**THÔNG TIN CHUNG CỦA NHÓM**

* Link github nhóm:

[lekhanhhoang/influence-of-color](https://github.com/lekhanhhoang/influence-of-color)

|  |  |
| --- | --- |
| * Họ và tên: Phan Hoàng Vũ * MSSV: 3123410436 | * Lớp: DCT1235 * Tự đánh giá (điểm tổng kết môn): 9.5/10 * Số buổi vắng: 0 * Link github: [Phan Hoang Vu](https://phanvu1.github.io/phanvu.github.io/). * Mô tả công việc và đóng góp cá nhân cho kết quả nhóm:   + Đóng góp ý tưởng cho đề tài.   + Làm slide.   + Viết mục tiêu.   + Hỗ trợ viết đề cương. |
| * Họ và tên: Nguyễn Tiến Trung * MSSV: 3123410396 | * Lớp: DCT1233 * Tự đánh giá (điểm tổng kết môn): 9.0/10 * Số buổi vắng: 0 * Link github: [Nguyen Tien Trung](https://github.com/uncletientrung) * Mô tả công việc và đóng góp cá nhân cho kết quả nhóm:   + Đóng góp ý tưởng cho đề tài.   + Tìm hiểu code   + Làm slide.   + Viết mục tiêu.   + Hỗ trợ viết đề cương. |
| * Họ và tên: Nguyễn Minh Thuận * MSSV: 3123410365 | * Lớp: DCT1235 * Tự đánh giá (điểm tổng kết môn): 9.5/10 * Số buổi vắng: 0 * Link github: [Nguyễn Minh Thuận](https://github.com/thuanstealer) * Mô tả công việc và đóng góp cá nhân cho kết quả nhóm:   + Đóng góp ý tưởng cho đề tài.   + Làm slide.   + Viết mục tiêu.   + Hỗ trợ viết đề cương. |
| * Họ và tên: Lê Khánh Hoàng * MSSV: 3122410125 | * Lớp: DCT1226 * Tự đánh giá (điểm tổng kết môn): 9.5/10 * Số buổi vắng: 0 * Link github: [Lê Khánh Hoàng](https://github.com/lekhanhhoang) * Mô tả công việc và đóng góp cá nhân cho kết quả nhóm:   + Đóng góp ý tưởng cho đề tài.   + Soạn nội dung và viết toàn bộ báo cáo đề tài nghiên cứu. |

**ĐỀ CƯƠNG NGHIÊN CỨU**

|  |
| --- |
| **TÊN ĐỀ TÀI**  **TÔ MÀU ẢNH TRẮNG ĐEN DỰA TRÊN MÔ HÌNH HỌC SÂU** |
| **TÊN ĐỀ TÀI BẰNG TIẾNG ANH**  **COLORIZING GRAYSCALE IMAGES USING DEEP LEARNING MODELS** |
| **TÓM TẮT**  Trong thời đại Cách mạng Công nghiệp 4.0, trí tuệ nhân tạo (AI) và học sâu (Deep Learning) đang trở thành động lực cốt lõi thúc đẩy sự phát triển của các ứng dụng công nghệ, đặc biệt trong lĩnh vực xử lý hình ảnh. Các mạng nơ-ron tích chập (CNN) đã chứng minh hiệu quả vượt trội trong việc trích xuất đặc trưng và xử lý dữ liệu thị giác, mở ra tiềm năng lớn cho các bài toán như tô màu ảnh tự động (image colorization) . Tô màu ảnh, với mục tiêu tái tạo màu sắc từ hình ảnh xám (grayscale), không chỉ mang ý nghĩa trong lĩnh vực đồ họa số mà còn đóng vai trò quan trọng trong các ứng dụng thực tiễn như phục hồi ảnh lịch sử, tăng cường dữ liệu thị giác, và cải thiện hiệu suất các tác vụ thị giác máy tính.  Với 1 bức ảnh trắng đen làm đầu vào, bài báo này sẽ giải quyết vấn đề tưởng tượng ra bản màu hợp lý của bức đó. Đây là một bài toán thiếu dữ kiện, vì vậy các phương pháp trước đây thường dựa vào các tương tác của người dùng hoặc cho ra các kết quả màu bị nhạt. Chúng tôi để xuất phương pháp tự động hoàn toàn, có thể tạo ra các màu sống động và chân thực. Sử dụng kỹ thuật cân bằng lại các lớp trong quá trình huấn luyện để tăng sự đa dạng về màu sắc trong kết quả. Hệ thống của chúng tôi được triển khai như một quá trình lan truyền xuôi (feed-forward) trong mạng nơ-ron tích chập (CNN) ở giai đoạn kiểm tra, và được huấn luyện trên hơn 1 triệu ảnh màu. Chúng tôi đánh giá thuật toán bằng cách sử dụng một " bài kiểm tra Turing tô màu". Yêu cầu người tham gia chọn giữa bức ảnh được tạo ra và bức ảnh gốc Phương pháp của chúng tôi đánh lừa con người trong 32% số lần thử, cao hơn so với các phương pháp trước. |
