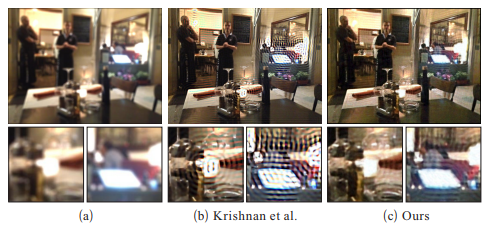
**Deep Convolutional Neural Network for Image Deconvolution (Mạng lưới thần kinh Convolutional sâu cho hình ảnh Giải mã)**

**Abstrac**

Nhiều vấn đề cơ bản liên quan đến hình ảnh liên quan đến toán tử giải mã. Thực suy giảm độ mờ hiếm khi tuân thủ một mô hình tích chập tuyến tính lý tưởng do tiếng ồn máy ảnh, độ bão hòa, nén hình ảnh, để đặt tên một vài. Thay vì mô hình hóa hoàn hảo các ngoại lệ, điều khá thách thức từ quan điểm mô hình tổng quát, chúng tôi phát triển một mạng lưới thần kinh tích chập sâu để nắm bắt các đặc điểm của sự xuống cấp. Chúng tôi lưu ý trực tiếp áp dụng các mạng thần kinh sâu hiện tại không tạo ra kết quả hợp lý. Giải pháp của chúng tôi là thiết lập kết nối giữa các sơ đồ dựa trên tối ưu hóa điều kiện và kiến trúc mạng thần kinh trong đó một cấu trúc mới, có thể tách rời được giới thiệu như một sự hỗ trợ đáng tin cậy cho việc giải mã mạnh mẽ chống lại các tạo tác. Mạng của chúng tôi chứa hai mô hình con, cả hai đều được đào tạo về một cách giám sát với khởi tạo thích hợp. Chúng mang lại hiệu suất tốt đối với việc giải mã hình ảnh không mù so với các phương pháp dựa trên mô hình thế hệ trước.

**1.Introduction**

Nhiều quá trình suy giảm hình ảnh và video có thể được mô hình hóa thành tích chập dịch bất biến. Để khôi phục các dữ liệu trực quan này, quá trình nghịch đảo, tức là giải mã, trở thành một công cụ quan trọng trong chuyển động làm mờ [1, 2, 3, 4], siêu phân giải [5, 6] và độ sâu trường mở rộng [7]. Trong các ứng dụng liên quan đến hình ảnh được chụp bởi máy ảnh, các ngoại lệ như bão hòa, hình ảnh bị hạn chế ranh giới, tiếng ồn, hoặc các tạo tác nén là không thể tránh khỏi. Nghiên cứu trước đây đã chỉ ra rằng việc xử lý không đúng các vấn đề này có thể tạo ra một loạt các hiện vật liên quan đến nội dung hình ảnh, trong đó rất khó để loại bỏ. Vì vậy, đã có công việc dành riêng cho mô hình và giải quyết từng cụ thể loại tạo tác trong giải mã không mù để triệt tiêu các tạo tác chuông [8], loại bỏ tiếng ồn [9], và đối phó với các vùng bão hòa [9, 10]. Những phương pháp này có thể được cải tiến thêm bằng cách kết hợp thống kê cấp độ bản vá [11] hoặc các sơ đồ khác [4]. Bởi vì mỗi phương pháp đều có đặc sản riêng của nó. như giới hạn, không có giải pháp nào để giải quyết thống nhất tất cả các vấn đề này. Một ví dụ được hiển thị trong Hình 1 - một hình ảnh mờ bão hòa một phần với các lỗi nén có thể đã thất bại nhiều cách tiếp cận. Một khả năng để loại bỏ các cổ vật này là thông qua việc sử dụng các mô hình thế hệ. Tuy nhiên, những mô hình thường được thực hiện dựa trên các giả định mạnh mẽ, chẳng hạn như tiếng ồn phân tán giống hệt và độc lập, mà có thể không giữ cho hình ảnh thực sự. Điều này cho thấy thực tế là ngay cả các thuật toán tiên tiến cũng có thể bị ảnh hưởng khi các thuộc tính mờ hình ảnh được thay đổi một chút.



Hình 1: Một ví dụ giải mã đầy thách thức. (a) là đầu vào mờ với độ bão hòa một phần vùng. (b) là kết quả của [3] khi sử dụng hyper-Laplacian trước đó. (c) là kết quả của chúng tôi.

