việc sử dụng các tập dữ liệu lớn để huấn luyện. Các tập dữ liệu lớn này thường bao gồm các cặp dữ liệu gồm một hình ảnh gốc không nhiễu, và một hình ảnh nhiễu. Tuy nhiên, trong hầu hết các bối cảnh thực tế, việc thu thập các cặp dữ liệu như vậy là rất khó khăn [4]. Do đó, để giải quyết các tình huống này, các phương pháp học tự giám sát đã được giới thiệu [5], nhằm sử dụng

**Abstract** - Poisson noise removal in digital images is a significant challenge in image processing, particularly when the noise is signal-dependent and varies with signal intensity. Current denoising methods mainly rely on the assumption of Gaussian noise, making them ineffective for Poisson noise. This study introduces a method for image denoising through Supervised Contrastive Learning (PDSCL) to address this issue. The proposed method leverages neural networks to learn sparse representations, model similar and repetitive image regions, and thereby enhance noise reduction. Additionally, supervised contrastive learning is applied by using noisy or restored images as negative examples in the consensus space, improving denoising efficiency. To overcome ambiguity during the learning process, the authors propose a strategy to arrange negative examples based on difficulty levels, optimizing the proposed method's performance.

**Key words** - Image Denoising; Poisson noise; Neural network; enhanced contrastive learning; Sparse representation; Consensus space

các mạng NN khởi tạo ngẫu nhiên để nắm bắt các thống kê hình ảnh ở mức thấp. Các mạng này được sử dụng làm tiền đề để giải quyết các bài toán nghịch đảo như khử nhiễu cho hình ảnh mà không yêu cầu hình ảnh gốc của nó [5, 6]. Bên cạnh đó, học có giám sát sử dụng các bộ khử nhiễu dựa trên học sâu với các mạng NN tích chập, và dựa trên một tập dữ liệu gồm các hình ảnh gốc và hình ảnh nhiễu để học phép ánh xạ giữa chúng [1, 7]. Các nghiên cứu gần đây cũng đã chỉ ra rằng, ngay cả khi không có dữ liệu nhãn rõ ràng, các mô hình học sâu vẫn có thể đạt được hiệu suất cao trong nhiều tác vụ [8, 9].

Trong những năm gần đây, học tương phản đã được đề xuất và ứng dụng trong thị giác máy tính [10, 11]. Ý tưởng đằng sau học tương phản là kéo một điểm neo (anchor) gần lại điểm tích cực (positive) trong khi đồng thời đẩy nó ra xa điểm tiêu cực (negative) thông qua hàm mất mát tương phản. Một số ít các nghiên cứu đã áp dụng học tương phản vào các bài toán thị giác cấp thấp. Tuy nhiên, hầu hết các điểm tiêu cực này không có sự đồng thuận và do đó được biểu diễn xa so với điểm tích cực, dẫn đến một không gian nghiệm chưa bị ràng buộc rõ ràng.

Đóng góp chính của bài báo này là đề xuất một phương pháp mới, PDSCL, để khử nhiễu Poisson cho hình ảnh, tận dụng khả năng của mạng NN học sâu dựa trên ý tưởng từ [6] sử dụng cho nhiễu Gaussian. Cụ thể, mạng NN được sử dụng để học các biểu diễn thưa có khả năng khử nhiễu đầu

<sup>1</sup> The University of Danang - University of Science and Technology, Vietnam (Pham Cong Thang, Phan Minh Nhat)

vào bằng cách mô hình hóa các vùng ảnh tương tự và lặp lại. Sau đó, các biểu diễn thưa này được sử dụng trong quá trình tái tạo hình ảnh và tiếp tục loại bỏ nhiễu đầu vào. Ngoài ra, phương pháp này được kết hợp với học tương phản tăng cường, sử dụng các hình ảnh nhiễu ban đầu hoặc đã được khôi phục làm các ví dụ tiêu cực trong không gian đồng thuận cho việc khử nhiễu ảnh. Mức độ khó của các ví dụ tiêu cực trong không gian đồng thuận có thể ảnh hưởng đến hiệu suất khôi phục hình ảnh của PDSCL, vì vậy một chiến lược học được đề xuất để sắp xếp các ví dụ tiêu cực này để giảm bớt sự mơ hồ trong quá trình huấn luyện.

