

Tìm hiểu các phương pháp phân đoạn đối tượng cho bài toán tách ký tự trên biển số xe máy

Ngô Văn Hải

Ngày 26 tháng 5 năm 2024

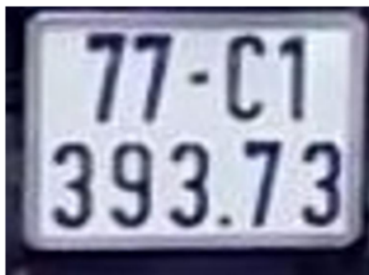
Tóm tắt nội dung

Bài toán phân tách ký tự là bước quan trọng trong hệ thống nhận diện biển số xe tự động. Báo cáo này sẽ trình bày cụ thể hai phương pháp giải quyết bài toán này đó là phương pháp phân đoạn đối tượng dựa trên ngưỡng và phương pháp phân đoạn đối tượng dựa trên cạnh. Việc so sánh giữa hai phương pháp trên cũng được thực hiện nhằm mục đích cho thấy hiệu suất, ưu điểm và nhược điểm của từng phương pháp. Từ đó đưa ra một số thảo luận và phân tích về hai phương pháp trên.

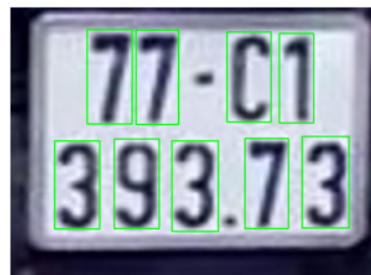
1 Giới thiệu

Hệ thống nhận diện biển số xe tự động hiện đang được ứng dụng nhiều trong thực tế, có thể kể đến các ứng dụng như hệ thống nhà xe tự động [6] và phát hiện phương tiện vi phạm giao thông [4]. Việc nhận diện được thông tin biển số xe giúp cho việc quản lý các phương tiện một cách hiệu quả hơn.

Hệ thống nhận diện biển số xe tự động bao gồm ba bước chính đó là phát hiện biển số, phân tách ký tự trên biển số xe và nhận diện các ký tự. Bài toán phân tách ký tự trên biển số xe là một phần quan trọng trong hệ thống nhận diện biển số xe tự động. Mục tiêu của bài toán là xác định vị trí và ranh giới của từng ký tự trên biển số xe, từ đó tách riêng từng ký tự để có thể nhận diện và giải mã thông tin biển số.



(a) Biển số xe ban đầu.



(b) Biển số xe sau khi phân tách ký tự.

Hình 1: Bài toán phân tách ký tự có đầu vào là hình (a) và đầu ra là hình (b).

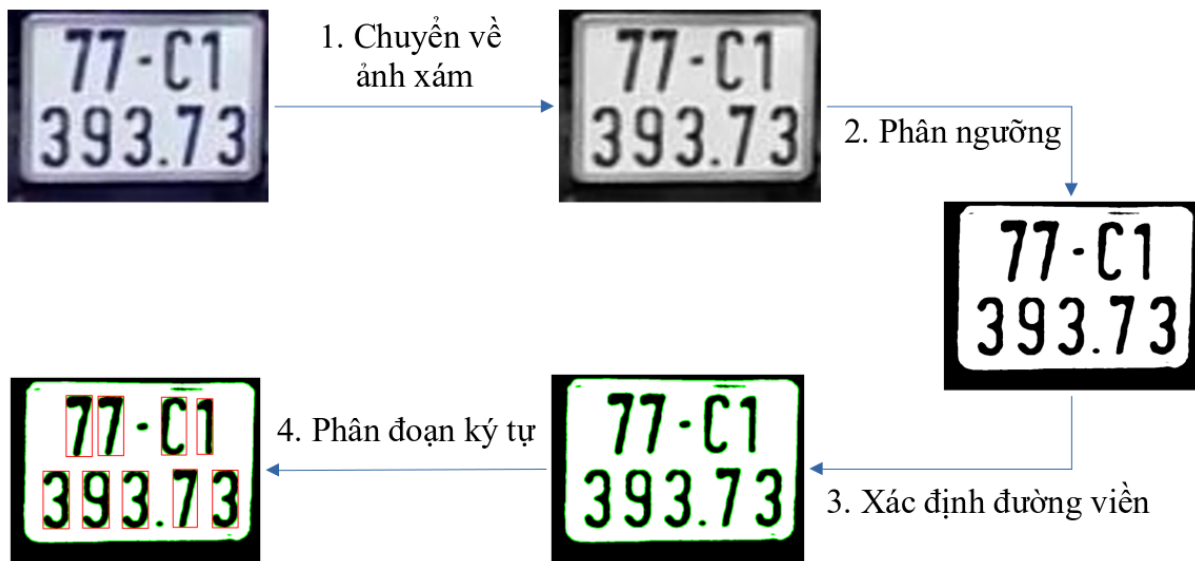
Bài báo cáo này sẽ tập trung tìm hiểu các phương pháp phân đoạn đối tượng trong bài toán phân tách ký tự trên biển số xe máy và tiến hành so sánh các phương pháp trên tập dữ liệu cụ thể. Từ đó có thể thấy được độ hiệu quả của các phương pháp trong bài toán.