Trong bài báo này, chúng tôi bắt đầu quy trình giải mã hình ảnh tự nhiên không dựa trên vật lý của chúng hoặc đặc điểm dựa trên toán học. Thay vào đó, chúng tôi cho thấy một hướng mới để xây dựng một hướng dữ liệu hệ thống sử dụng các mẫu hình ảnh có thể dễ dàng được tạo ra từ máy ảnh hoặc được thu thập trực tuyến. Chúng tôi sử dụng mạng nơ ron tích chập (CNN) để tìm hiểu hoạt động giải mã mà không cần cần biết nguyên nhân của tạo tác thị giác. Chúng tôi cũng không dựa vào bất kỳ quá trình tiền xử lý nào để làm mờ hình ảnh, Không giống như cách tiếp cận dựa trên học tập trước đây [12, 13]. Trong thực tế, việc tìm một mạng phù hợp là không tầm thường kiến trúc để giải mã. Mạng thần kinh khử nhiễu trước đây [14, 15, 16] không thể trực tiếp được thông qua kể từ khi giải mã có thể liên quan đến nhiều pixel lân cận và dẫn đến kết quả rất phức tạp hàm năng lượng với suy giảm phi tuyến. Điều này làm cho việc học thông số khá khó khăn. Trong công việc của chúng tôi, chúng tôi thu hẹp khoảng cách giữa một mạng lưới thần kinh tích chập xác định theo kinh nghiệm và các cách tiếp cận hiện có với các mô hình thế hệ trong bối cảnh giả nghịch đảo của giải mã. Nó cho phép một hệ thống thực tế và quan trọng hơn là cung cấp một chiến lược hiệu quả theo kinh nghiệm để khởi tạo các trọng số trong mạng mà không thể dễ dàng đạt được trong thông thường thủ tục đào tạo khởi tạo ngẫu nhiên. Các thử nghiệm cho thấy hệ thống của chúng tôi vượt trội so với trước đây những người đặc biệt là khi hình ảnh đầu vào mờ được bão hòa một phần.

**2. Related Work**

Deconvolution được nghiên cứu trong các lĩnh vực khác nhau do tính cơ bản của nó trong phục hồi hình ảnh. Phần lớn các phương pháp trước đây giải quyết vấn đề từ góc độ tổng quát giả sử nhiễu hình ảnh đã biết mô hình và độ dốc hình ảnh tự nhiên sau khi phân phối nhất định. Trong phương pháp Richardson-Lucy [17], nhiễu hình ảnh được giả sử tuân theo phân phối Poisson. Wiener Deconvolution [18] áp đặt giả định Gaussian tương đương cho cả độ nhiễu và độ dốc hình ảnh. Những cách tiếp cận ban đầu này phải chịu đựng các cạnh quá mịn và tạo tác chuông. Sự phát triển gần đây về giải mã cho thấy các thuật ngữ chính quy hóa với các mục sư hình ảnh thưa thớt là quan trọng để bảo tồn các cạnh sắc nét và triệt tiêu các hiện vật. Các linh mục hình ảnh thưa thớt theo đuôi nặng các bản phân phối, chẳng hạn như Mô hình hỗn hợp Gaussian [1, 11] hoặc siêu Laplacian [7, 3], có thể là tối ưu hóa hiệu quả bằng cách sử dụng phân tách nửa bậc hai (HQ) [3]. Để chụp số liệu thống kê hình ảnh lớn hơn Hỗ trợ không gian, năng lượng được mô hình hóa thêm trong khuôn khổ Trường ngẫu nhiên có điều kiện (CRF) [19] và trên các bản vá hình ảnh [11]. Trong khi bước cuối cùng của phương pháp HQ là tối ưu hóa bậc hai, Schmidt et al. [4] cho thấy có thể huấn luyện trực tiếp CRF Gaussian từ dữ liệu mờ tổng hợp. Để xử lý các ngoại lệ như bão hòa, Cho et al. [9] đã sử dụng EM đa dạng để loại trừ các vùng xa hơn từ một khả năng Gaussian. Whyte et al. [10] đã giới thiệu một biến phụ trợ trong phương pháp RichardsonLucy. Một đường chuyền khử nhiễu rõ ràng được thêm vào để giải mã, trong đó cách tiếp cận khử nhiễu được thiết kế cẩn thận [20] hoặc được đào tạo từ dữ liệu ồn ào [12]. Các cách tiếp cận khái quát thường có khó khăn để xử lý các ngoại lệ phức tạp không độc lập và phân phối giống hệt nhau.