## 2. Phương pháp nghiên cứu

### 2.1. Mô hình khử nhiễu Poisson

Với ảnh nhiễu Poisson là  $\mathbf{X}_0 \in \mathbb{R}^{d \times d}$  và dạng vector hóa của nó là  $\mathbf{x}_0 \in \mathbb{R}^{d^2}$ . Các giá trị điểm ảnh của vector nhiễu  $\mathbf{x}_0$  là các biến ngẫu nhiên lấy mẫu từ phân phối Poisson, được tham số hóa bởi các giá trị điểm ảnh của vector nhân thực  $\mathbf{x}$  tại mỗi chỉ số  $i$  của ảnh. Điều này cho phép mô hình hóa  $\mathbf{x}_0[i]$  dưới dạng  $\mathbf{x}_0[i] \sim \varphi(\mathbf{x}[i])$  [12], trong đó  $\varphi$  là quá trình lấy mẫu từ phân phối Poisson được định nghĩa như sau:

$$\varphi_{\mathbf{x}[i]}(\mathbf{x}_0[i]) = \frac{\mathbf{x}[i]^{\mathbf{x}_0[i]} \exp(-\mathbf{x}[i])}{\mathbf{x}_0[i]!} \quad (1)$$

Hàm log-likelihood của phương trình (1) được tối ưu hóa nhằm ước lượng một vector khử nhiễu  $\mathbf{x}^*$  bằng cách tối thiểu hóa bài toán tối ưu sau [13]:

$$\mathbf{x}^* = \min_{\mathbf{x}} (\mathbf{1}^T \mathbf{x} - \mathbf{x}_0^T \log(\mathbf{x})) \text{ s.t. } \mathbf{x} > 0 \quad (2)$$

Trong đó,  $\mathbf{1} \in \mathbb{R}^{d^2}$  là một vector giá trị 1, ký hiệu  $>$  biểu thị bất đẳng thức theo từng phần tử, và hàm  $\log(\cdot)$  được áp dụng theo từng phần tử.

### 2.2. Thuật toán ISTA

Bài toán tối ưu (2) là không xác định rõ ràng [14]. Để giải quyết vấn đề này, việc ước lượng  $\mathbf{x}^*$  được tính toán thông qua một vector thưa  $\boldsymbol{\alpha} \in \mathbb{R}^{d^2 \times k}$  và một không gian thưa  $\mathbf{S} \in \mathbb{R}^{d^2 \times k}$  [15]:

$$\begin{aligned} \min_{\mathbf{S}, \boldsymbol{\alpha}} (\mathbf{1}^T (\mathbf{S}\boldsymbol{\alpha}) - \mathbf{x}_0^T \log(\mathbf{S}\boldsymbol{\alpha})) \\ \text{s.t. } \|\mathbf{x}_0 - \mathbf{S}\boldsymbol{\alpha}\| \leq \zeta, \mathbf{S}\boldsymbol{\alpha} > 0 \end{aligned} \quad (3)$$

với  $\zeta = \sigma^2 F_m^{-1}(\tau)$  là ngưỡng xác định,  $F_m^{-1}$  là hàm nghịch đảo của hàm phân phối tích lũy  $F_m$  của phân phối  $\chi_m^2$ ,  $\tau = 0,9$  [15].

Đặt  $\mathbf{S}\boldsymbol{\alpha} = \exp(\mathbf{S}\boldsymbol{\alpha})$  để xử lý ràng buộc không âm [12]. Ràng buộc  $l_0$  trên  $\boldsymbol{\alpha}$  khiến phương trình (3) trở thành một bài toán NP-khó (NP-hard problem) [16]. Do đó, phép biến đổi  $l_1$  được sử dụng để đưa về bài toán sau tối ưu sau:

$$\min_{\mathbf{S}, \boldsymbol{\alpha}} (\mathbf{1}^T \exp(\mathbf{S}\boldsymbol{\alpha}) - \mathbf{x}_0^T (\mathbf{S}\boldsymbol{\alpha})) + \lambda \|\boldsymbol{\alpha}\|_1, \quad (4)$$

Bài toán tối ưu trong phương trình (4) có thể được giải quyết bằng thuật toán co giãn ngưỡng lặp (Iterative Shrinkage Thresholding Algorithm, ISTA) [17], trong đó  $\boldsymbol{\alpha}$  và  $\mathbf{S}$  được tìm bằng cách lặp luân phiên. Việc ước lượng  $\boldsymbol{\alpha}$  được đưa về bài toán sau:

$$\operatorname{argmin}_{\boldsymbol{\alpha}} \left( \frac{1}{2} \|\mathbf{x}_0 - \mathbf{S}\boldsymbol{\alpha}\|_2^2 + \lambda \|\boldsymbol{\alpha}\|_1 \right) \quad (5)$$

Thuật toán ISTA nhằm giải quyết phương trình (5)

thông qua bước cập nhật  $\boldsymbol{\alpha} = \operatorname{Soft} \left( \left( \boldsymbol{\alpha} + \frac{1}{L} \mathbf{S}^T (\mathbf{x}_0 - \mathbf{S}\boldsymbol{\alpha}) \right) \right)$ , trong đó  $L \leq \sigma \max(\mathbf{S}^T \mathbf{S})$  và  $\operatorname{Soft}$  là toán tử ngưỡng mềm (soft threshold operator) với ngưỡng  $\zeta$ , được định nghĩa là  $\operatorname{Soft}_{\zeta}(\mathbf{x}) = \operatorname{sign}(\mathbf{x}) \max(|\mathbf{x}| - \zeta, 0)$  [17]. Thuật toán ISTA tinh chỉnh mã thưa với tính toán một cách lặp lại cho đến khi đạt tiêu chí hội tụ.

