

ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN
KHOA KHOA HỌC MÁY TÍNH



BÁO CÁO ĐỒ ÁN
NHẬP MÔN THỊ GIÁC MÁY TÍNH
CS231.M21.KHCL
ĐỀ TÀI: MỘT SỐ KỸ THUẬT PHÂN ĐOẠN ẢNH

GIẢNG VIÊN HƯỚNG DẪN: TS. MAI TIẾN DŨNG

SINH VIÊN THỰC HIỆN: NGUYỄN PHAN QUỐC THIỆN – 20520775

ĐỖ ĐỨC THỊNH - 20520780

TP. HỒ CHÍ MINH, 6/2022

MỤC LỤC

Phần 1. Giới thiệu môn học và đề tài	2
Phần 2. Giới thiệu về phân đoạn ảnh	3
2.1 Khái niệm	3
2.2 Ứng dụng	3
Phần 3. Các kỹ thuật phân đoạn hình ảnh	4
3.1 Phân đoạn dựa trên cạnh (Edge Based Segmentation).....	5
3.2 Phân đoạn dựa trên khu vực (Region-based Segmentation)	6
Phần 4. Thực nghiệm	8
4.1 Input và output.....	8
4.2 Độ đo IoU	8
4.3 Thực nghiệm phân đoạn dựa trên cạnh	9
4.4 Thực nghiệm phân đoạn dựa trên khu vực.....	11
Phần 5: Kết quả	15
Phần 6. So sánh	18
Tài liệu tham khảo.....	19

Phần 1. Giới thiệu môn học và đề tài

Thị giác máy tính là một chuyên ngành hẹp trong trí tuệ nhân tạo để giúp máy tính có thể nhìn thấy được, hiểu được những hình ảnh như con người. Nếu AI là bộ não giúp cho máy tính biết suy nghĩ thì thị giác máy tính như là đôi mắt cho phép máy tính quan sát, ghi nhận hình ảnh và hiểu chúng. Hiện nay với sự phát triển mạnh mẽ của công nghệ 4.0, các ứng dụng của thị giác máy tính đang được ứng dụng rộng rãi trong nhiều lĩnh vực của đời sống. Một số ứng dụng có thể kể đến như: phân loại hình ảnh, nhận diện vật thể, truy xuất dữ liệu,

Phân đoạn ảnh (image segmentation) là một kỹ thuật xử lý ảnh quan trọng trong thị giác máy tính, là bước tiền xử lý ảnh để có thể dễ dàng giải quyết những bước tiếp theo trong một bài toán về thị giác máy tính.

Trong đồ án môn học lần này, nhóm chúng em tập trung tìm hiểu về 2 kỹ thuật phân đoạn hình ảnh cơ bản, cụ thể là phân đoạn theo cạnh và phân đoạn theo vùng (region growing) vì những kỹ thuật khác đã được giới thiệu qua ở trên lớp. Do là lần đầu tìm hiểu về lĩnh vực thị giác máy tính nên nhóm chúng em có những thiếu sót, hạn chế, rất mong nhận được những nhận xét của thầy về đồ án này để hoàn thiện thêm.

Phần 2. Giới thiệu về phân đoạn ảnh

2.1 Khái niệm

Phân đoạn ảnh (Image Segmentation) là một phương pháp phân chia một hình ảnh kỹ thuật số thành nhiều vùng ảnh nhỏ (tập hợp các pixel). Mục tiêu của phân đoạn là đơn giản hóa hoặc thay đổi hình ảnh đại diện thành một thứ gì đó có ý nghĩa hơn và dễ phân tích hơn. Chính xác hơn, phân đoạn hình ảnh là quá trình gán nhãn cho mọi pixel trong một hình ảnh sao cho các pixel có cùng nhãn có những đặc điểm tương đồng nhất định.

2.2 Ứng dụng

Trong Thị giác máy tính, phân đoạn ảnh là một kỹ thuật quan trọng, giúp giải nhiều bài toán thuộc các lĩnh vực khác nhau như xử lý ảnh y tế, phát hiện và nhận dạng đối tượng, hệ thống camera thông minh... Đây là tiền đề của quá trình xử lý dữ liệu hình ảnh. Kết quả phân đoạn tốt sẽ tạo điều kiện thuận lợi cho các khâu xử lý về sau, tiết kiệm thời gian, gia tăng mức độ chính xác, đồng thời giảm thiểu nguồn lực tính toán.

- Y học: phát hiện khối u, phân đoạn và phát hiện các cơ quan, bộ phận,...
- Giao thông: xe tự hành, nhận diện làn đường, xe cộ, người đi bộ, đèn giao thông, biển báo,...
- Xử lý ảnh vệ tinh: phát hiện sông, núi, tuyến đường,...

Phần 3. Các kỹ thuật phân đoạn hình ảnh

Có 2 cách tiếp cận phân đoạn ảnh:

- Cách tiếp cận tương đồng (Similarity approach), có nghĩa là phát hiện sự tương đồng giữa các pixel hình ảnh để tạo thành một phân đoạn, dựa trên một ngưỡng. Các thuật toán học máy như phân cụm thường dựa trên kiểu tiếp cận này để phân đoạn một hình ảnh.
- Cách tiếp cận gián đoạn (Discontinuity approach): Cách tiếp cận này dựa trên sự gián đoạn của các giá trị cường độ pixel trong hình ảnh. Các kỹ thuật phát hiện đường, điểm và cạnh sử dụng kiểu tiếp cận gián đoạn để thu được các kết quả phân đoạn trung gian. Kết quả này sau đó có thể được xử lý để cho ra hình ảnh được phân đoạn cuối cùng.

