# **Bài 1: GAN Prior based Null-Space Learning for Consistent Super-Resolution**

Link paper: https://arxiv.org/pdf/2211.13524v1.pdf

Link git : https://github.com/wyhuai/rnd

Star : 36\*

Năm 2022

Benmarks:

Tóm tắt:

A screenshot of a computer

Description automatically generated with low confidence

Nó sử dụng mô hình GAN để học không gian con trống dựa trên tiên đề trong việc tái tạo hình ảnh phân giải cao từ hình ảnh phân giải thấp.

Để giải quyết vấn đề siêu phân giải liên tục trong hình ảnh, bài báo đề xuất sử dụng tiên đề rằng một ảnh phân giải cao có thể được tái tạo bằng cách thêm vào một ảnh phân giải thấp và một phần không gian con trống. Trong đó, ảnh phân giải thấp là đầu vào và không gian con trống là phần còn lại.

Để học không gian con trống này, bài báo đề xuất sử dụng mô hình GAN, bao gồm một mạng sinh (Generator) và một mạng phân biệt (Discriminator). Mạng sinh được đào tạo để tạo ra các ảnh phân giải cao từ không gian con trống và ảnh phân giải thấp đầu vào, trong khi mạng phân biệt được đào tạo để phân biệt giữa các ảnh phân giải cao được tạo ra bởi mạng sinh và các ảnh phân giải cao thực tế.

Để đảm bảo tính ổn định và đồng nhất trong việc tái tạo hình ảnh, bài báo đề xuất sử dụng một số kỹ thuật như:

Kiểm soát tỷ lệ cân bằng giữa độ lớn của ảnh phân giải cao và không gian con trống trong mô hình GAN.

Sử dụng hàm mất mát kết hợp giữa hàm mất mát của mạng sinh và hàm mất mát của mạng phân biệt.

Áp dụng kỹ thuật tái cấu trúc định hướng (guided reconstruction) để hướng dẫn mạng sinh tạo ra các ảnh phân giải cao có chất lượng tốt hơn.

Kết quả:

Phương pháp được kiểm tra trên các bộ dữ liệu hình ảnh khác nhau và so sánh với các phương pháp khác trong việc tái tạo hình ảnh phân giải cao. Kết quả cho thấy rằng phương pháp đề xuất đạt được kết quả tốt hơn so với các phương pháp khác và cho thấy tính ổn định và đồng nhất trong việc tái tạo hình ảnh.

Graphical user interface, chart

Description automatically generated

Table

Description automatically generated

8 x SR

Nhược điểm

nếu sự suy giảm mà LR gặp phải làm hư hại nội dung tần số thấp, PD sẽ truyền các nội dung tần số thấp bị hư hại trực tiếp vào kết quả SR (ảnh siêu phân giải). Một giải pháp đơn giản là tiền xử lý LR bị hư hại thành LR sạch

# **Bài 2: Implicit Diffusion Models for Continuous Super-Resolution**

Link paper: https://arxiv.org/pdf/2303.16491v1.pdf

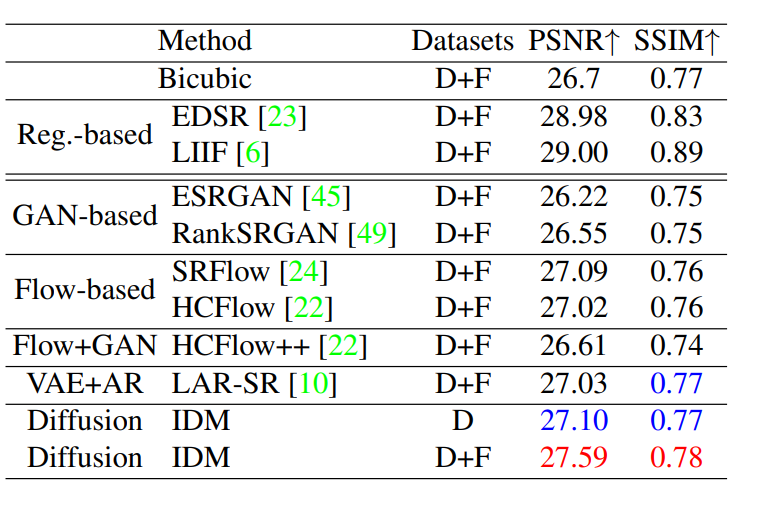
Link git: https://github.com/ree1s/idm

Năm 2023

Star: 26\*

Benmarks:

DIV2K : D  
Flicker2K : F



* CelebA-HQGraphical user interface, application

  Description automatically generated

Tóm tắt:

IDM kết hợp những ưu điểm của các mô hình diffusion và Implicit neural representations (INRs) trong một end-to-end framework, do đó đạt được ảnh siêu phân giải có chất lượng như ảnh thực.

Diagram

Description automatically generated

Các Mô hình Xác suất Khuếch tán (DM) đã cho thấy những kết quả tiên tiến nhất trong việc tổng hợp hình ảnh và giọng nói. IDM kết hợp ưu điểm của các mô hình khuếch tán và các biểu diễn nơ-ron ngầm trong một khung thực tế từ đầu đến cuối để thu được hình ảnh SR chân thực với độ phân giải liên tục, giải quyết các vấn đề như tạo tác không tự nhiên, tỷ lệ phóng đại cố định và đường dẫn hai giai đoạn phức tạp trong các mô hình hiện có.

