


THÔNG TIN CHUNG CỦA BÁO CÁO

- Link YouTube video của báo cáo (tối đa 5 phút):
(ví dụ: <https://www.youtube.com/watch?v=AWq7uw-36Ng>)
- Link slides (dạng .pdf đặt trên Github):
(ví dụ: <https://github.com/mynameuit/CS2205.APR2023/TenDeTai.pdf>)
- Mỗi thành viên của nhóm điền thông tin vào một dòng theo mẫu bên dưới
- Sau đó điền vào Đề cương nghiên cứu (tối đa 5 trang), rồi chọn Turn in

<ul style="list-style-type: none">• Họ và Tên: Nguyễn Lê Nam Anh• MSSV: 230101070 	<ul style="list-style-type: none">• Lớp: CS2205.CH181• Tự đánh giá (điểm tổng kết môn): 8/10• Số buổi vắng: 1• Số câu hỏi QT cá nhân: 1• Link Github: https://github.com/mynameuit/CS2205.APR2023/
---	---

ĐỀ CƯƠNG NGHIÊN CỨU

TÊN ĐỀ TÀI (IN HOA)

TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)

OPTICAL CHARACTER RECOGNITION APPLICATIONS AND ALGORITHMS

TÓM TẮT (Tối đa 400 từ)

- Optical character recognition (ocr) là quá trình chuyển đổi văn bản viết tay, đánh máy hoặc in thành hình ảnh kỹ thuật số mà máy có thể đọc được. Công nghệ này thường được sử dụng để nhận diện và tìm kiếm văn bản trong các tài liệu điện tử hoặc để xuất bản văn bản trên các trang web. Chúng tôi tập trung vào 3 ứng dụng chính và quan trọng của Optical character recognition (ocr) là Captcha, Institutional Repository và Optical Music Character. Chúng tôi sử dụng một thuật toán phân đoạn hình ảnh nâng cao dựa trên cân bằng biểu đồ và thuật toán di truyền cho nhận dạng ký tự quang học
- Chúng tôi sử dụng một thuật toán phân đoạn hình ảnh nâng cao dựa trên cân bằng biểu đồ và thuật toán di truyền cho nhận dạng ký tự quang học để thay đổi thuật toán hướng tiếp cận và cải thiện độ chính xác của ocr so với cách sử dụng thuật toán tiếp cận thông thường của nó là tiếp cận dựa trên biểu đồ được sử dụng để trích xuất thông tin từ hình ảnh hai chiều

GIỚI THIỆU (Tối đa 1 trang A4)

- Vào những năm 1950, Nhận dạng Ký tự Quang học (OCR) được giới thiệu, đã cách mạng hóa quy trình quản lý tài liệu trong nhiều ngành công nghiệp khác nhau. OCR cho phép các tài liệu được quét trở thành các tệp văn bản có thể tìm kiếm, khi máy tính nhận dạng nội dung của chúng. Nó trích xuất thông tin liên quan và nhập vào cơ sở dữ liệu điện tử, loại bỏ nhu cầu gõ lại thủ công.
- OCR có nhiều ứng dụng, bao gồm xử lý hóa đơn, tài liệu pháp lý, ngân hàng và chăm sóc sức khỏe. Nó cũng được sử dụng trong Captcha, kho lưu trữ thể chế, thư viện số, Nhận dạng Âm nhạc Quang học, nhận dạng biển số xe tự động và nhận dạng chữ viết tay.
- Nghiên cứu tập trung vào các thành phần chính của OCR là captcha được sử dụng để phân biệt con người với các tập lệnh tự động nhằm ngăn chặn thư rác. Các kho lưu trữ thể chế và thư viện số nâng cao khả năng hiển thị và tác động toàn cầu của các kết quả đầu ra của một tổ chức, nhận dạng ký hiệu âm nhạc và số hóa dữ liệu âm nhạc, đáp ứng các ký hiệu âm nhạc đa dạng.