| **GIỚI THIỆU**  Tô màu bức ảnh là việc gán vào mỗi pixel của bức ảnh trắng đen một màu nhất định. Để tô màu cho bức ảnh đen trắng thành những bức ảnh sống động và và đẹp thì đây là một bài toán khó vì phần lớn các thông đã bị mất. Trước thời kì Deep Learning thì có các cách tiếp cận thông thường là con người cần phải can thiệp vào việc tô màu như **người dùng cần phải vẽ vài nét màu lên ảnh trắng đen [4]**, **Chuyển màu từ ảnh màu sang ảnh đen trắng bằng cách so sánh đặc trưng giữa các vùng ảnh [5]**. Trong bối cảnh học sâu (Deep Learning) đang phát triển mạnh mẽ, chúng tôi đề xuất một phương pháp mới trong lĩnh vực thị giác máy tính, tập trung vào bài toán tô màu ảnh trắng đen. **Cách tiếp cận của chúng tôi sử dụng mô hình học sâu để tự động tạo ra các phiên bản màu sắc sống động và chân thực từ ảnh xám**, khai thác sức mạnh của mạng nơ-ron tích chập (CNN) và dữ liệu quy mô lớn nhằm đạt được kết quả vượt trội so với các phương pháp truyền thống.  Trong phạm vi bài báo cáo này có mục tiêu là không phải là khôi phục chính xác màu gốc thực tế, mà là tạo ra một phiên bản màu hợp lý, có thể khiến con người tin nó là thật. Trong không gian màu CIE Lab, hệ thống chúng tôi có nhiệm vụ sử dụng kênh L ( độ sáng) làm đầu vào, và dự đoán 2 kênh màu a và b tương ứng để tái tạo màu.    Vì dự đoán màu là một bài toán mang tính đa mô thức - nhiều đối tượng nên có thể có nhiều cách tô màu hợp lý khác nhau. Ví dụ như 1 quả táo có thể là màu đỏ, vàng, xanh lá và có khi là cam. Và để làm được điều này chúng tôi phát triển mô hình dự đoán một phối các màu khả thi cho mỗi điểm ảnh rồi lấy giá trị trình của chúng. Kết quả là là các ảnh được tô màu sống động và chân thực hơn so với phương pháp trước đó.  Trong phương pháp của Zhang, thay vì coi bài toán dự đoán màu là một bài toán hồi quy đơn giản (như sử dụng L2 loss), ông đã chuyển nó thành một bài toán phân loại, bằng cách lượng hóa không gian màu ab thành 313 lớp rời rạc. Mỗi điểm ảnh được gán một lớp màu tương ứng, và mô hình sẽ dự đoán xác suất cho từng lớp này. Để huấn luyện mô hình, hàm mất mát cross-entropy được sử dụng nhằm đo sự khác biệt giữa phân phối dự đoán và nhãn màu thật. Cách tiếp cận này giúp mô hình linh hoạt hơn trong việc xử lý tính đa mô thức của màu sắc — khi một điểm ảnh có thể có nhiều màu hợp lý khác nhau.    Ứng dụng chủ yếu của việc tô màu ảnh trắng đen là khôi phục cho các bức ảnh thời xưa hoặc các bộ phim tài liệu cũ |
| **MỤC TIÊU**  Sau khi hoàn thành, đề tài sẽ đáp ứng được các yêu cầu sau:   * Phát triển thành công một thuật toán tô màu ảnh trắng đen tự động sử dụng mạng nơ-ron tích chập (CNN), có khả năng tạo ra các hình ảnh màu sắc sống động và chân thực từ đầu vào ảnh xám, được thử nghiệm trên các bộ dữ liệu lớn như ImageNet trong các tác vụ thị giác máy tính (Image Colorization). * Cải thiện chất lượng màu sắc và tính đa dạng của kết quả: áp dụng kỹ thuật phân loại màu sắc với hàm mất mát được cân bằng lại (class-rebalancing) và tích hợp phương pháp lấy trung bình ủ (annealed-mean) để tối ưu hóa độ sống động và tính nhất quán không gian của hình ảnh. * Cung cấp các phân tích, đánh giá thực nghiệm và cơ sở lý thuyết về hiệu quả của phương pháp đề xuất, bao gồm việc kiểm tra tính chân thực thị giác thông qua bài kiểm tra Turing tô màu (Colorization Turing Test) và khả năng ứng dụng trong học biểu diễn tự giám sát (self-supervised representation learning) cho các tác vụ phân loại, phát hiện, và phân đoạn ảnh. |