2 Phương pháp

Hai phương pháp phân đoạn đối tượng cho bài toán phân tách ký tự trên biển số xe máy được đề cập trong báo cáo này là phân đoạn đối tượng dựa trên ngưỡng (Threshold-based Segmentation) và phân đoạn đối tượng dựa trên cạnh (Edge-based Segmentation).

2.1 Phân đoạn đối tượng dựa trên ngưỡng

Bài toán phân tách ký tự trên biển số xe máy sử dụng phương pháp phân đoạn đối tượng dựa trên ngưỡng bao gồm bốn bước như trong hình 2.



Hình 2: Các bước trong bài toán phân tách ký tự trên biển số xe máy sử dụng phương pháp phân đoạn dựa trên ngưỡng.

2.1.1 Chuyển về ảnh xám

Ảnh đầu vào là ảnh có hệ màu RGB. Mục đích của việc chuyển đổi ảnh có hệ màu RGB thành ảnh xám là để giảm số lượng màu và giúp cho các phương pháp dựa trên ảnh xám hoạt động đúng.



Hình 3: Chuyển ảnh ban đầu về ảnh xám.

2.1.2 Phân ngưỡng

Có hai phương pháp phân ngưỡng [2] có thể áp dụng ở bước này là phân ngưỡng toàn cục và phân ngưỡng cục bộ. Với đặc điểm của biển số xe máy là nền trắng chữ đen thì cả hai phương pháp đều hoạt động tốt ở trong bài toán. Báo cáo này sử dụng phương pháp phân ngưỡng toàn cục dựa trên phương pháp Otsu [9].

Phương pháp Otsu là phương pháp xác định giá trị ngưỡng toàn cục tối ưu từ biểu đồ histogram của hình ảnh.



(a) Ảnh xám ban đầu.



(b) Ảnh phân ngưỡng bằng phương pháp Otsu.

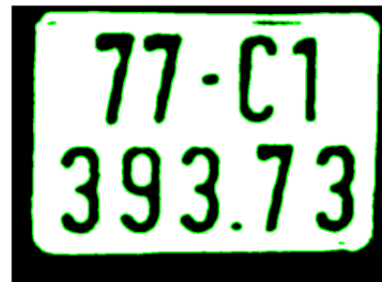
Hình 4: Hình ảnh trước và sau khi áp dụng phương pháp Otsu.

2.1.3 Xác định đường viền

Thuật toán Suzuki [11] được sử dụng để xác định đường viền trong ảnh đã được phân ngưỡng trước đó. Các đường viền là những đường khép kín bao những đối tượng có màu sắc tách biệt với xung quanh. Điều này giúp cho việc xác định vị trí và hình dạng của các ký tự trên biển số xe một cách dễ dàng hơn. Hình 5 là một ví dụ sử dụng thuật toán Suzuki để xác định đường viền cho một ảnh biển số xe được nhị phân hoá.



(a) Biển số xe được nhị phân hoá.