- Bước đầu tiên trong việc triển khai OCR là phân đoạn đối tượng. Nhiều phương pháp đã được phát triển cho việc phân đoạn, chẳng hạn như sử dụng biểu đồ để xác định ngưỡng. Một số phương pháp đòi hỏi nhiều tính toán, sử dụng các kỹ thuật như làm sắc nét thung lũng và phương pháp chênh lệch biểu đồ. Các phương pháp tiên tiến như Logic Mờ và Mạng Nơ-ron Nhân tạo cũng đã được khám phá cho OCR. Thuật toán di truyền đặc biệt hiệu quả trong việc giải quyết các vấn đề quy mô lớn và có thể được sử dụng để xác định ngưỡng cho phân đoạn hình ảnh.
- Hệ thống OCR được mô tả trong nghiên cứu này sử dụng cân bằng biểu đồ và thuật toán di truyền để nâng cao và trích xuất các ký tự từ hình ảnh. Hiệu suất của hệ thống được thử nghiệm trên các hình ảnh có nhiễu, mờ và thay đổi ánh sáng

MỤC TIÊU

(Viết trong vòng 3 mục tiêu, lưu ý về tính khả thi và có thể đánh giá được)

- Hiểu được các ứng dụng quan trọng của optical character recognition (ocr)
- Thông tin về thuật toán thông thường và cải thiện
- Đánh giá kết quả

NỘI DUNG VÀ PHƯƠNG PHÁP

(Viết nội dung và phương pháp thực hiện để đạt được các mục tiêu đã nêu)

1. Captcha:

- CAPTCHA là một chương trình tạo ra và đánh giá các bài kiểm tra mà con người có thể vượt qua nhưng các chương trình máy tính hiện tại không thể. Việc bị hack là một mối đe dọa nghiêm trọng đối với không gian mạng
- Ngày nay, nhiều hoạt động của con người như giao dịch kinh tế, tuyển sinh giáo dục, đăng ký, đặt vé du lịch, v.v., được thực hiện qua internet và tất cả đều cần mật khẩu, điều này có thể bị hacker lợi dụng.
- Hacker sử dụng các phương pháp như tấn công từ điển và đăng ký tự động giả mạo, làm lãng phí bộ nhớ và tài nguyên của trang web. Ví dụ như tấn công từ điển (dictionary attacks), tự động khoá trượt trang, ...
- Tấn công từ điển (dictionary attacks) là một cuộc tấn công vào các hệ thống xác thực mật khẩu, nơi hacker viết chương trình để thử nhiều mật khẩu khác nhau từ một danh sách các mật khẩu phổ biến. Trong CAPTCHA, một hình ảnh gồm một chuỗi các chữ cái hoặc số được tạo ra và bị che mờ bằng các kỹ thuật như biến dạng hình ảnh, thay đổi kích thước và phông chữ, nền gây mất tập trung, đoạn ngẫu nhiên, điểm nổi bật và nhiễu trong hình ảnh. Hệ thống này có thể được sử