Tuy nhiên, trong bài toán nghiên cứu, ảnh đầu vào có kích thước  $d \times d$ , và không gian  $\mathbf{S}$  phụ thuộc vào kích thước của hình ảnh đầu vào. Để giải quyết vấn đề này, mô hình mã hóa thưa bằng mạng tích chập thay thế phép nhân ma trận-vector bằng phép tích chập (ký hiệu là  $*$ ) [6]:

$$\mathbf{S}\boldsymbol{\alpha} = \sum_j^M \mathbf{S}_j * \mathbf{F}_j = \mathbf{S} * \mathbf{F} \quad (6)$$

với  $\mathbf{S}_j \in \mathbb{R}^{k \times k}$  là bộ lọc tích chập quanh bản đồ đặc trưng thưa  $\mathbf{F}_j \in \mathbb{R}^{d \times d}$  với số lượng  $M$  trong không gian  $\mathbf{S}$ . Dạng mới này của mã thưa (6) và phương pháp áp dụng không gian thưa tách biệt với kích thước ảnh đầu vào giúp loại bỏ yêu cầu mở rộng mô hình theo kích thước ảnh và hỗ trợ bước cập nhật tiếp theo như sau:

$$\mathbf{F}_i = \operatorname{Soft}(\mathbf{F}_{i-1} + \mathbf{S}^T * (\mathbf{X}_0 - \mathbf{S} * \mathbf{F}_{i-1})) \quad (7)$$

Do đó, việc tối ưu hóa  $\boldsymbol{\alpha}$  trong phương trình (4) được viết lại như sau:

$$\min_{\mathbf{S}} (\exp(\mathbf{S} * \mathbf{F}) - \mathbf{X}_0 \odot (\mathbf{S} * \mathbf{F})) \quad (8)$$

Trong đó,  $\odot$  là tích Hadamard [18]. Để giải phương trình (8),  $\mathbf{S} * \mathbf{F}$  được biểu diễn qua một mạng NN  $\mathfrak{F}_{\theta}$ . Mạng  $\mathfrak{F}_{\theta}$  gồm một bộ mã hóa và bộ giải mã duy nhất, tính toán biểu diễn thưa  $\mathbf{F}$  dựa trên các tham số của mạng, cho phép học các không gian  $\mathbf{S}$  thông qua lan truyền ngược, được biểu diễn như sau:

$$\mathbf{S} * \mathbf{F} = \mathfrak{F}_{\theta}(\mathbf{X}_0)$$

Bài toán tối ưu (8) được điều chỉnh nhằm mục đích áp dụng tính toán thông qua cấu trúc mạng một cách gián tiếp như sau:

$$\min_{\theta} (\exp(\mathfrak{F}_{\theta}(\mathbf{X}_0)) - \mathbf{X}_0 \odot \mathfrak{F}_{\theta}(\mathbf{X}_0))$$

Việc tính toán bước cập nhật trong thuật toán ISTA có thể xấp xỉ bằng cách thay thế  $\mathbf{S}$  và  $\mathbf{S}^T$  lần lượt bằng bộ giải mã và bộ mã hóa cho phương trình (7) như sau [9]:

$$\mathbf{F}_i = \operatorname{Soft} \left( \mathbf{F}_{i-1} + \operatorname{Encoder}(\mathbf{X}_0 - \operatorname{Decoder}(\mathbf{F}_{i-1})) \right) \quad (9)$$

Thuật toán ISTA được xấp xỉ bằng một số lượng bước lặp hữu hạn, trong đó quá trình cập nhật truyền thông được điều chỉnh dựa trên phương trình (9). Sự điều chỉnh này dẫn đến việc thiết kế một cấu trúc tương tự mạng no-ron hồi quy, trong đó mã thưa được tinh chỉnh qua một số bước hữu hạn cụ thể [9]. Quá trình này được thể hiện dưới dạng truyền xuôi của mạng NN, cho phép mã thưa được cập nhật dần qua một số bước hữu hạn, thay vì tiếp tục lặp đến khi hội tụ như ISTA truyền thống.

### 2.3. Học tương phản

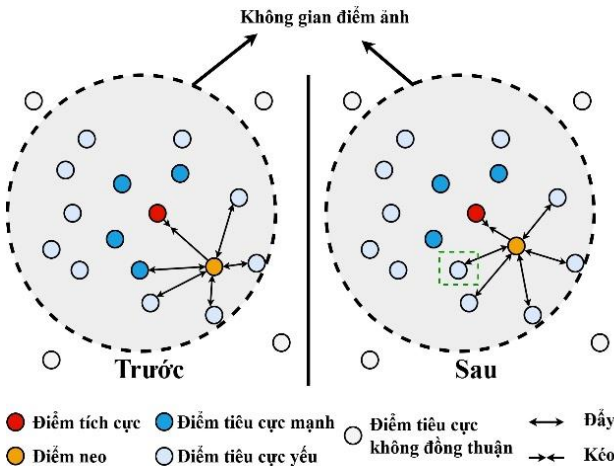
Để áp dụng học tương phản [19] vào bài toán khử nhiễu Poisson cường độ lớn, điểm neo (anchor) là kết quả khôi phục từ mạng PDSCL, điểm tích cực (positive) là ảnh thực (ground truth), còn các điểm tiêu cực (negative) gồm ảnh nhiễu đầu vào và các ảnh nhiễu không tương đồng với ảnh thực. Mục tiêu là giảm khoảng cách giữa đặc trưng của

