

NGHIÊN CỨU VÀ PHÁT TRIỂN MÔ HÌNH GAN TRONG TÁI TẠO ẢNH Y TẾ

Sinh viên thực hiện:
Phạm Huỳnh Quý An

Ngày 14 tháng 1 năm 2026

Mục lục

1	Cơ sở lý thuyết	2
1.1	Bài toán Super-Resolution	2
1.2	Hàm mất mát (Loss Functions)	2
1.2.1	Content Loss	2
1.2.2	Adversarial Loss	2
1.2.3	Tổng hợp hàm mục tiêu	2
2	Thực nghiệm và Kết quả	3
2.1	Độ đo đánh giá (Metrics)	3
2.2	Kết quả so sánh	3
3	Kết luận	4

Chương 1

Cơ sở lý thuyết

1.1 Bài toán Super-Resolution

Mục tiêu của bài toán là khôi phục ảnh độ phân giải cao I^{HR} từ hình ảnh độ phân giải thấp I^{LR} . Mối quan hệ giữa chúng được mô hình hóa như sau:

$$I^{LR} = \mathcal{D}(I^{HR}; \delta) \quad (1.1)$$

Trong đó \mathcal{D} là hàm giảm chất lượng (degradation function), và δ là tham số nhiễu.

1.2 Hàm mất mát (Loss Functions)

Để huấn luyện mạng Generator G, chúng ta kết hợp nhiều thành phần (Perceptual Loss).

1.2.1 Content Loss

Thay vì tính toán trên từng pixel, ta tính khoảng cách giữa các đặc trưng (feature maps) được trích xuất bởi mạng VGG19:

$$L_{content} = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I^{HR})_{x,y} - \phi_{x,y}(G(I^{LR}))_{x,y})^2 \quad (1.2)$$

1.2.2 Adversarial Loss

Thành phần này giúp ảnh sinh ra trông "thật" hơn, đánh lừa được Discriminator:

$$L_{adv} = \sum_{n=1}^N -\log D(G(I^{LR})) \quad (1.3)$$

1.2.3 Tổng hợp hàm mục tiêu

Hàm loss cuối cùng là tổng có trọng số:

$$L_{total} = \lambda_1 L_{content} + \lambda_2 L_{adv} + \lambda_3 L_{pixel} \quad (1.4)$$

Chương 2

Thực nghiệm và Kết quả

2.1 Độ đo đánh giá (Metrics)

Chúng ta sử dụng chỉ số PSNR (Peak Signal-to-Noise Ratio) để đánh giá chất lượng ảnh tái tạo. Công thức như sau:

$$\text{PSNR} = 10 \cdot \log_{10}\left(\frac{MAX_I^2}{MSE}\right) \quad (2.1)$$

Trong đó MAX_I là giá trị pixel lớn nhất của ảnh (thường là 255).

2.2 Kết quả so sánh

Dưới đây là bảng so sánh hiệu năng giữa mô hình đề xuất (Our GAN) với các phương pháp truyền thống như Bicubic và SRGAN [1]

Bảng 2.1: So sánh chỉ số PSNR và SSIM trên bộ dữ liệu X-ray

Phương pháp	PSNR (dB) ↑	SSIM ↑	Thời gian (s)
Bicubic Interpolation	24.50	0.72	0.01
SRGAN (2017)	28.15	0.85	0.15
ESRGAN (2018)	29.40	0.88	0.22
Our Method (Ours)	30.12	0.91	0.18

Như ta thấy ở Bảng 2.1, mô hình của chúng tôi đạt chỉ số PSNR cao nhất là 30.12 dB, vượt trội hơn so với SRGAN [1]

Chương 3

Kết luận

Nghiên cứu này chứng minh tính hiệu quả của việc kết hợp Content Loss và Adversarial Loss

Tài liệu tham khảo

- [1] R. Gupta, A. Sharma, and A. Kumar, “Super-resolution using gans for medical imaging,” *Procedia Computer Science*, vol. 173, pp. 28–35, 2020.