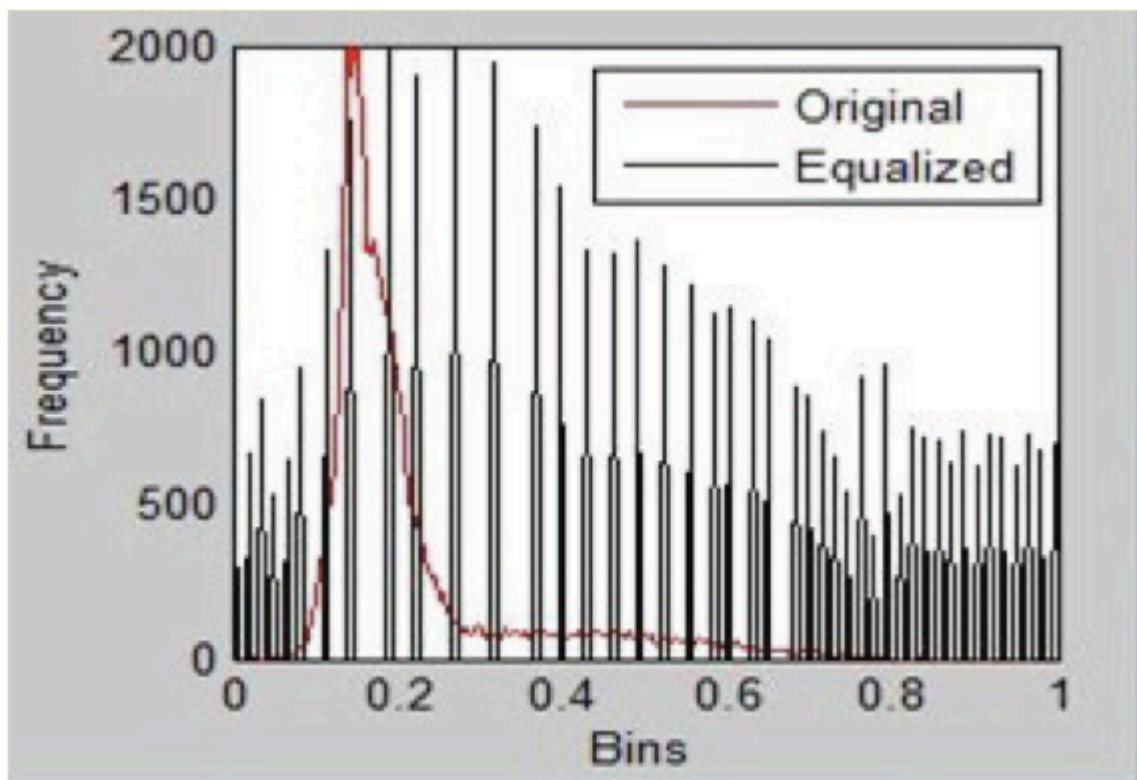
dụng để loại bỏ nhiễu và phân đoạn hình ảnh, làm cho hình ảnh có thể xử lý được bởi các hệ thống OCR (Nhận dạng Ký tự Quang học)

2. Institutional repositories

- Kho lưu trữ thể chế (institutional repositories) là các bộ sưu tập kỹ thuật số của các công trình được tạo ra trong một trường đại học hoặc viện nghiên cứu.
- Đây là các trung tâm trực tuyến lưu trữ dữ liệu trí tuệ của một tổ chức, đặc biệt là các viện nghiên cứu, nơi dữ liệu được thu thập, bảo quản và phát sóng.
- Chúng giúp mở rộng khả năng tiếp cận các công trình của một tổ chức và tăng cường tầm ảnh hưởng trên toàn cầu. Kho lưu trữ này hỗ trợ nghiên cứu liên ngành và phát triển cũng như chia sẻ tài liệu giảng dạy kỹ thuật số.
- Các bộ sưu tập này thường bao gồm các bài báo trên tạp chí được bình duyệt, các kỷ yếu hội nghị, dữ liệu nghiên cứu, chuyên khảo, sách, luận văn và luận án, cùng các bài thuyết trình.
- Chức năng chính của chúng là cung cấp tài liệu truy cập mở.
- Việc triển khai thực tế bao gồm thiết lập một hệ thống với máy quét để số hóa tài liệu, sau đó các tài liệu được xử lý bởi hệ thống Nhận dạng Ký tự Quang học để thu thập và lưu trữ thông tin dưới dạng kỹ thuật số.

3. Optical Music Recognition

- Nhận dạng âm nhạc quang học (Optical Music Recognition) là một lĩnh vực tiên tiến xuất hiện vào những năm 1950, tập trung vào việc chuyển đổi các bản nhạc in thành các định dạng có thể chỉnh sửa và chơi được bằng các phương pháp điện tử và điện hóa.
- Hệ thống Optical Music Recognition có nhiều ứng dụng khác nhau, bao gồm xử lý các loại nhạc khác nhau, số hóa dữ liệu âm nhạc quy mô lớn và hỗ trợ các ký hiệu âm nhạc đa dạng.