Một xu hướng khác để phục hồi hình ảnh là tận dụng cấu trúc mạng thần kinh sâu và dữ liệu lớn để đào tạo chức năng phục hồi. Do đó, sự xuống cấp không còn giới hạn trong một mô hình liên quan đến nhiễu hình ảnh. Burger và cộng sự. [14] cho thấy rằng các tri giác đa lớp đơn giản có thể tạo ra phong nha kết quả và xử lý các loại tiếng ồn khác nhau. Xie et al. [15] cho thấy cấu trúc bộ khử nhiễu tự động (SDAE) xếp chồng lên nhau [21] là một lựa chọn tốt cho khử nhiễu và inpainting. Agostinelli et al. [22] khái quát hóa nó bằng cách kết hợp nhiều SDAE để xử lý các loại tiếng ồn khác nhau. Trong [23] và [16], kiến trúc mạng nơ ron tích chập (CNN) [24] đã được sử dụng để xử lý nhiễu mạnh nhưhạt mưa và bụi bẩn ống kính. Schuler và cộng sự. [13] đã thêm MLP vào một giải mã trực tiếp để loại bỏ các vật phẩm. Mặc dù cấu trúc mạng hoạt động tốt đối với khử nhiễu, nhưng nó không hoạt động tương tự cho quá trình giải mã. Làm thế nào để thích ứng kiến trúc là vấn đề chính cần giải quyết trong bài viết này.

**3. Blur Degradation**

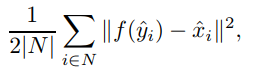
Chúng tôi xem xét mờ hình ảnh trong thế giới thực bị một số loại suy thoái bao gồm cắt xén cường độ (bão hòa), tiếng ồn máy ảnh và tạo tác nén. Mô hình mờ được đưa ra bởi :

(1)

Trong đó αx đại diện cho hình ảnh sắc nét tiềm ẩn. Ký hiệu  là để chỉ ra thực tế rằng αx có thể có các giá trị vượt quá phạm vi động của cảm biến camera và do đó được cắt bớt. k là người biết hạt nhân chập, hay thường được gọi là hàm lây lan điểm (PSF), n mô hình phụ gia tiếng ồn máy ảnh.  là một hàm cắt để bão hòa mô hình, được định nghĩa là trong đó  là một ngưỡng phạm vi.  là toán tử nén phi tuyến (ví dụ: JPEG). Chúng tôi lưu ý rằng ngay cả với  và kernel k, việc khôi phục αx là không thể truy cập được, đơn giản là vì thông tin mất mát do cắt. Về vấn đề này, mục tiêu của chúng tôi là khôi phục đầu vào bị cắt , trong đó Mặc dù giải x bằng hàm năng lượng phức tạp liên quan đến phương trình. (1) là khó khăn, việc tạo ra hình ảnh mờ từ đầu vào x khá đơn giản bằng cách tổng hợp hình ảnh theo mô hình tích chập lấy tất cả các loại suy giảm hình ảnh có thể thành thế hệ. Điều này thúc đẩy một thủ tục học tập để giải mã, sử dụng các cặp hình ảnh đào tạo trong đó chỉ số 

**4. Analysis**

Mục tiêu là đào tạo một kiến trúc mạng f (·) để giảm thiểu.

(2)

Trong đó | N | là số lượng cặp ảnh trong tập mẫu. Chúng tôi đã sử dụng hai mạng lưới thần kinh sâu gần đây để giải quyết vấn đề này, nhưng không thành công. Một là Bộ giải mã tự động xếp chồng thưa thớt (SSDAE) [15] và cái còn lại là mạng nơ ron tích chập (CNN) được sử dụng trong [16]. Cả hai đều được thiết kế cho hình ảnh con rùa. Đối với SSDAE, chúng tôi sử dụng kích thước bản vá 17 × 17 như đề xuất trong [14]. Việc thực hiện CNN được cung cấp bởi các tác giả của [16]. Chúng tôi thu thập hai triệu bản vá sắc nét cùng với các phiên bản mờ của chúng trong đào tạo. Một ví dụ được hiển thị trong Hình 2, trong đó (a) là một hình ảnh mờ. Hình 2 (b) và (c) hiển thị kết quả của SSDAE và CNN. Kết quả của SSDAE trong (b) vẫn còn mờ. Cấu trúc CNN hoạt động tương đối tốt hơn. Nhưng nó phải chịu các cạnh mờ còn lại và các tạo tác bóng ma mạnh mẽ. Điều này là do những cấu trúc mạng là dành cho khử nhiễu và không xem xét các thuộc tính giải mã cần thiết. Hơn giải thích được cung cấp từ một quan điểm khái quát trong những gì sau đây.