# **Bài 3: Image Super-Resolution via Iterative Refinement**

Link paper: <https://arxiv.org/pdf/2104.07636v2.pdf>

Link git: https://github.com/Janspiry/Image-Super-Resolution-via-Iterative-Refinement

Năm 2021

Star: 2,720\*

Benmarks:

# **Bài 4: GLEAN: Generative Latent Bank for Image Super-Resolution and Beyond**

Link paper: https://arxiv.org/pdf/2303.16491v1.pdf

Link git: https: https://github.com/open-mmlab/mmediting

Năm 2022

Star: 4,255\*

Benmarks:

Text

Description automatically generated

Tóm tắt:

* Sử dụng 3 phần chính:

+ encoder: tạo ra các vector đặc trưng để đưa vào latent bank

+ Generative Latent Bank:

* Biến đổi những mạng trước đó như StyleGAN, BigGAN

1. thay vì lấy một vectơ tiềm ẩn duy nhất làm đầu vào, mỗi khối của trình tạo có một vectơ tiềm ẩn khác nhau để cải thiện tính biểu cảm
2. Để cho phép điều hòa trên các tính năng bổ sung từ bộ mã hóa, chúng tôi sử dụng tích chập bổ sung trong mỗi khối
3. Thay vì tạo trực tiếp các đầu ra từ trình tạo, chúng tôi xuất ra các tính năng và chuyển chúng cho bộ giải mã để hợp nhất các tính năng tốt hơn từ ngân hàng tiềm ẩn và bộ mã hóa

+ Decoder: sử dụng bộ giải mã bổ sung với sự kết hợp tiến bộ để tích hợp các đặc trưng từ bộ mã hóa và bộ nhớ ẩn để tạo ra ảnh đầu ra. Nó lấy các đặc trưng RRDBNet làm đầu vào và tiến hành kết hợp các đặc trưng này với các đặc trưng đa phân giải từ bộ nhớ ẩn. Sau mỗi lớp tích chập là một lớp pixelshuffle. Với sự kết nối trực tiếp giữa bộ mã hóa và bộ giải mã, thông tin được bắt được bởi bộ mã hóa có thể được tăng cường, và do đó bộ nhớ ẩn có thể tập trung tốt hơn vào việc tạo ra các chi tiết và cấu trúc của ảnh đầu ra.

* Không yêu cầu tối ưu hóa lặp lại trong quá trình đạo hàm

LightGLEAN, đạt được hiệu suất tương đương trong khi chỉ có 21% tham số và 35% FLOPs.

"latent bank" thường được sử dụng để chỉ một tập hợp các vector đặc trưng (feature vectors) được học từ dữ liệu đầu vào bởi mô hình mạng nơ-ron sâu.

Cụ thể, mỗi lần mô hình được huấn luyện trên một lượng lớn dữ liệu, các layer ẩn (hidden layers) của mô hình sẽ học được một số vector đặc trưng riêng biệt, tương ứng với các tính năng quan trọng trong dữ liệu đó. Những vector này được gọi là "latent vectors" hoặc "embedding vectors".

Chart, box and whisker chart

Description automatically generated

GLEAN

A picture containing chart

Description automatically generated

LightGLEAN

# **So sánh mạng GAN và Diffusion**

|  |  |  |
| --- | --- | --- |
|  | GAN | Diffusion |
| Ưu điểm | * Mạng GAN có khả năng tạo ra những ảnh chất lượng cao và có tính thẩm mỹ tốt. * Có thể áp dụng mạng GAN để tạo ra ảnh có độ phân giải cao từ ảnh gốc có độ phân giải thấp mà không cần thông tin chi tiết về ảnh ban đầu. * Để huấn luyện mạng GAN cần ít dữ liệu hơn so với một số phương pháp khác. | * Phương pháp diffusion có khả năng tạo ra ảnh có chất lượng cao, giữ được nhiều chi tiết và độ phân giải cao hơn so với nhiều phương pháp khác. * Diffusion có khả năng kiểm soát độ phân giải của ảnh kết quả. |
| Nhược điểm | * Mạng GAN có xu hướng tạo ra những ảnh giống với tập dữ liệu huấn luyện, điều này có thể dẫn đến overfitting và ảnh kết quả không đa dạng. * Việc huấn luyện mạng GAN khó khăn và tốn nhiều thời gian. | * Diffusion yêu cầu tính toán phức tạp hơn so với một số phương pháp khác. * Để đạt được kết quả tốt, diffusion cần phải có một lượng lớn dữ liệu huấn luyện. |

* Cả mạng GAN và diffusion đều có ưu điểm và nhược điểm của riêng mình. Tuy nhiên, tùy thuộc vào yêu cầu và điều kiện cụ thể của bài toán siêu phân giải ảnh mà một phương pháp có thể phù hợp hơn phương pháp khác.