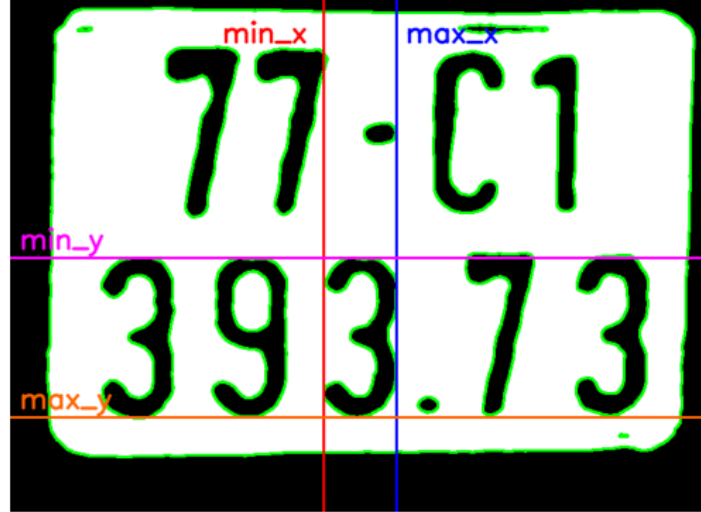


(b) Các đường viền xanh lá trên biển số xe.

Hình 5: Hình ảnh trước và sau khi áp dụng thuật toán Suzuki.

2.1.4 Phân đoạn ký tự

Việc xác định xem đối tượng đó có phải là ký tự số trong biển số xe hay không thì phải dựa vào một số tiêu chí về kích thước và diện tích. Với mỗi đường viền khép kín, xác định xem tọa độ nhỏ nhất và lớn nhất của hai chiều ngang và dọc.



Hình 6: Xác định các tọa độ của đối tượng dựa vào đường viền (min_x là tọa độ nhỏ nhất theo chiều ngang, max_x là tọa độ lớn nhất theo chiều ngang, min_y là tọa độ nhỏ nhất theo chiều dọc, max_y là tọa độ lớn nhất theo chiều dọc).

Gọi w_{box} và h_{box} lần lượt là chiều rộng và chiều cao của đối tượng, w và h lần lượt là chiều rộng và chiều cao của biển số xe. Khi đó, $w_{box} = \text{max_x} - \text{min_x}$ và $h_{box} = \text{max_y} - \text{min_y}$. Hai tiêu chí được áp dụng để phân loại đối tượng trong biển số xe máy:

- Tỷ lệ chiều rộng trên chiều cao:

$$\text{min}_1 \leq \frac{w_{box}}{h_{box}} \leq \text{max}_1$$

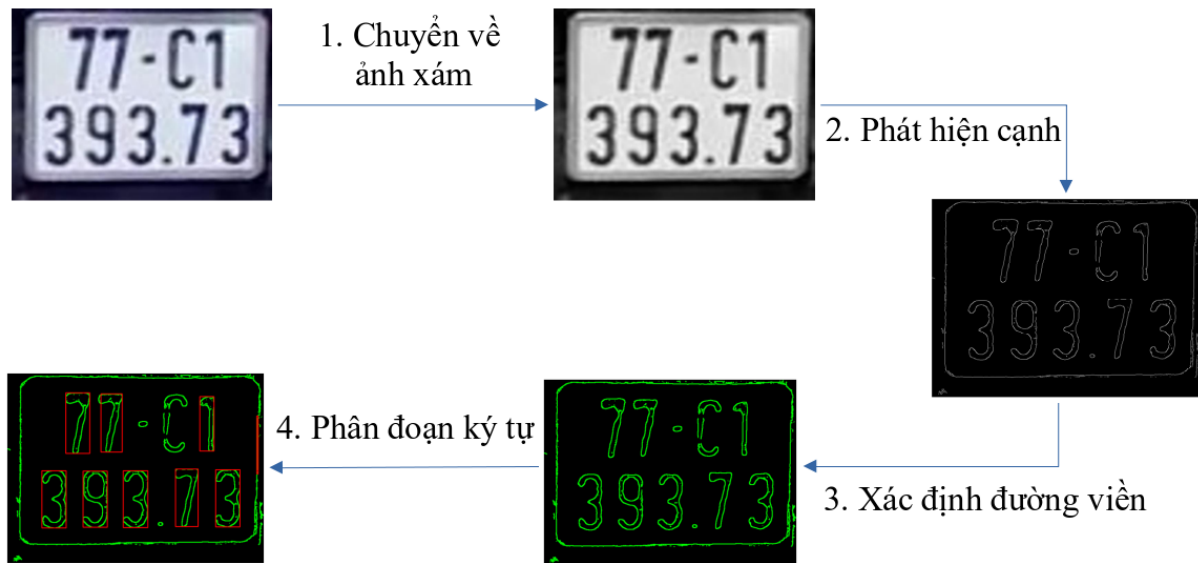
- Tỷ lệ chiều cao của đối tượng trên chiều cao của biển số xe:

$$\text{min}_2 \leq \frac{h_{box}}{h} \leq \text{max}_2$$

với min_i và max_i ($i \in 1, 2$) là các biên của các tỉ lệ.