điểm neo và điểm tích cực, đồng thời tăng khoảng cách giữa điểm neo và điểm tiêu cực, được mô hình hóa qua hàm mục tiêu sau với  $\mathbf{X} \in \mathbb{R}^{d \times d}$  là nhân thực của ảnh nhiễu Poisson  $\mathbf{X}_0 \in \mathbb{R}^{d \times d}$ :

$$R = \sum_{i=1}^n \xi_i \frac{\|V_i(\mathbf{X}) - V_i(\mathfrak{I}_\theta(\mathbf{X}_0))\|_1}{\|V_i(\mathbf{U}_q) - V_i(\mathfrak{I}_\theta(\mathbf{X}_0))\|_1 + E_i} \quad (10)$$

Trong đó,  $E_i = \|V_i(\mathbf{X}) - V_i(\mathfrak{I}_\theta(\mathbf{X}_0))\|_1$ ,  $\mathfrak{I}_\theta(\cdot)$  biểu thị mạng khử nhiễu Poisson với các tham số  $\theta$ ,  $V_i(\cdot)$ ,  $i = 1, 2, \dots, n$  là lớp trích xuất đặc trưng ẩn thứ  $i$  từ mạng VGG19 đã được huấn luyện trước [20], số lượng các điểm tiêu cực không đồng thuận (Non-consensual Negative)  $\{\mathbf{U}_q\}$  là  $r$ , và  $\{\xi_i\}$  là tập hợp các siêu tham số. Không gian không đồng thuận là tập hợp các ảnh có nội dung giống nhau nhưng khác biệt về phân bố và mức nhiễu. Điểm tiêu cực không đồng thuận là các ảnh thuộc không gian này [19].

Như minh họa ở Hình 1, khoảng cách giữa điểm neo và các điểm tiêu cực không đồng thuận không tạo được giới hạn dưới thỏa đáng cho không gian nghiệm. Vì các điểm tiêu cực này thường cách xa điểm tích cực, không gian nghiệm bị giới hạn kém, làm giảm chất lượng khôi phục ảnh.



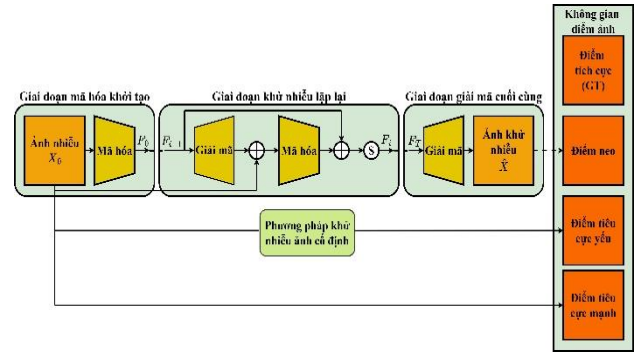
Hình 1. Minh họa về học tương phân tăng cường

#### 2.4. Phương pháp đề xuất

Nghiên cứu này đề xuất một phương pháp học tương phân tăng cường cho bài toán khử nhiễu Poisson, tận dụng các điểm tiêu cực trong không gian đồng thuận. Các điểm tiêu cực bao gồm các kết quả khử nhiễu từ các phương pháp khác, nhằm mở rộng không gian nghiệm. Mục tiêu là tối đa hóa khoảng cách (10) giữa điểm neo và các điểm tiêu cực có chất lượng. Tuy nhiên, hai thách thức quan trọng cần giải quyết là xác định độ khó của các điểm tiêu cực, sắp xếp các điểm tiêu cực theo mức độ khó trong quá trình huấn luyện.

Dựa trên ảnh đầu vào nhiễu Poisson, ảnh khử nhiễu được tạo ra tương ứng thông qua mô hình được đề xuất cho khử nhiễu Poisson (PDSCL) xây dựng từ phương trình (9) theo mô tả ở Hình 2. Với ảnh nhiễu Poisson là  $\mathbf{X}_0 \in \mathbb{R}^{d \times d}$  và dạng vector hóa của nó là  $\mathbf{x}_0 \in \mathbb{R}^{d^2}$ . Ảnh nhiễu đầu vào được mã hóa khởi tạo (Initial Encode Phase) thành các biểu diễn thưa, sau đó thuật toán tối ưu lặp (Iterative Denoise Phase) được sử dụng dựa trên các lớp tích chập  $3 \times 3$  với kích thước đệm 1 cho các lớp Encoder và tích chập ngược  $3 \times 3$  với kích thước đệm 1 cho các lớp Decoder để học

các biểu diễn thưa được mã hóa qua hình ảnh nhiễu. Cuối cùng, các biểu diễn thưa được giải mã để nhận được ảnh khử nhiễu cuối cùng (Final Decode Phase).