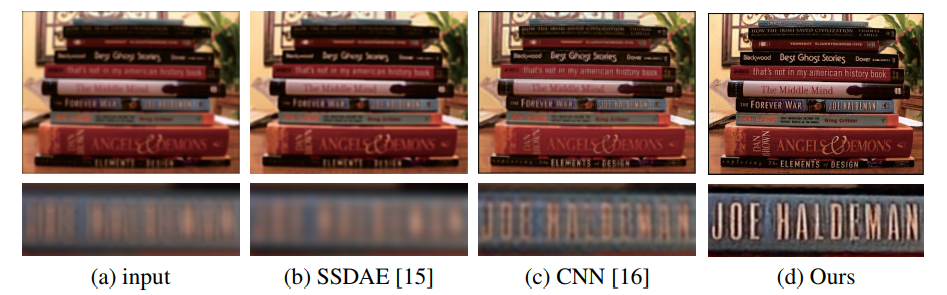
**4.1. Pseudo Inverse Kernels**

Nhiệm vụ giải mã có thể được xấp xỉ bởi một mạng chập theo bản chất. Chúng tôi xem xét theo mô hình mờ tuyến tính đơn giản:

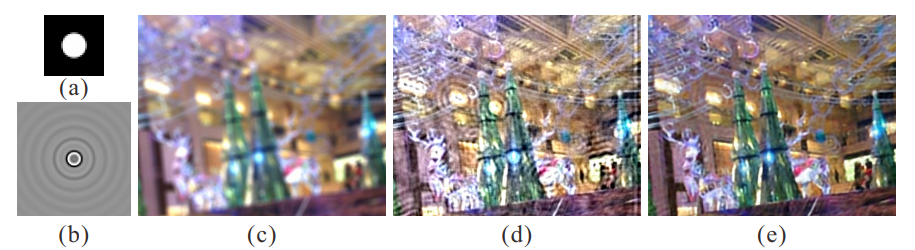


Sự tích chập không gian có thể được chuyển thành phép nhân miền tần số, mang lại:





Hình 2: Cấu trúc mạng tự động khử nhiễu xếp chồng và cấu trúc mạng nơ ron hiện tại không thể giải bài toán giải mã.

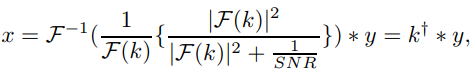


Hình 3: Các ví dụ về nhân ngược và giải mã giả.

 biểu thị biến đổi Fourier rời rạc (DFT). Toán tử · là phép nhân phần tử. Trong Miền Fourier, x có thể được lấy là:



trong  là biến đổi Fourier rời rạc ngược. Trong khi bộ giải cho x được viết dưới dạng tích chập không gian với một hạt nhân , hạt nhân thực sự là một tín hiệu lặp đi lặp lại toàn bộ miền không gian mà không có sự hỗ trợ nhỏ gọn. Khi tiếng ồn phát sinh, các điều khoản chính quy là thường liên quan để tránh chia cho 0 trong miền tần số, điều này làm cho giả nghịch đảo rơi ra nhanh chóng trong miền không gian [25]. Việc giải mã Wiener cổ điển tương đương với việc sử dụng bộ chỉnh tần Tikhonov [2]. Wiener giải mã có thể được thể hiện như:



Trong đó SNR là tỷ lệ tín hiệu trên tạp âm. k† Biểu thị hạt nhân nghịch đảo giả. Tiếng ồn mạnh dẫn đến một lớn , tương ứng với đảo ngược thường xuyên mạnh mẽ. Chúng tôi lưu ý rằng với sự giới thiệu của SNR, k Trở nên nhỏ gọn với sự hỗ trợ hữu hạn. Hình 3 (a) cho thấy một nhân mờ đĩa có bán kính 7, thường được sử dụng để mô hình mờ tiêu cự. Hạt nhân giả k† Với SNR = 1E – 4 được đưa ra trong hình 3 (b). Một hình ảnh mờ với hạt nhân này được hiển thị trong Hình 3 (c). Kết quả giải mã với k† đang ở (d). Một mức độ mờ được loại bỏ khỏi hình ảnh. Nhưng tiếng ồn và bão hòa gây ra thị giác hiện vật, phù hợp với sự hiểu biết của chúng tôi về giải mã Wiener. Mặc dù phương thức Wiener không phải là hiện đại, nhưng sản phẩm phụ của nó là hạt nhân nghịch đảo với hỗ trợ không gian hữu hạn nhưng lớn trở nên hữu ích to lớn trong hệ thống mạng thần kinh của chúng ta, điều này thể hiện sự giải mã đó có thể được xấp xỉ bằng cách tích chập không gian với các hạt nhân đủ lớn. Điều này giải thích việc áp dụng SSDA và CNN không thành công trực tiếp vào việc giải mã trong Hình 2 như sau:

• SSDA không nắm bắt tốt bản chất của tích chập với các cấu trúc được kết nối đầy đủ của nó.