2.2 Phân đoạn đối tượng dựa trên cạnh

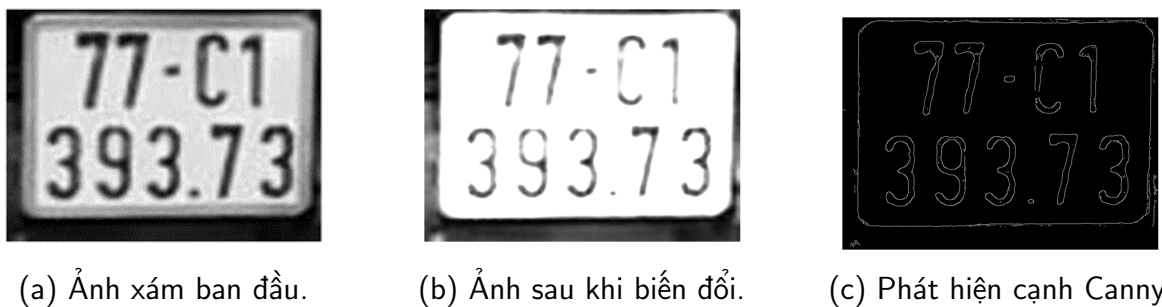
Bài toán phân tách ký tự trên biển số xe máy sử dụng phương pháp phân đoạn đối tượng dựa trên cạnh bao gồm bốn bước như trong hình 7.



Hình 7: Các bước trong bài toán phân tách ký tự trên biển số xe máy sử dụng phương pháp phân đoạn dựa trên cạnh.

Các bước chuyển về ảnh xám, xác định đường viền và phân đoạn ký tự lần lượt tương tự các bước 2.1.1, 2.1.3 và 2.1.4.

Phát hiện cạnh Phương pháp phát hiện cạnh Canny [3] được sử dụng trong bước này để tìm ra cạnh của các đối tượng trên ảnh xám của biển số xe máy. Để tăng khả năng phát hiện cạnh, ảnh xám ban đầu được áp dụng thêm một số phép biến đổi như khử nhiễu bằng bộ lọc Gaussian và tăng độ tương phản.



Hình 8: Các bước trong quá trình phát hiện cạnh bằng phương pháp Canny.

3 Thí nghiệm

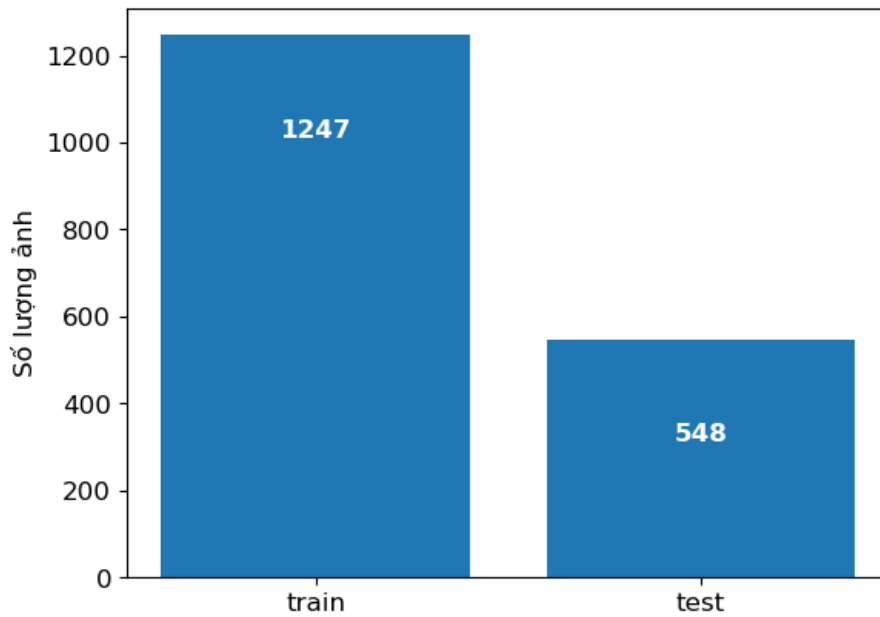
Thí nghiệm này được thực hiện để so sánh hai phương pháp phân đoạn đối tượng trên một tập dữ liệu về biển số xe Việt Nam. Code của thí nghiệm được lưu tại [7].