Hình 2. Kiến trúc mô hình PDSCL đề xuất

Độ khó của điểm tiêu cực được chia thành hai mức: yếu (weak negative) và mạnh (strong negative). Ảnh nhiễu đầu vào là điểm tiêu cực yếu, trong khi độ khó của các điểm tiêu cực khác được xác định linh động trong quá trình huấn luyện. Trước mỗi giai đoạn xử lý, giá trị trung bình của thang đo chất lượng hình ảnh PSNR, gọi là tỷ số tín hiệu cực đại trên nhiễu (peak signal-to-noise ratio, PSNR) [21], được xác định. Nếu PSNR của một điểm cao hơn mức trung bình, nó là điểm mạnh; nếu thấp hơn, nó là điểm yếu. Các điểm tiêu cực được gán trọng số khác nhau tùy theo độ khó. Điểm mạnh giúp thu hẹp không gian nghiệm nhưng có thể gây mơ hồ, trong khi điểm yếu cần trọng số đủ lớn để hướng lực tác động về phía điểm tích cực. Trọng số của một điểm tiêu cực  $S_q$  tại giai đoạn thứ  $t$  được định nghĩa như sau:

$$W_t(S_q) = \begin{cases} 1 + \gamma & \text{if } \text{avgPSNR}(\mathfrak{I}_{\theta-1}(\{\mathbf{X}_0\})) \geq \text{PSNR}(S_q) \\ 1 - \gamma & \text{otherwise} \end{cases} \quad (11)$$

Trong đó,  $\{\mathbf{X}_0\}$  biểu thị tập ảnh nhiễu đầu vào,  $q = 1, 2, \dots$  là chỉ số của các mẫu tiêu cực mạnh, và  $\gamma$  là một siêu tham số.

Trọng số của các mẫu tiêu cực mạnh và yếu lần lượt được thiết lập trong (11) là  $1 + \gamma$  và  $1 - \gamma$ . Trong đó, trọng số của mẫu tiêu cực yếu lớn hơn trọng số của mẫu tiêu cực mạnh, nhằm giảm bớt sự mơ hồ trong quá trình học. Hơn nữa, tính linh hoạt trong việc xác định các mức độ khó khăn cho phép các mẫu tiêu cực mạnh chuyển thành mẫu tiêu cực yếu trong các bước tiếp theo của quá trình huấn luyện. Khi chất lượng của điểm neo được cải thiện, sự mơ hồ do các mẫu tiêu cực mạnh gây ra sẽ giảm, và tầm quan trọng của chúng nên được tăng cường. Theo cách này, các mẫu tiêu cực mạnh có thể được xem như các giới hạn dưới tối ưu hơn để ràng buộc không gian nghiệm một cách hiệu quả. Do đó, việc điều chỉnh tăng cường  $R^*$  được xác định như sau:

$$R^* = \sum_{i=1}^n \xi_i \frac{\|V_i(\mathbf{X}) - V_i(\mathfrak{I}_\theta(\mathbf{X}_0))\|_1}{\sum_{q=1}^r W_t(S_q) \|V_i(\mathbf{U}_q) - V_i(\mathfrak{I}_\theta(\mathbf{X}_0))\|_1 + E_i} \quad (12)$$

Cuối cùng, hàm mục tiêu  $L$  của mô hình bao gồm khoảng cách  $l_1$  và điều chỉnh tăng cường  $R^*$  (12) với trọng số  $\lambda > 0$ , được biểu diễn như sau:

$$L = \|\mathbf{X} - \mathfrak{I}_\theta(\mathbf{X}_0)\|_1 + \lambda R^* \quad (13)$$

Thành phần  $\|\mathbf{X} - \mathfrak{I}_\theta(\mathbf{X}_0)\|_1$  trong (13) đo lường sự khác biệt giữa đầu ra của mô hình  $\mathfrak{I}_\theta(\mathbf{X}_0)$  và dữ liệu thực tế  $\mathbf{X}$  bằng cách sử dụng chuẩn  $l_1$ , đảm bảo rằng mô hình tạo kết quả gần sát nhất với dữ liệu thực tế. Việc thiết kế hàm

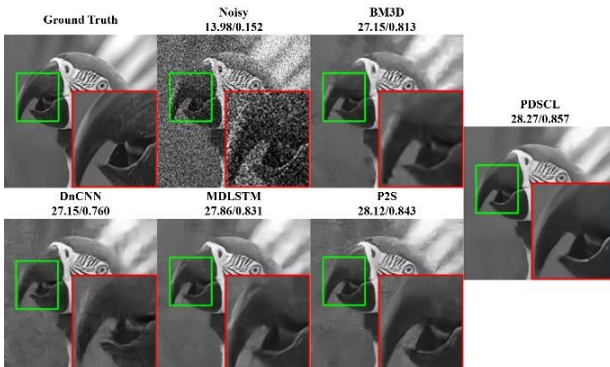
mục tiêu  $L$  bằng cách phối hợp hai thành phần này giúp đảm bảo mô hình vừa tái tạo chính xác dữ liệu đầu vào, vừa đáp ứng ràng buộc  $R^*$ .