• CNN hoạt động tốt hơn do quá trình giải mã có thể được xấp xỉ bằng phép chập hạt nhân lớn như đã giải thích ở trên.

• CNN trước đó sử dụng hạt nhân chập nhỏ. Tuy nhiên, đây không phải là một cấu hình phù hợp trong bài toán giải mã của chúng tôi.

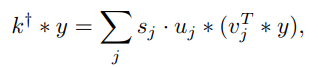
Do đó, có thể tóm tắt rằng việc sử dụng các mạng nơ ron sâu để thực hiện giải mã là không có nghĩa là đơn giản Chỉ cần sửa đổi mạng bằng cách sử dụng các hạt tích chập lớn sẽ dẫn đến để khó khăn cao hơn trong đào tạo. Chúng tôi trình bày một cấu trúc mới để cập nhật mạng theo những gì sau đây. Kết quả của chúng tôi trong hình 3 được hiển thị trong (e).

**5. Network Architecture**

Chúng tôi chuyển đổi hạt nhân giả đơn giản để giải mã thành một mạng chập dựa trên định lý phân tách nhân. Nó làm cho mạng biểu cảm hơn với ánh xạ tới kích thước cao hơn để phù hợp với phi tuyến. Hệ thống này được hưởng lợi từ dữ liệu đào tạo lớn.

**5.1. Kernel Separability**

Khả năng phân tách hạt nhân đạt được thông qua phân rã giá trị số ít (SVD) [26]. Cho nghịch đảo hạt nhân k†, phân hủy k† = USV tồn tại Chúng tôi biểu thị bởi uj và vj j cột thứ U và V, s j giá trị số ít. Việc giải mã giả ban đầu có thể được biểu thị bằng:

(3)

cho thấy tích chập 2D có thể được coi là tổng của các bộ lọc 1D có thể tách rời. Trong thực tế, chúng ta cũng có thể xấp xỉ k† Bằng một số lượng nhỏ các bộ lọc có thể tách rời bằng cách loại bỏ các hạt nhân liên quan với sj bằng 0 hoặc rất nhỏ. Chúng tôi đã thử nghiệm với các nhân mờ thực sự để bỏ qua các giá trị số ít nhỏ hơn 0,01. Số lượng trung bình của hạt nhân có thể tách được là khoảng 30 [25]. Sử dụng một tỷ lệ SNR nhỏ hơn, hạt nhân nghịch đảo có hỗ trợ không gian nhỏ hơn. Chúng tôi cũng thấy rằng một nghịch đảo hạt nhân có độ dài 100 thường đủ để tạo ra kết quả giải mã trực quan hợp lý. Điều này là thông tin quan trọng trong việc thiết kế kiến trúc mạng.

**5.2. Image Deconvolution CNN (DCNN)**

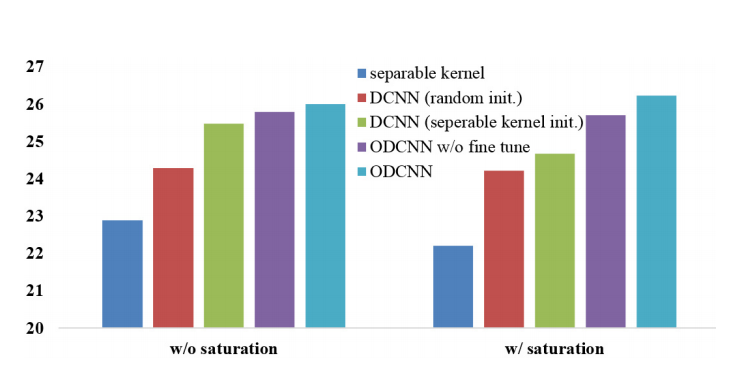
Chúng tôi mô tả mạng nơ ron giải mã tích chập hình ảnh (DCNN) dựa trên phân tách hạt nhân. Mạng này được thể hiện là:



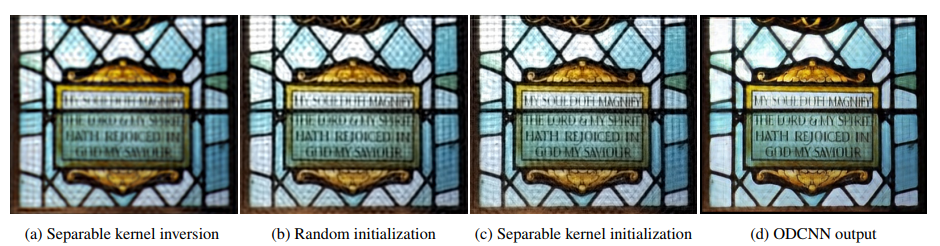
ở đâu  là trọng số ánh xạ lớp đến l thứ một và  là độ lệch giá trị vectơ.  là hàm phi tuyến, có thể là tiếp tuyến sigmoid hoặc hyperbolic. Mạng của chúng tôi chứa hai lớp ẩn tương tự như cài đặt đảo ngược hạt nhân có thể tách rời. Đầu tiên lớp ẩn h1 được tạo bằng cách áp dụng 38 hạt nhân một chiều quy mô lớn có kích thước 121 × 1, theo phân tích trong mục 5.1. Các giá trị 38 và 121 được xác định theo kinh nghiệm, trong đó có thể được thay đổi cho các đầu vào khác nhau. Lớp ẩn thứ hai h2 được tạo bằng cách áp dụng 38 1 × 121 hạt nhân chập cho mỗi 38 bản đồ trong h1. Để tạo kết quả, hạt nhân 1 × 1 × 38 được áp dụng, tương tự như sự kết hợp tuyến tính sử dụng giá trị số ít sj. Kiến trúc có một số lợi thế cho việc giải mã. Đầu tiên, nó lắp ráp đảo ngược hạt nhân có thể tách rời để giải mã và do đó được đảm bảo là tối ưu. Thứ hai, các điều khoản phi tuyến và cấu trúc chiều cao làm cho mạng biểu cảm hơn so với giả ngược truyền thống. Đó là hợp lý mạnh mẽ để ngoại lệ.

**5.3. Training DCNN**

Mạng có thể được đào tạo bằng cách khởi tạo trọng lượng ngẫu nhiên hoặc bằng cách khởi tạo từ đảo ngược kernel tách rời, vì chúng chia sẻ cùng một cấu trúc. Chúng tôi thử nghiệm cả hai chiến lược trên các hình ảnh tự nhiên, tất cả đều bị suy giảm bởi Gaussian phụ gia nhiễu (AWG) và nén JPEG. Những hình ảnh này có hai loại - một loại có màu sắc mạnh bão hòa và không có. Lưu ý bão hòa ảnh hưởng đến nhiều thuật toán giải mã hiện có rất nhiều.



Hình 4: Các PSNR được tạo ra trong các giai đoạn khác nhau của kiến trúc mạng thần kinh tích chập của chúng tôi.



Hình 5: So sánh kết quả trong các giai đoạn khác nhau của CNN giải mã của chúng tôi.

Các PSNR được hiển thị dưới dạng ba thanh đầu tiên trong Hình 4. Chúng tôi thu được các quan sát sau đây.

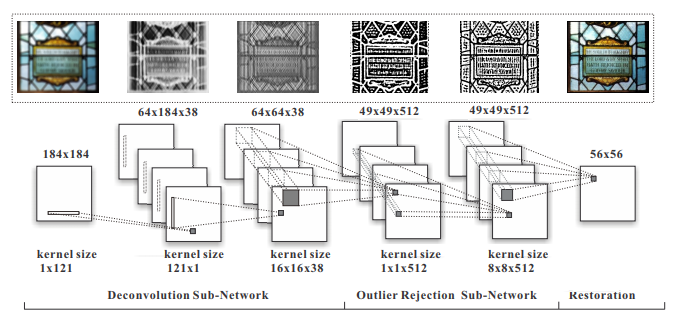
• Mạng được đào tạo có một lợi thế so với việc thực hiện đảo ngược hạt nhân đơn giản, không có vấn đề với khởi tạo ngẫu nhiên hoặc khởi tạo từ giả ngược. Giải thích của chúng tôi là mạng, với ánh xạ chiều cao và phi tuyến, có tính biểu cảm cao hơn so với đảo ngược hạt nhân đơn giản.

• Phương thức khởi tạo đảo ngược nhân tách rời mang lại PSNR cao hơn so với phương thức với khởi tạo ngẫu nhiên, cho thấy rằng các giá trị ban đầu ảnh hưởng đến mạng này và do đó có thể điều chỉnh.

So sánh trực quan được cung cấp trong Hình 5 (a) - (c), trong đó kết quả đảo ngược hạt nhân có thể tách rời, đào tạo với trọng số ngẫu nhiên và đào tạo với khởi tạo đảo ngược nhân tách rời được hiển thị. Kết quả trong (c) rõ ràng chứa các cạnh sắc nét và nhiều chi tiết hơn. Lưu ý rằng DCNN được đào tạo cuối cùng không tương đương với bất kỳ hàm nghịch đảo nào hiện có ngay cả với việc chính quy hóa khác nhau, do liên quan đến ánh xạ chiều cao với phi tuyến. Hiệu suất của CNN giải mã giảm cho các hình ảnh có độ bão hòa màu. Hiện vật trực quan cũng có thể được mang lại do tiếng ồn và nén. Chúng tôi trong phần tiếp theo chuyển sang một cấu trúc sâu hơn để giải quyết những vấn đề còn lại này, bằng cách kết hợp một mô-đun CNN khử nhiễu.