3.1 Thiết lập thí nghiệm

3.1.1 Tập dữ liệu

Tập dữ liệu sử dụng trong thí nghiệm được tổng hợp từ tập dữ liệu có tên là Vietnamese License Plate Recognition tại nguồn [1].

Tập dữ liệu được chia thành hai tập là tập train và tập test có số lượng hình ảnh như hình 9.



Hình 9: Biểu đồ thể hiện số lượng hình ảnh trong tập dữ liệu.

Kích thước của tất cả hình ảnh trong tập dữ liệu là 760 x 560 pixels. Các hình ảnh trong tập dữ liệu đều là biển số xe máy có hai hàng như hình 10. Các biển số xe đã được gán nhãn theo định dạng YOLO [10].



Hình 10: Một số loại biển số xe trong tập dữ liệu.

3.1.2 Các tham số được dùng

Đối với phương pháp phân đoạn dựa trên cạnh:

- Kích thước kernel của bộ lọc Gaussian là 5×5 .
- Sử dụng hàm `addWeighted` của thư viện OpenCV [8] với tham số $\alpha = 2.0$ để tăng độ tương phản của hình ảnh.
- Sử dụng hàm `Canny` của thư viện OpenCV với tham số $\text{minVal} = 50$ và $\text{maxVal} = 100$ để phát hiện cạnh.

Các khoảng tiêu chí tại bước 2.1.4 được xác định dựa vào tập train như sau:

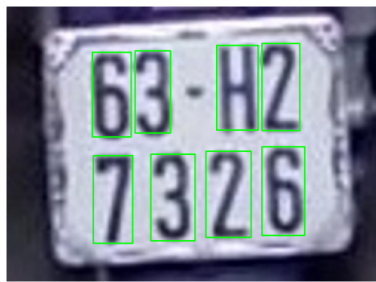
- min_1 là giá trị $\frac{w_{box}}{h_{box}}$ tối thiểu của tất cả đối tượng trong tập train.
- max_1 là giá trị $\frac{w_{box}}{h_{box}}$ tối đa của tất cả đối tượng trong tập train.
- min_2 là giá trị $\frac{h_{box}}{w_{box}}$ tối thiểu của tất cả đối tượng trong tập train.
- max_2 là giá trị $\frac{h_{box}}{w_{box}}$ tối đa của tất cả đối tượng trong tập train.

Và các con số cụ thể của khoảng tiêu chí sau khi xác định dựa vào tập train trong thí nghiệm như sau:

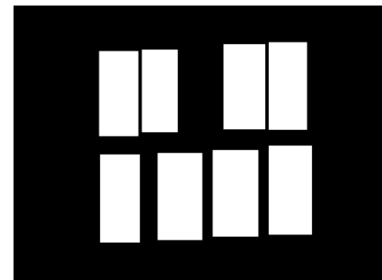
- $[\text{min}_1, \text{max}_1] = [0.225, 0.642]$
- $[\text{min}_2, \text{max}_2] = [0.119, 0.611]$

3.1.3 Các độ đo dùng để so sánh

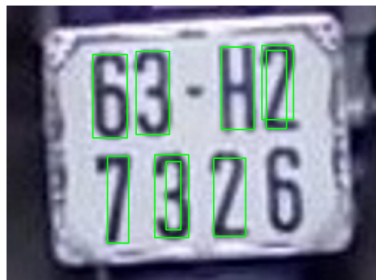
Trong thí nghiệm này, độ đo IoU [5] được sử dụng để so sánh hiệu suất phân đoạn ký tự của hai phương pháp. Đánh giá IoU giữa hai hình ảnh ground truth như 11b và hình ảnh mặt nạ dự đoán như 11d cho biết được hiệu suất của phương pháp đang sử dụng.



