

BÁO CÁO SAU PHẢN BIỆN

TÔ MÀU ẢNH ĐỘ XÁM DỰA TRÊN MÔ HÌNH KHUẾCH TÁN

DIFFUSION-BASED GRAYSCALE IMAGE COLORIZATION

1 THÔNG TIN CHUNG

Người hướng dẫn:

- PGS. TS. Lý Quốc Ngọc (Khoa Công nghệ thông tin)
- ThS. Đỗ Thị Thanh Hà (Khoa Công nghệ thông tin)

Người phản biện:

- ThS. Võ Hoài Việt (Khoa Công nghệ thông tin)

Nhóm sinh viên thực hiện:

1. Phạm Thái Huy (MSSV: 21120081)
2. Tiêu Ân Tuấn (MSSV: 21120161)

Loại đề tài: Nghiên cứu

Thời gian thực hiện: Từ 01/2025 đến 07/2025

2 NỘI DUNG BÁO CÁO

2.1 Tổng kết nội dung buổi phản biện

Ở phần này, nhóm xin phép tổng kết các nhận xét của giáo viên phản biện, và các nhiệm vụ được yêu cầu sau buổi phản biện.

2.1.1 Nhận xét chung

Các khuyết điểm trong việc thực hiện khóa luận của nhóm sinh viên, bao gồm:

- Phát biểu bài toán không đúng
 - Đề tài nghiên cứu là "Tô màu ảnh độ xám", nhưng mô hình lại nhận đầu vào là kênh L (của không gian màu Lab).
 - Đồng thời, sinh viên chưa trình bày được lý do chọn kênh màu này thay cho ảnh độ xám, cả về mặt lý thuyết lẫn thực nghiệm.
- Chưa giải quyết toàn bộ, triệt để các hạn chế hiện tại của các nghiên cứu về tô màu ảnh độ xám như đã trình bày trong báo cáo.
- Chưa hiểu rõ và trình bày mạch lạc kiến trúc của mô hình, cách hoạt động của hệ thống.
- Dữ liệu không được mô tả rõ ràng
 - Dữ liệu huấn luyện và đánh giá được sử dụng có đúng với triết lý học máy hay không ?
 - Miền dữ liệu được sử dụng và miền dữ liệu của bài toán có tương thích với nhau hay không ? Cần nêu rõ giới hạn miền dữ liệu.
- Kết quả nghiên cứu không đạt kết quả vượt trội hơn các nghiên cứu SOTA trước đó.
 - Mục đích và đóng góp nghiên cứu của nhóm thật sự là gì ?

2.1.2 Các yêu cầu sau phản biện

Với những nhận xét trên, thầy yêu cầu nhóm tiến hành thực nghiệm cho các vấn đề cốt lõi của nghiên cứu, bao gồm:

- Trình bày lý do sử dụng kênh L thay cho ảnh độ xám
 - Giải thích ý nghĩa trực quan giữa kênh L và ảnh độ xám.
 - Đánh giá sai số đầu vào của mô hình: Đánh giá mối tương quan số học giữa kênh L và ảnh độ xám.
- Đánh giá sai số các độ đo (FID, Colorfulness) trong thực nghiệm đối với hai đầu vào là kênh L và ảnh độ xám.

Hai yêu cầu trên nhằm trả lời câu hỏi lớn của giáo viên phản biện:

Kênh L có tương đương và có thể thay thế cho ảnh độ xám đối với bài toán được nêu hay không ?

Ngoài ra, nhóm còn được yêu cầu viết lại báo cáo hoàn chỉnh bao gồm chi tiết cách hoạt động của hệ thống và mô tả dữ liệu được sử dụng.

2.2 Báo cáo sau phản biện của nhóm

Với những nhận xét và yêu cầu từ giáo viên phản biện, nhóm tiến hành nghiên cứu, thực nghiệm và báo cáo lại **một cách ngắn gọn** về khóa luận, bao gồm các mục chính sau:

1. Trình bày sự khác biệt giữa kênh L (trong không gian màu Lab) và ảnh độ xám.
2. Đánh giá sai số đầu vào của bài toán.
3. Đánh giá sai số của kết quả các độ đo trong các thực nghiệm.
4. Phát biểu lại bài toán.