4. Thuật toán thông thường của OCR

- Thuật toán sử dụng phương pháp dựa trên biểu đồ để trích xuất thông tin từ hình ảnh hai đỉnh. Thuật toán di truyền được áp dụng lên biểu đồ của hình ảnh hai đỉnh để lấy thông tin hữu ích từ nền.
- Biểu đồ là đồ thị hiển thị tần suất xuất hiện của mỗi mức xám trong hình ảnh. Một quần thể ngẫu nhiên có kích thước N được khởi tạo, với các phần tử có giá trị từ 0 đến 255.
- Các thao tác được thực hiện trên các cặp cha mẹ được chọn ngẫu nhiên. Các giá trị xác suất lai ghép và xác suất đột biến thích hợp được thiết lập.
- Sau khi tính toán các giá trị độ thích nghi của con, chiến lược chọn lọc giải đấu cho phép con cạnh tranh với cha mẹ. Một cuộc cạnh tranh giữa các cá nhân được chọn ngẫu nhiên được thực hiện và cá nhân có độ thích nghi tốt nhất được chọn.
- Thuật toán di truyền sử dụng phương pháp này để chọn lọc từ quần thể. Hai cha mẹ và hai con cạnh tranh để chọn ra hai cá nhân tốt nhất.
- Phương pháp này được gọi là mạng lưới đông đúc. Các phần tử kết quả được đặt vào các lớp tương ứng của chúng.
- Phương pháp này thay thế các phần tử cũ trong quần thể bằng các phần tử có độ thích nghi tốt nhất trong thế hệ mới, giúp giảm lỗi thay thế. Lặp lại quá trình này cho tất cả các phần tử dẫn đến hội tụ. Giá trị hội tụ là giá trị mức xám tại điểm cực tiểu giữa hai đỉnh, được sử dụng làm ngưỡng để phân đoạn hình ảnh.

5. Thuật toán cải thiện OCR

- Phương pháp đề xuất được chia thành hai bước. Đầu tiên, biểu đồ hình ảnh được cân bằng, dẫn đến việc phân phối lại các mức độ sáng như thể hiện trong Hình 1. Đối với biểu đồ đầu vào chuẩn hóa, biểu đồ được hiển thị trong Hình 1, nó cho thấy sự phân phối lại các cường độ điểm ảnh. Điều này dẫn đến việc tăng cường độ của các điểm ảnh có mức độ sáng thấp và giảm cường độ của các điểm ảnh có giá trị cao.
- Các cường độ với giá trị thấp được nâng lên các giá trị cao hơn và ngược lại. Sự thay đổi trong các giá trị cường độ này dẫn đến việc cải thiện hình ảnh hai đỉnh. Với sự cải thiện trong biểu đồ, các giá trị cường độ tăng lên cũng như giảm xuống cho mỗi điểm ảnh do sự phân phối lại của biểu đồ. Điều này dẫn đến việc nâng cao các giá trị xám của lớp có giá trị xám gần với dải trên (129 đến 225). Các điểm ảnh thuộc dải trên di chuyển về phía dải trên. Để trích xuất ngưỡng, thuật toán di truyền được áp dụng cho biểu đồ đã cân bằng. Do tăng cường độ điểm ảnh, ngưỡng được chọn sẽ cao hơn so với phương pháp trước và ngưỡng cao hơn sẽ cần thiết để trích xuất các điểm ảnh lớp thấp từ các điểm ảnh lớp cao.
- **Các giá trị điểm ảnh tương ứng với lớp trên có thể dễ dàng tách biệt** khỏi các điểm ảnh của lớp dưới từ các điểm ảnh lớp cao. Khi áp dụng ngưỡng, các giá trị điểm ảnh tương ứng với lớp trên có thể dễ dàng tách biệt khỏi các điểm ảnh của lớp dưới vì cân bằng biểu đồ đã di chuyển các điểm ảnh có giá trị thấp nhưng thuộc về dải trên về phía giá trị cao hơn. Phương pháp này rất hiệu quả và có thể được sử dụng một cách hiệu quả để trích xuất thông tin từ hình ảnh trong các điều kiện khác nhau. Kỹ thuật sử dụng không yêu cầu bất kỳ kỹ thuật làm sắc nét thung lũng nào.