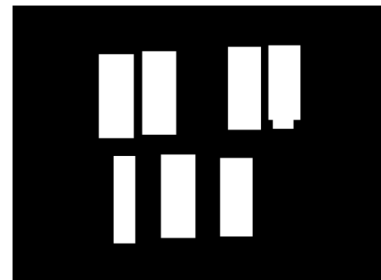
(a) Nhãn thực tế.



(b) Ground truth.



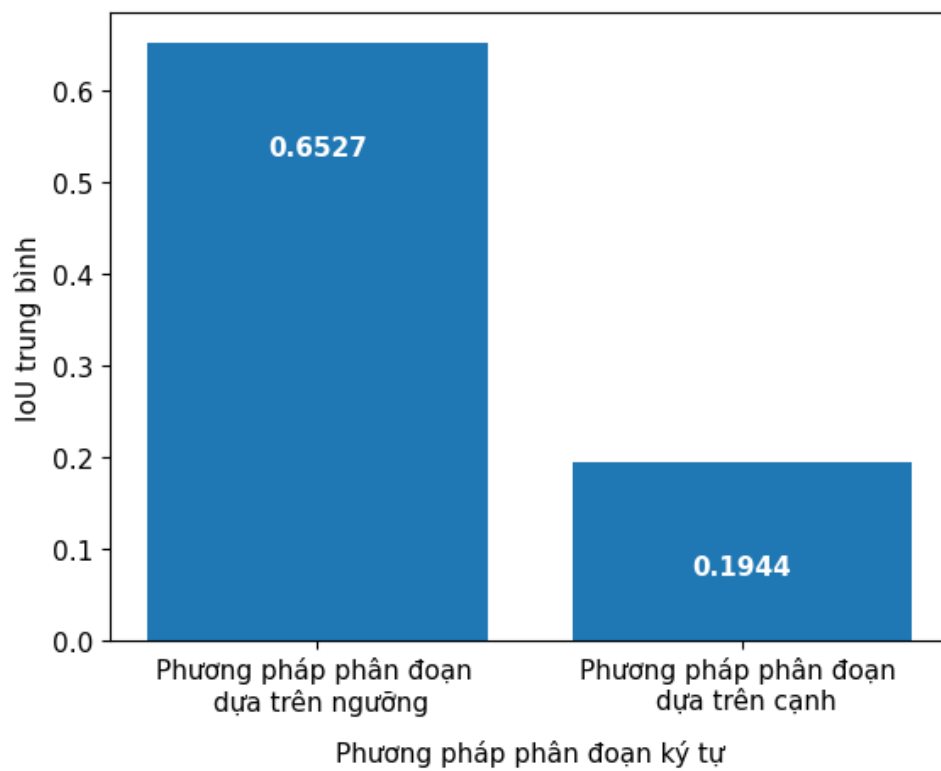
(c) Nhãn dự đoán được.



(d) Mặt nạ của nhãn dự đoán.

Hình 11: Chuyển đổi nhãn về mặt nạ.

3.2 Kết quả



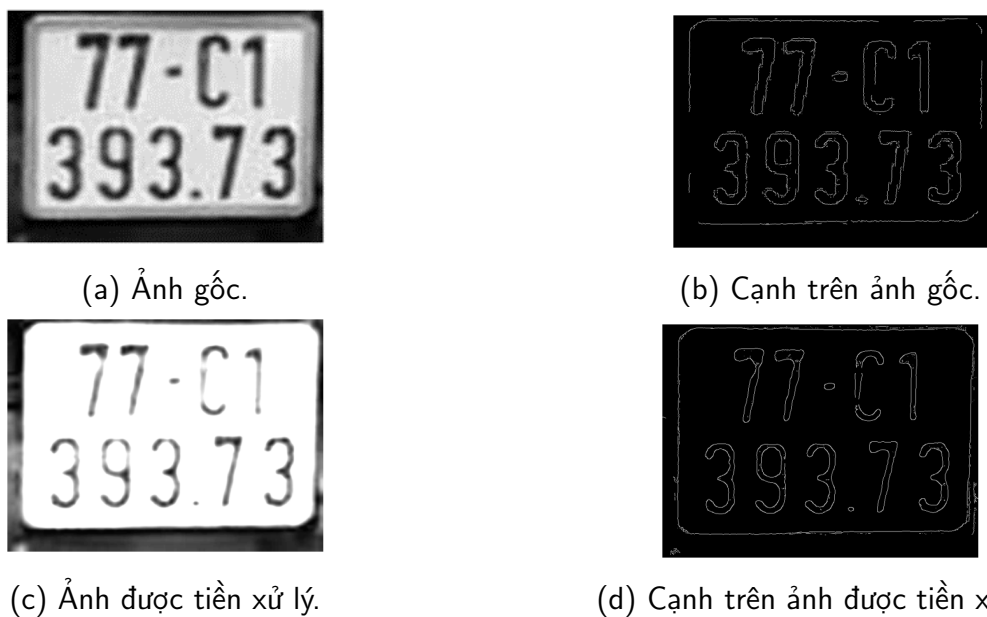
Hình 12: IoU trung bình của hai phương pháp khi thử nghiệm trên tập test.