**5.4. Outlier-rejection Deconvolution CNN (ODCNN)**

Mạng hoàn chỉnh của chúng tôi được hình thành như là sự kết hợp của mô đun CNN giải mã với một khử nhiễu CNN [16]. Cấu trúc tổng thể được hiển thị trong Hình 6. Mô-đun CNN khử nhiễu có hai các lớp ẩn với 512 bản đồ tính năng. Hình ảnh đầu vào được tích hợp với 512 nhân kích thước 16 × 16 được đưa vào lớp ẩn. Hai mô-đun mạng được nối với nhau trong hệ thống của chúng tôi bằng cách kết hợp lớp CNN giải mã cuối cùng với đầu vào của CNN khử nhiễu. Điều này được thực hiện bằng cách hợp nhất kernel 1 × 1 × 36 với 512 Hạt nhân 16 × 16 để tạo ra hạt nhân 512 có kích thước 16 × 16 × 36. Lưu ý rằng không có phi tuyến khi kết hợp hai mô-đun. Mặc dù số lượng trọng lượng tăng lên do hợp nhất, nó cho phép thủ tục linh hoạt và đạt được hiệu suất tốt, bằng cách kết hợp thêm tinh chỉnh.



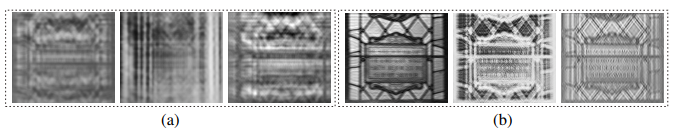
Hình 6: Kiến trúc mạng hoàn chỉnh của chúng tôi để giải mã sâu

**5.5. Training ODCNN**

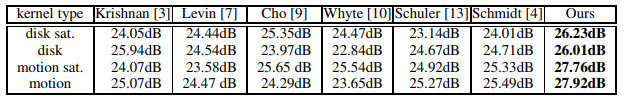
Chúng tôi làm mờ hình ảnh tự nhiên để đào tạo - do đó dễ dàng có được một số lượng lớn dữ liệu. Đặc biệt, chúng tôi sử dụng 2.500 hình ảnh tự nhiên được tải xuống từ Flickr. Hai triệu bản vá được lấy mẫu ngẫu nhiên từ họ. Nối hai mô-đun mạng có thể mô tả quá trình giải mã và tăng cường khả năng ngăn chặn các cấu trúc không mong muốn. Chúng tôi đào tạo các mạng con riêng biệt. Các CNN giải mã được đào tạo bằng cách sử dụng khởi tạo từ đảo ngược có thể tách rời như mô tả trước đây. Đầu ra của CNN giải mã sau đó được lấy làm đầu vào của CNN khử nhiễu. Tinh chỉnh được thực hiện bằng cách cung cấp một trăm nghìn bản vá 184 × 184 vào toàn bộ mạng. Các mẫu đào tạo chứa tất cả các bản vá có thể có các tạo tác nhiễu, bão hòa và nén. Số liệu thống kê về việc thêm CNN khử nhiễu cũng được vẽ trong hình 4. CNN loại bỏ ngoại lệ sau khi phạt tiền điều chỉnh cải thiện hiệu suất tổng thể lên đến 2dB, đặc biệt là đối với những vùng bão hòa.