| **NỘI DUNG VÀ PHƯƠNG PHÁP**  **Nội dung 1: Tìm hiểu tổng quan về tô màu ảnh trắng đen sử dụng học sâu Phương pháp thực hiện:**   * Tiến hành nghiên cứu lý thuyết về các phương pháp tô màu ảnh trắng đen, bao gồm các cách tiếp cận không tham số (non-parametric) dựa trên ảnh tham chiếu và các phương pháp tham số (parametric) sử dụng học sâu. * Tìm hiểu các mô hình học sâu, đặc biệt là mạng nơ-ron tích chập (CNN), đã được áp dụng cho bài toán tô màu ảnh, tập trung vào cách sử dụng không gian màu CIE Lab và các hàm mất mát (loss functions) như hồi quy L2 và phân loại đa thức (multinomial classification).   **Kết quả dự kiến:**   * Tóm tắt lý thuyết về tô màu ảnh trắng đen và các phương pháp học sâu liên quan. * Xây dựng nền tảng lý thuyết cho phương pháp đề xuất, tập trung vào phân loại màu sắc và xử lý tính đa thức của bài toán.   **Nội dung 2: Tìm kiếm dữ liệu và huấn luyện mô hình tô màu ảnh**  **Phương pháp thực hiện:**   * Tìm kiếm các bộ dữ liệu ảnh màu lớn, chẳng hạn như ImageNet, để sử dụng làm nguồn huấn luyện. Dữ liệu cần bao gồm ảnh màu chất lượng cao, phù hợp cho việc tạo ảnh xám làm đầu vào và sử dụng kênh màu (a, b) làm tín hiệu giám sát. * Thiết kế và huấn luyện một mô hình CNN dựa trên kiến trúc VGG, bổ sung các lớp tích chập mở rộng (dilated convolutions) để tăng trường tiếp nhận (receptive field). * Thực hiện huấn luyện từ đầu hoặc sử dụng các trọng số khởi tạo (k-means initialization) để tối ưu hóa hiệu suất.   Kết quả dự kiến:   * Thu thập ít nhất một bộ dữ liệu thỏa mãn yêu cầu (ví dụ: ImageNet với 1.3M ảnh). * Huấn luyện thành công mô hình CNN với độ chính xác cao trong việc dự đoán kênh màu a, b từ kênh sáng L.   **Nội dung 3: Phát triển phương pháp phân loại màu sắc với class-rebalancing**  **Phương pháp thực hiện:**   * Xây dựng hàm mất mát phân loại đa thức (multinomial classification loss) để dự đoán phân phối màu sắc cho mỗi pixel, sử dụng không gian màu a, b được lượng tử hóa thành 313 giá trị trong gamut. * Áp dụng kỹ thuật class-rebalancing để xử lý sự mất cân bằng trong phân phối màu sắc tự nhiên, ưu tiên các màu hiếm (saturated colors) bằng cách điều chỉnh trọng số mất mát dựa trên xác suất màu sắc trong ImageNet. * So sánh kết quả với các mô hình sử dụng hàm mất mát hồi quy L2 truyền thống.   **Kết quả dự kiến:**   * Mô phỏng thành công phương pháp phân loại màu sắc với class-rebalancing.   **Nội dung 4: Tìm hiểu và áp dụng kỹ thuật annealed-mean để tối ưu hóa đầu ra**  **Phương pháp thực hiện:**   * Nghiên cứu lý thuyết về cách ánh xạ phân phối xác suất màu sắc sang giá trị điểm (point estimate) trong không gian a, b, tập trung vào kỹ thuật lấy trung bình ủ (annealed-mean) với tham số nhiệt độ T. * Thử nghiệm các giá trị T khác nhau để cân bằng giữa độ sống động (vibrancy) và tính nhất quán không gian (spatial consistency) của kết quả tô màu.   **Kết quả dự kiến:**   * Tóm tắt lý thuyết về annealed-mean và tác động của tham số T. * Xác định giá trị T tối ưu (ví dụ: T = 0.38) để đạt được kết quả tô màu chân thực và nhất quán.   **Nội dung 5: Đánh giá tính chân thực của kết quả tô màu**  **Phương pháp thực hiện:**   * Thực hiện bài kiểm tra Turing tô màu (Colorization Turing Test) trên Amazon Mechanical Turk, yêu cầu người tham gia phân biệt giữa ảnh màu gốc và ảnh được tô màu bởi mô hình. * Đánh giá hiệu suất của mô hình trên các tác vụ thị giác máy tính khác (như phân loại đối tượng sử dụng VGG) để kiểm tra tính ứng dụng của ảnh tô màu. * So sánh kết quả với các phương pháp khác (ví dụ: hồi quy L2, phân loại không rebalancing).   **Kết quả dự kiến:**   * Đạt tỷ lệ đánh lừa người tham gia (fooled rate) cao trong bài kiểm tra Turing (ví dụ: 32% hoặc cao hơn). * Bảng tóm tắt so sánh hiệu suất giữa các phương pháp trên các chỉ số như tính chân thực thị giác và độ chính xác phân loại.   **Nội dung 6: Ứng dụng mô hình trong học biểu diễn tự giám sát**  **Phương pháp thực hiện:**   * Tìm hiểu lý thuyết về học biểu diễn tự giám sát (self-supervised representation learning) và cách bài toán tô màu ảnh có thể được sử dụng như một tác vụ pretext. * Đánh giá hiệu suất của mô hình tô màu trên các tác vụ phân loại, phát hiện, và phân đoạn trên các bộ dữ liệu như PASCAL VOC, so sánh với các phương pháp tự giám sát khác.   **Kết quả dự kiến:**   * Tóm tắt lý thuyết về học biểu diễn tự giám sát thông qua tô màu ảnh. * Kết quả cạnh tranh hoặc vượt trội so với các phương pháp tự giám sát khác trên các tác vụ PASCAL VOC.   **Nội dung 7: Tóm tắt các nội dung, viết báo cáo nghiên cứu và đóng gói chương trình**  **Phương pháp thực hiện:**   * Tổng hợp toàn bộ quá trình nghiên cứu, từ lý thuyết, thực nghiệm đến kết quả. * Chỉnh sửa và đóng gói mã nguồn của mô hình, đảm bảo khả năng tái hiện (reproducibility).   **Kết quả dự kiến:**   * Bản báo cáo nghiên cứu hoàn chỉnh, trình bày rõ ràng các nội dung và kết quả. * Chương trình được đóng gói, sẵn sàng chia sẻ và sử dụng. |
| **KẾT QUẢ MONG ĐỢI**  Trong phạm vi nghiên cứu của đề tài, chúng tôi mong đợi ứng dụng sau khi hoàn thành sẽ được các kết quả sau:   1. Xây dựng thành công mô hình tô màu ảnh trắng đen tự động bằng mạng nơ-ron tích chập (CNN), có khả năng tạo ra ảnh màu sống động, hợp lý và mang tính chân thực cao từ ảnh đầu vào xám. 2. Cải thiện chất lượng và độ đa dạng màu sắc của ảnh tô màu, thông qua việc áp dụng kỹ thuật cân bằng lớp màu (class-rebalancing) và phương pháp lấy trung bình ủ (annealed-mean) nhằm tăng tính phong phú và tự nhiên của hình ảnh. 3. Thực hiện đánh giá khách quan hiệu quả mô hình bằng các chỉ số định lượng như RMSE, PSNR và bài kiểm tra Turing tô màu, chứng minh tính khả thi và độ tin cậy của mô hình. |
| **TÀI LIỆU THAM KHẢO**  [1] Colorful Image Colorization (Oct 2016, ECCV, Author Richard Zhang, Phillip Isola(B), and Alexei A. Efros)  [2] Image Colorization Using the Global Scene-Context Style and Pixel-Wise Semantic Segmentation (Author: Tram-Tran Nguyen-Quynh, S.H. Kim, Nhu-Tai Do )  [3] Learning Representations for Automatic Colorization (Gustav Larsson, Michael Maire, Gregory Shakhnarovich)  [4] Colorization using Optimization (Anat Levin, Dani Lischinski ,Yair Weiss)  [5] Transferring Color to Grayscale Images (C. E. Welsh, M. Ashikhmin, K. Mueller)  [6] Dahl, R.: Automatic colorization (2016). |