Hình 13: Một số kết quả thực tế. Mỗi kết quả gồm ba ảnh, ảnh bên trái là nhãn thực tế, ảnh giữa là nhãn được dự đoán bằng phương pháp phân đoạn dựa trên ngưỡng, ảnh bên phải nhãn được dự đoán bằng phương pháp phân đoạn dựa trên cạnh.

4 Thảo luận và phân tích

Dựa vào biểu đồ trong hình 12, phương pháp phân đoạn dựa trên ngưỡng hoạt động tốt hơn phương pháp phân đoạn dựa trên cạnh. Điều này là do ảnh hưởng của các yếu tố như ánh sáng và màu sắc của biển số xe, làm cho thuật toán phát hiện cạnh Canny hoạt động không chính xác. Cạnh ở hình 14d không được liền nét hơn so với cạnh ở hình 14b, nên khi sang bước phát hiện đường viền 2.1.3 sẽ không đạt được hiệu suất cao.



Hình 14: Phát hiện cạnh Canny bị ảnh hưởng bởi ánh sáng và màu sắc.

Phương pháp phân đoạn theo ngưỡng hoạt động khá tốt trong nhiều điều kiện khác nhau của biển số xe như hình 13. Nhược điểm của phương pháp này là không thể phát hiện được những ký tự không liền nét như hình 13d hoặc phát hiện các vùng không phải là ký tự như vòng tròn ở trong ký tự 0 tại hình 13a, 13b, 13c và 13e. Để khắc phục được điều này thì có thể đặt thêm một số tiêu chí tại bước 2.1.4 nhằm nâng cao khả năng nhận diện ký tự của phương pháp.

Tài liệu

- [1] Real-time auto license plate recognition with jetson nano. <https://github.com/winter2897/Real-time-Auto-License-Plate-Recognition-with-Jetson-Nano>. Accessed: 2024-05-18.
- [2] K Bhargavi and S Jyothi. A survey on threshold based segmentation technique in image processing. *International Journal of Innovative Research and Development*, 3(12):234–239, 2014.
- [3] John Canny. A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, (6):679–698, 1986.
- [4] Duong Duc Cuong, Nguyen Tuan Anh, Pham Mai Anh, Ha Minh Duc, Nguyen Minh, et al. Automatic traffic red-light violation detection using ai. *Ingenierie des Systemes d’Information*, 27(1), 2022.
- [5] Mark Everingham, Luc Van Gool, Christopher KI Williams, John Winn, and Andrew Zisserman. The pascal visual object classes (voc) challenge. *International journal of computer vision*, 88:303–338, 2010.
- [6] John Patrick Fabros, D Tabañag, A Espira, and OJ Gerasta. Automated car park management system. In *IOP Conference Series: Materials Science and Engineering*, volume 79, page 012009. IOP Publishing, 2015.
- [7] Ngo Van Hai. License plate character segmentation. <https://github.com/VHSkillPro/license-plate-character-segmentation>, 2024. Accessed: 2024-05-25.
- [8] Itseez. Open source computer vision library. <https://github.com/itseez/opencv>, 2015. Accessed: 2024-05-20.
- [9] Nobuyuki Otsu et al. A threshold selection method from gray-level histograms. *Automatica*, 11(285-296):23–27, 1975.
- [10] Joseph Redmon, Santosh Kumar Divvala, Ross B. Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. *CoRR*, abs/1506.02640, 2015.
- [11] Satoshi Suzuki et al. Topological structural analysis of digitized binary images by border following. *Computer vision, graphics, and image processing*, 30(1):32–46, 1985.