5. Mô tả tường minh cách hoạt động của hệ thống.
6. Mô tả dữ liệu được sử dụng.

2.2.1 Sự khác biệt giữa kênh L và ảnh độ xám

1. Lý thuyết về kênh L và ảnh độ xám
2. Cách áp dụng kênh L và ảnh độ xám trong nghiên cứu của nhóm

2.2.2 Đánh giá sai số đầu vào

Chứng minh mối quan hệ giữa kênh L và ảnh độ xám cả về mặt trực quan hay số học một cách tổng quát rất phức tạp, và nhóm cũng không đủ khả năng để thực hiện. Thay vào đó, nhóm tiến hành đánh giá mối quan hệ của chúng trong ngữ cảnh riêng biệt của nghiên cứu này. Cụ thể, nhóm sẽ đánh giá về mặt số học của hai đối tượng này.

Theo lý thuyết, kênh L có giá trị từ 0 đến 100, còn các điểm ảnh trong ảnh độ xám có giá trị từ 0 đến 255. Để so sánh công bằng, các giá trị của hai đối tượng sẽ được chuẩn hóa về khoảng từ 0 đến 1, theo kỹ thuật chuẩn hóa min-max. Sau đó, nhóm khảo sát mối quan hệ giữa kênh L và ảnh độ xám thông qua các giá trị được mô tả trong bảng bên dưới.

Các độ đo	Ý nghĩa
Các giá trị thống kê phổ biến: trung bình (mean), trung vị (median), độ lệch chuẩn (std)	Đo sự tương đồng tổng thể và độ phân tán dữ liệu trong cả ảnh độ xám và kênh L.
Sai số bình phương trung bình (MSE) Sai số căn bậc hai bình phương trung bình (RMSE)	Định lượng mức độ khác biệt giữa hai đối tượng. Giá trị càng thấp cho thấy sự tương đồng càng cao và ngược lại.
Tỉ lệ đỉnh tín hiệu trên nhiễu (PSNR)	Là một tiêu chuẩn để đánh giá chất lượng tái tạo. Được sử dụng với ngữ cảnh giả sử kênh L là ảnh tái tạo của ảnh độ xám và ngược lại. Giá trị càng cao cho thấy chúng càng giống nhau, có thể xem như tương đồng.
Hệ số tương quan Pearson	Đo mức độ tương quan tuyến tính giữa ảnh độ xám và kênh L. Miền giá trị từ -1 (tương quan âm hoàn hảo) đến 1 (tương quan dương hoàn hảo), trong đó giá trị 0 đại diện cho sự không tương quan. Giá trị càng cao cho thấy sự tương đồng tuyến tính càng cao và ngược lại.
Hệ Số Tương Quan Spearman	Đo sự tương quan đơn điệu (không nhất thiết tuyến tính), dựa trên thứ hạng các giá trị. Miền giá trị từ -1 (đơn điệu âm hoàn hảo) đến 1 (đơn điệu dương hoàn hảo), trong đó giá trị 0 đại diện cho sự không tương quan. Giá trị càng cao cho thấy kênh L và ảnh độ xám có xu hướng thay đổi đồng bộ. Hệ số này bổ sung cho hệ số tương quan Pearson với thông tin về sự đơn điệu.
SSIM	Đo sự tương đồng cấu trúc, ánh sáng và độ tương phản. Miền giá trị từ -1 (khác biệt hoàn toàn) đến 1 (giống nhau hoàn toàn), trong đó giá trị 0 đại diện cho sự không tương quan.
Thông tin chung (Mutual Information)	Đo lượng thông tin chung giữa ảnh độ xám và kênh L, tính trên dữ liệu đã được rời rạc hóa. Giá trị nhỏ nhất là 0 (tức là không có thông tin chung) và không có giới hạn trên. Giá trị càng cao cho thấy hai đối tượng có càng nhiều thông tin chung, nhưng trong thực tế, các giá trị này thường nhỏ vì phụ thuộc vào entropy của dữ liệu (Nhóm sẽ không trình bày chi tiết về khía cạnh này).