**6. More Discussions**

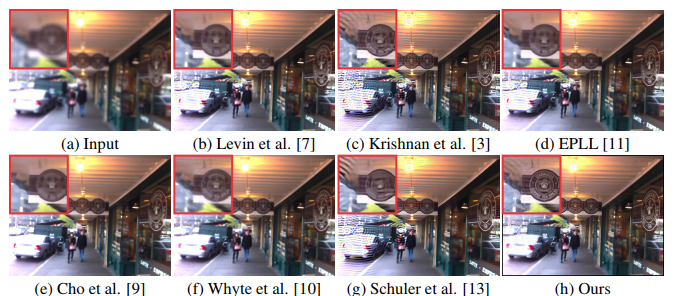
Cách tiếp cận của chúng tôi khác với những cách trước đây theo nhiều cách. Đầu tiên, chúng tôi xác định sự cần thiết của việc sử dụng một hỗ trợ hạt nhân tương đối lớn cho mạng nơ ron tích chập để đối phó với quá trình giải mã. Tránh mở rộng kích thước trọng lượng nhanh chóng, chúng tôi ủng hộ việc sử dụng hạt nhân 1D. Thứ hai, chúng tôi đề xuất một giám sát đào tạo trước trên mạng con tương ứng với việc diễn giải lại quá trình giải mã Wiener. Thứ ba, chúng tôi áp dụng giải mã truyền thống để khởi tạo mạng, trong đó người giải quyết thế hệ có thể hướng dẫn học tập mạng lưới thần kinh và cải thiện đáng kể hiệu suất. Hình 6 cho thấy một kiến ​​trúc mạng nơ ron tích chập mới có khả năng xử lý sự phân giải. Tuy nhiên, nếu không có sự hiểu biết tốt về chức năng của từng mạng con và thực hiện đào tạo trước được giám sát, rất khó để làm cho mạng hoạt động rất tốt. Đào tạo toàn bộ mạng với khởi tạo ngẫu nhiên ít được ưu tiên hơn vì thuật toán đào tạo dừng lại giữa chừng mà không giảm thêm năng lượng. Các kết quả tương ứng mờ tương tự như hình ảnh đầu vào. Để hiểu nó, chúng tôi hình dung các kết quả trung gian từ mạng con CNN giải mã, trong đó tạo ra 38 bản đồ trung gian. Các kết quả được hiển thị trong Hình 7, trong đó (a) là lựa chọn ba kết quả thu được từ đào tạo khởi tạo ngẫu nhiên và (b) là kết quả tương ứng các nút từ quá trình khởi tạo tốt hơn của chúng tôi. Các bản đồ trong (a) trông giống như phần tần số cao của đầu vào mờ, cho biết khởi tạo ngẫu nhiên có khả năng tạo các bộ lọc thông cao. Nếu không có bắt đầu các giá trị, cơ hội của nó là rất nhỏ để đạt được các bản đồ thành phần được hiển thị trong (b) nơi sắc nét hơn các cạnh hiện diện, hoàn toàn có thể sử dụng để tiếp tục loại bỏ nhiễu và tạo tác. Zeiler et al. [27] cho thấy giải mã thường xuyên thưa thớt có thể được sử dụng để trích xuất hữu ích đại diện trung cấp trong mạng giải mã của họ. CNN giải mã của chúng tôi có thể được sử dụng để gần đúng cấu trúc này, thống nhất quá trình trong một mạng lưới thần kinh tích chập sâu hơn.



Hình 7: So sánh các kết quả trung gian từ CNN giải mã. (a) Bản đồ ngẫu nhiên khởi tạo. (b) Thêm bản đồ thông tin với sơ đồ khởi tạo của chúng tôi.



Bảng 1: So sánh định lượng trên bộ ảnh đánh giá.



Hình 8: So sánh trực quan các kết quả giải mã.

**7. Experiments and Conclusion**

Chúng tôi đã trình bày một số kết quả giải mã. Ở đây chúng tôi hiển thị đánh giá định lượng của phương pháp của chúng tôi chống lại các phương pháp tiên tiến, bao gồm giải mã trước thưa thớt [7], phương pháp trước hyperLaplacian [3], EM đa dạng cho các ngoại lệ [9], phương pháp nhận biết bão hòa [10], học tập phương pháp dựa trên [13] và phương pháp phân biệt đối xử [4]. Chúng tôi so sánh hiệu suất bằng cả hai đĩa và hạt nhân chuyển động. Các PSNR trung bình được liệt kê trong Bảng 1. Hình 8 cho thấy một so sánh trực quan. Phương pháp của chúng tôi đạt được kết quả tốt về số lượng và trực quan. Việc thực hiện, cũng như dữ liệu, có sẵn tại trang web của dự án.

Để kết thúc bài báo này, chúng tôi đã đề xuất một cấu trúc mạng tích chập sâu mới cho nhiệm vụ giải mã hình ảnh đầy thách thức. Đóng góp chính của chúng tôi là để cho các chương trình giải mã truyền thống hướng dẫn các mạng thần kinh và giải mã gần đúng bằng một loạt các bước tích chập. Hệ thống của chúng tôi novelly sử dụng hai mô-đun tương ứng để loại bỏ giải mã và tạo tác. Trong khi mạng rất khó để đào tạo nói chung, chúng tôi áp dụng hai bước đào tạo trước được giám sát để khởi tạo các mạng con. Kết quả giải mã chất lượng cao mang lại hiệu quả của phương pháp này.