### 3. Thực nghiệm

Tập dữ liệu hình ảnh nhiễu tổng hợp (15,882 hình ảnh trắng đen) được cấu thành từ các bộ dữ liệu chuẩn: 432 hình ảnh từ BSD [22], 800 hình ảnh từ DIV2K [23], và 2,650 hình ảnh từ Flickr2K [24]. Để so sánh kết quả, tập kiểm thử gồm Set12 [25] và BSD68 [22] được sử dụng. Bộ dữ liệu FMD cũng được dùng để đánh giá, gồm hình ảnh kính hiển vi huỳnh quang với nhiễu Poisson [26]. Thử nghiệm được thực hiện trên bốn mẫu kính hiển vi hai photon (MICE, BPAE-B, BPAE-G, BPAE-R), với 250 ảnh FOV thứ 20, giá trị thực được lấy bằng cách trung bình 50 lần chụp. Mô hình được huấn luyện trong 1500 vòng lặp. Hệ số học ban đầu là  $10^{-3}$  và tăng dần lên  $5 \times 10^{-3}$  trong 100 vòng lặp đầu tiên, sau đó giảm dần về  $10^{-5}$  theo hàm cosine trong phần còn lại của quá trình huấn luyện. Chất lượng hình ảnh khử nhiễu được đánh giá thông qua các thang đo PSNR (dB) và SSIM [21].

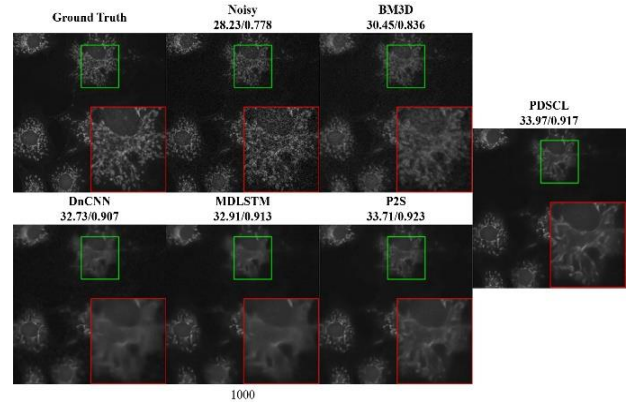
Thực nghiệm được thực hiện trên nền tảng Ubuntu 20.04, với PyTorch 1.12.1, Python 3.10.4, và GPU NVIDIA DGX A100 hỗ trợ CUDA 12.1. Hệ thống sử dụng CPU Intel Xeon Platinum 8470Q. Trong thực nghiệm, mô hình PDSCL được so sánh với các phương pháp khử nhiễu Poisson nổi bật như BM3D [27], DnCNN [25], Multi-directional LSTM (MDLSTM) [28], và Poisson2Sparse (P2S) [29].

So sánh trực quan trên tập BSD68 được trình bày trong Hình 3. Các tạo tác giả và hiện tượng mờ xuất hiện rõ ràng trong kết quả của BM3D và DnCNN. Kết quả từ MDLSTM và P2S cũng khá mờ và chứa một số chi tiết không rõ ràng. Nhờ khả năng tổng quát hóa vượt trội, PDSCL đạt hiệu quả trực quan tốt nhất so với các phương pháp được so sánh. Kết quả từ Bảng 1 cho thấy, PDSCL vượt trội về SSIM ở nhiều giá trị  $\lambda$ , thể hiện khả năng khôi phục chi tiết hình ảnh tốt hơn. Phương pháp PDSCL đạt PSNR cao hơn ở hầu hết các trường hợp, đặc biệt trên tập Set12. Với  $\lambda = 1$ , PDSCL đạt PSNR=23,65 dB và SSIM=0,687 cao hơn đáng kể so với Poisson2Sparse với SSIM = 0,647. Trên BSD68, PDSCL cho PSNR= 22,20 dB và SSIM=0,540 tại  $\lambda = 1$ , tốt hơn so với các phương pháp được so sánh. Với  $\lambda = 10$ , PDSCL đạt PSNR= 25,60 dB và SSIM=0,708. Điều này chứng minh tính ổn định và hiệu quả khử nhiễu của phương pháp đề xuất.