Bảng 1: Mô tả các độ đo để đánh giá mối quan hệ của ảnh độ xám và kênh L.

Nhóm tiến hành đo trên từng tấm ảnh và tổng hợp tên toàn bộ tập dữ liệu, chi tiết ở bảng bên dưới.

Tập dữ liệu	Thống kê	Các độ đo												
		g_mean	g_median	g_std	L_mean	L_median	L_std	MSE	RMSE	PSNR	Pearson	Spearman	Mual_Info	SSIM
val5k	mean	0.35046	0.34239	0.27801	0.37005	0.37067	0.28941	0.00092	0.02808	31.60181	0.99822	0.99872	3.26391	0.99058
	std	0.11773	0.1855	0.05772	0.12053	0.19575	0.05581	0.00101	0.01141	3.06679	0.00315	0.00446	0.47562	0.00826
	min	0.01507	0.0	0.07997	0.01416	0.0	0.08256	2e-05	0.00415	17.06456	0.87476	0.8507	0.75083	0.88107
	max	0.95069	1.0	0.47148	0.95391	1.0	0.47192	0.01966	0.14021	47.63545	0.99996	1.0	5.28947	0.99972
ctest	mean	0.34778	0.33988	0.27746	0.36736	0.36811	0.28891	0.00091	0.0281	31.57133	0.99826	0.99879	3.25005	0.99047
	std	0.11595	0.18312	0.05734	0.11891	0.19366	0.0555	0.00097	0.01109	2.99609	0.00251	0.00351	0.45413	0.00841
	min	0.01741	0.0	0.04721	0.01773	0.0	0.04723	4e-05	0.00598	17.18224	0.92739	0.84725	0.44393	0.81927
	max	0.95596	1.0	0.4738	0.96013	1.0	0.4759	0.01913	0.13832	44.46363	0.99997	1.0	5.17154	0.99949
COCO-Stuff	mean	0.32594	0.30085	0.28618	0.34343	0.32599	0.29759	0.00075	0.02609	32.06156	0.99852	0.99923	3.25002	0.99164
	std	0.09611	0.16268	0.0525	0.09899	0.17562	0.05069	0.00061	0.00845	2.56665	0.00199	0.00217	0.42089	0.00571
	min	0.01145	0.0	0.02502	0.01134	0.0	0.02346	1e-05	0.0037	19.35504	0.94979	0.92472	0.49256	0.89094
	max	0.84333	0.96471	0.46333	0.85314	0.96863	0.4654	0.0116	0.10771	48.6306	0.99998	1.0	5.29685	0.99903

Bảng 2: Bảng sai số đầu vào trên các độ đo được khảo sát.

Tập dữ liệu	val5k		ctest		COCO-Stuff	
Loại đầu vào Độ đo	FID	Colorfulness	FID	Colorfulness	FID	Colorfulness
Kênh L	10.4125	44.2871	6.8341	44.7854	10.4632	45.2589
Ảnh độ xám (g)	10.5105	43.5132	6.9232	43.8667	10.5635	44.4984
Phần trăm thay đổi ($\frac{g-L}{L} * 100$)	0.9412	-1.7475	1.3038	-2.0513	0.9586	-1.6803

Bảng 3: Bảng so sánh kết quả độ đo với hai loại đầu vào.

2.2.3 Đánh giá sai số độ đo trong thực nghiệm với hai đầu vào khác nhau

2.2.4 Phát biểu bài toán

2.2.5 Cách hoạt động của hệ thống

2.2.6 Mô tả dữ liệu được sử dụng

3 LỜI KẾT