KẾT QUẢ MONG ĐỢI

(Viết kết quả phù hợp với mục tiêu đặt ra, trên cơ sở nội dung nghiên cứu ở trên)

- Các kết quả thu được từ mô phỏng giúp minh họa khả năng của phương pháp áp dụng cho hình ảnh hai đỉnh. Độ bền và hiệu quả của phương pháp đã được kiểm tra thêm trên hai hình ảnh khác nhau bằng máy Intel Core 2 Duo 2.20 GHz. Kết quả từ quá trình xử lý hình ảnh cho phép đánh giá kỹ lưỡng phương pháp đã áp dụng. Nghiên cứu sử dụng các hình ảnh có biểu đồ hai đỉnh. Thuật toán được mã hóa trong MATLAB (Phiên bản 7.10.0.499) trên máy Intel Core 2 Duo 2.20 GHz. Các hình ảnh hai đỉnh của bản nhạc được sử dụng để trích xuất các ký hiệu và mã hóa âm nhạc hữu ích với các thuật toán đã đề cập.

- Mã hóa âm nhạc, kho lưu trữ thể chế và CAPTCHA được đưa vào hệ thống và các thuật toán lần lượt được áp dụng để lấy chữ ký cho cả ba ứng dụng. Để phân đoạn hình ảnh, hai đỉnh trong biểu đồ được lấy bằng cách sử dụng phương pháp crowding. Các tham số của thuật toán được đặt như sau: (i) kích thước quần thể "N" là 20, (ii) xác suất lai ghép $(P_c = 0.9)$, và (iii) xác suất đột biến (P_m) giảm dần theo cấp số nhân bắt đầu từ 0.05. Giá trị ngưỡng được xác định bằng thuật toán di truyền. Theo thuật toán đề xuất, biểu đồ của hình ảnh đầu vào được cân bằng như trong Hình 1. Việc cân bằng này làm cho hình ảnh sáng hơn nhiều so với hình ảnh gốc, di chuyển các điểm ảnh thuộc dải trên lên cường độ cao hơn. Phương pháp này có thể xử lý hiệu quả các thay đổi về ánh sáng và mờ chuyển động. Phân đoạn hình ảnh sau đó được thực hiện trên các hình ảnh được cải thiện bằng cách sử dụng phương pháp dựa trên GA như được mô tả trong thuật toán. Các tham số cho các thuật toán GA tiêu chuẩn cũng được áp dụng tương tự ở đây.

TÀI LIỆU THAM KHẢO (*Định dạng DBLP*)

- [1] K. Bachuwar, A. Singh, G. Bansa, S. Tiwari, "An Experimental Evaluation of Preprocessing Parameters for GA Based OCR Segmentation" in 3rd International Conference on Computational Intelligence and Industrial Applications (PACIIA 2010), 2010, proceedings, Vol. 2, pp. 417 -420.
- [2] M.D. Ganis, C.L. Wilson, J.L. Blue, "Neural network-based systems for handprint OCR applications" in IEEE Transactions on Image Processing, 1998, Vol: 7 , Issue: 8, p.p. 1097 - 1112
- [3] R. Gossweiler, M. Kamvar, S. Baluja, "What's Up CAPTCHA? A CAPTCHA Based On Image Orientation", in WWW, 2009.
- [4] B. Joanna," Building an institutional repository at Loughborough University: some experiences, program: Electronic library and information systems, 2009.
- [5] A. Singh, K. Bacchuwar, A. Choubey, S. Karanam, D. Kumar, "An OMR Based Automatic Music Player", in 3rd International Conference on Computer Research and Development(ICCRD 2011)in,(IEEE Xplore), 2011, Vol. 1, pp. 174-178.
- [6] S.L. Chang, T. Taiwan , L.S. Chen, Y.C. Chung, S.W. Chen, " Automatic license plate recognition" in IEEE Transactions on Intelligent Transportation Systems, 2004, Vol: 5 , Issue: 1, p.p. 42 - 53
- [7] R Plamondon, S. N. Srihari, "On-line and off-line handwriting recognition: a comprehensive survey" IEEE transaction on pattern Analysis and machine Intelligence, 2000, 22(1), 63-84