Hình 3. Kết quả khử nhiễu trên tập BSD68 với  $\lambda = 4$

So sánh trực quan trên tập FMD được biểu diễn trong Hình 4. Nhiễu nhỏ vẫn tồn tại trong kết quả của BM3D, làm giảm khả năng khử nhiễu Poisson của nó. Mặc dù, DnCNN, MDLSTM và P2S có thể giảm nhiễu Poisson hiệu quả, nhưng các cạnh trong kết quả của DnCNN và MDLSTM bị mờ. Nhìn chung, PDSCL cho kết quả khử nhiễu với các chi tiết rõ ràng hơn so với các phương pháp được so sánh. Kết quả Bảng 2 cũng cho thấy, hiệu năng nổi bật của PDSCL, đặc biệt trên mẫu BPAE-R và MICE, nơi PSNR và SSIM đạt cao nhất. Trên mẫu MICE, PDSCL đạt PSNR=33,96 dB và SSIM=0,919, cao nhất trong các phương pháp, chỉ kém nhẹ P2S về SSIM nhưng vượt trội về PSNR (SSIM=0,920, PSNR=33,69 dB). Với mẫu BPAE-R, PDSCL đạt PSNR=38,78 dB và SSIM= 0,941, vượt xa BM3D (PSNR= 37,17 và SSIM=0,929). Mặc dù trên mẫu BPAE-G, PDSCL không đạt hiệu suất cao nhất, nhưng vẫn giữ được hiệu suất cạnh tranh so với các phương pháp được so sánh. Điều này chứng minh thêm rằng PDSCL là một phương pháp mạnh mẽ, ổn định và hiệu quả trong nhiều điều kiện đánh giá.



Hình 4. Kết quả khử nhiễu trên tập FMD

Bảng 1. Kết quả đánh giá trên tập Set12 và BSD68 với thang đo PSNR và SSIM

Bộ dữ liệu	Phương pháp	$\lambda = 1$	$\lambda = 4$	$\lambda = 10$
Set12	BM3D	22,70/0,649	25,16/0,740	26,46/0,792
	DnCNN	22,58/0,619	25,55/0,729	26,99/0,766
	MDLSTM	22,89/0,639	26,01/0,749	27,65/0,800
	P2S	22,98/0,647	26,10/0,757	27,76/0,803
	PDSCL	<b>23,65/0,687</b>	<b>26,78/0,785</b>	<b>28,40/0,825</b>
BSD68	BM3D	21,33/0,511	23,22/0,611	24,35/0,672
	DnCNN	21,56/0,488	23,75/0,594	24,87/0,649
	MDLSTM	21,77/0,505	24,00/0,620	25,29/0,685
	P2S	21,81/0,510	24,08/0,625	25,29/0,682
	PDSCL	<b>22,20/0,540</b>	<b>24,38/0,650</b>	<b>25,60/0,708</b>

Bảng 2. Kết quả đánh giá trên tập FMD với thang đo PSNR và SSIM

Phương pháp	MICE	BPA E-B	BPAE-G	BPAE-R
BM3D	31,20/0,863	33,52/0,887	33,23/0,847	37,17/0,929
DnCNN	33,04/0,913	37,08/0,946	33,33/0,802	36,80/0,907
MDLSTM	33,41/0,915	37,39/0,948	33,69/0,812	37,27/0,914
P2S	33,69/0,920	37,41/0,949	34,84/0,857	38,68/0,940
PDSCL	<b>33,96/0,919</b>	<b>37,44/0,948</b>	<b>35,33/0,837</b>	<b>38,78/0,941</b>



#### 4. Kết luận

Nghiên cứu này giới thiệu phương pháp PDSCL để khử nhiễu Poisson trong ảnh đơn, dựa trên mạng NN sâu học biểu diễn thưa, khai thác các đặc điểm lặp lại và tương tự trong ảnh. Phương pháp tích hợp học tương phản tăng cường, sử dụng các ví dụ tiêu cực trong không gian đồng thuận để tối ưu hóa quá trình khử nhiễu. Bên cạnh đó, chiến lược học được đề xuất với sự sắp xếp ví dụ tiêu cực theo mức độ khó, cải thiện hiệu quả mô hình. Các kết quả thực nghiệm cho thấy, PDSCL khử nhiễu Poisson trên hình ảnh có hiệu suất cạnh tranh so với các phương pháp được so sánh.

#### TÀI LIỆU THAM KHẢO

- [1] L. Chen, X. Lu, J. Zhang, X. Chu, and C. Chen, "Hinet: Half instance normalization network for image restoration", in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Nashville, USA, 2021, pp. 182-192.
- [2] S. W. Zamir *et al.*, "Multi-stage progressive image restoration", in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Nashville, USA, 2021, pp. 14821-14831.
- [3] A. Moshkov, *Confocal Laser Scanning Microscopy of Living Cells: Fluorescence Methods for Investigation of Living Cells and Microorganisms*, IntechOpen publisher, 2020.
- [4] S. Aldaz, L. M. Escudero, and M. Freeman, "Live imaging of drosophila imaginal disc development", *Proc. Natl. Acad. Sci. U.S.A.*, vol. 107, no. 32, pp. 14217-14222, 2010. <https://doi.org/10.1073/pnas.1008623107>.
- [5] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep Image Prior", *Int. J. Comput. Vis.*, vol. 128, pp. 1867-1888, 2020. <https://doi.org/10.1007/s11263-020-01303-4>
- [6] D. Simon and M. Elad, "Rethinking the CSC model for natural images", in *Proc. Adv. Neural Inf. Process. Syst.*, Vancouver, Canada, 2019, pp. 2271-2281.
- [7] A. Aich, A. Gupta, R. Panda, R. Hyder, M. S. Asif, and A. K. R. Chowdhury, "Non-adversarial video synthesis with learned priors", in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Seattle, USA, 2020, pp. 6090-6099.
- [8] A. Krull, T. O. Buchholz, and F. Jug, "Noise2void-learning denoising from single noisy images", in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Long Beach, USA, 2019, pp. 2129-2137.
- [9] J. Lehtinen *et al.*, "Noise2noise: Learning image restoration without clean data", in *Proc. 35th Int. Conf. Mach. Learn.*, Stockholm Sweden, 2018, pp. 2965-2974.
- [10] Y. Guo *et al.*, "HCSC: Hierarchical Contrastive Selective Coding", in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, New Orleans, USA, 2022, pp. 9696-9705.
- [11] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, "Momentum Contrast for Unsupervised Visual Representation Learning", in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Seattle, USA, 2020, pp. 9726-9735.
- [12] J. Salmon, Z. Harmany, C.-A. Deledalle, and R. Willett, "Poisson noise reduction with non-local pca", *J. Math. Imaging Vis.*, vol. 48, pp. 279-294, 2014. <https://doi.org/10.1007/s10851-013-0435-6>
- [13] M. Raginsky, R. M. Willett, Z. T. Harmany, and R. F. Marcia, "Compressed sensing performance bounds under poisson noise", *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 3990-4002, 2010. <https://doi.org/10.1109/TSP.2010.2049997>
- [14] A. G. Yagola and Y. M. Korolev, "Error estimation in ill-posed problems in special cases", in *Proc. Appl. Inverse Probl., Springer Proc. Math. Stat.*, NY, USA, 2013, pp. 155-164.
- [15] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration", in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Kyoto, Japan, 2009, pp. 2272-2279.
- [16] D. Ge, X. Jiang, and Y. Ye, "A note on the complexity of lp minimization", *Math. Program.*, vol. 129, pp. 285-299, 2011. <https://doi.org/10.1007/s10107-011-0470-2>
- [17] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint", *Commun. Pure Appl. Math.*, vol. 57, no. 11, pp. 1413-1457, 2004. <https://doi.org/10.1002/cpa.20042>
- [18] R. Horn, "The Hadamard product", *Sympos. Appl. Math. Amer. Math. Soc.*, vol. 40, pp. 87-169, 1990. <https://doi.org/10.1090/psapm/040>
- [19] P. Khosla *et al.*, "Supervised contrastive learning", in *Proc. 34th Int. Conf. Neural Inf. Process. Syst.*, NY, USA, 2020, pp. 18661-18673.
- [20] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition", in *Proc. 3rd Int. Conf. Learn. Represent.*, San Diego, USA, 2015, pp. 1-14.
- [21] Z. Wang and A. C. Bovik, *Modern Image Quality Assessment: Synthesis Lectures on Image, Video, and Multimedia Processing*, Morgan and Claypool Publishers, 2006.
- [22] 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 *Proc. 8th IEEE Int. Conf. Comput. Vis. (ICCV 2001)*, Vancouver, Canada, 2001, pp. 416-423.
- [23] E. Agustsson and R. N. Timofte, "NTIRE 2017 Challenge on Single Image Super-Resolution: Dataset and Study", in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Honolulu, USA, 2017, pp. 1122-1131.
- [24] R. Timofte, E. Agustsson, L. Van Gool, M.-H. Yang, and L. Zhang, "Ntire 2017 challenge on single image super-resolution: Methods and results", in *Proc. IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Honolulu, USA, 2017, pp. 1110-1121.
- [25] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising", *IEEE Trans. Image Process.*, vol. 26, no. 7, 2017, pp. 3142-3155. <https://doi.org/10.1109/TIP.2017.2662206>
- [26] Y. Zhang *et al.*, "A poisson-gaussian denoising dataset with real fluorescence microscopy images". in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Long Beach, USA, 2019, pp. 11702-11710.
- [27] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering", *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080-2095, 2007. <https://doi.org/10.1109/TIP.2007.901238>
- [28] W. Kumwilaisak, T. Piriayatharawat, P. Lasang, and N. Thatphithakkul, "Image denoising with deep convolutional neural and multi-directional long short-term memory networks under poisson noise environments", *IEEE Access*, vol. 8, pp. 86998-87010, 2020. <https://doi.org/10.1109/ACCESS.2020.2991988>
- [29] C. K. Ta, A. Aich, A. Gupta, and A. K. Roy-Chowdhury, "Poisson2sparse: Self-supervised poisson denoising from a single image", in *Proc. Med. Image Comput. Comput. Assist. Interv.*, Singapore, 2022, pp. 557-567.