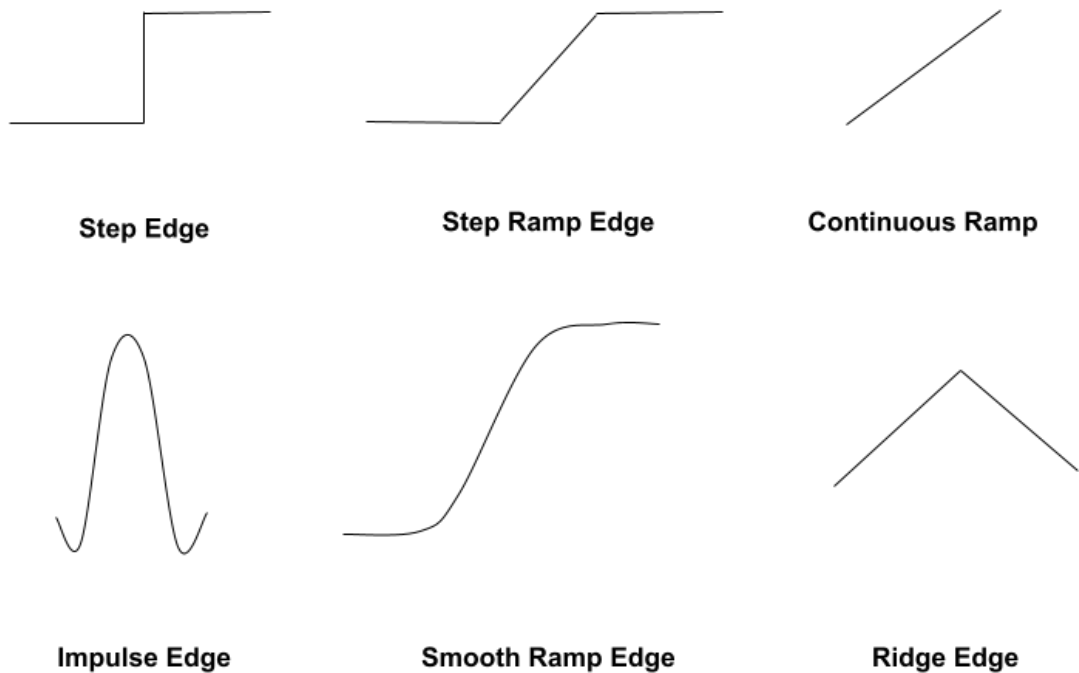
Có 5 kỹ thuật phân đoạn ảnh:

- Phân đoạn dựa trên ngưỡng (Threshold Based Segmentation)
- Phân đoạn dựa trên cạnh (Edge Based Segmentation)
- Phân đoạn dựa trên khu vực (Region-Based Segmentation)
- Phân đoạn dựa trên kỹ thuật phân cụm (Clustering Based Segmentation)
- Phân đoạn dựa trên mạng nơon nhân tạo (Artificial Neural Network Based Segmentation)

3.1 Phân đoạn dựa trên cạnh (Edge Based Segmentation)

Cạnh trong ảnh đánh dấu những vị trí hình ảnh không liên tục về mức xám, màu sắc, kết cấu, v.v. Khi di chuyển từ vùng này sang vùng khác, mức xám có thể thay đổi. Vì vậy, nếu tìm thấy sự gián đoạn đó, ta có thể tìm thấy cạnh. Thực tế, có nhiều toán tử phát hiện cạnh, nhưng hình ảnh thu được là kết quả phân vùng trung gian, và không nên nhầm lẫn với hình ảnh được phân vùng cuối cùng. Để ra được kết quả cuối, cần thực hiện một số bước bổ sung bao gồm: kết hợp các phân vùng cạnh thu được làm một, để giảm số lượng phân vùng và có được một đường viền liên mạch của đối tượng.

Như vậy, có thể thấy, phân vùng cạnh đưa ra một kết quả phân vùng trung gian. Kết quả này sau đó có thể áp dụng theo vùng hoặc bất kỳ kiểu phân đoạn nào khác, nhằm có được hình ảnh được phân vùng cuối.



Các cạnh thường được liên kết với “Độ lớn” và “Hướng”. Một số toán tử phát hiện cạnh cung cấp cả hai yếu tố này, chẳng hạn như Sobel edge operator, canny edge detector, Kirsch edge operator, Prewitt edge operator, Robert’s edge operator,....

3.2 Phân đoạn dựa trên khu vực (Region-based Segmentation)

Đối với kỹ thuật này thì có 2 thuật toán phổ biến: Region growing và Splitting and Merging.

Region growing (phát triển khu vực)

Đối với kỹ thuật này, bắt đầu với những pixel hạt giống (seed point) và sau đó kiểm tra các điểm liền kề. Nếu các điểm liền kề thỏa mãn 1 tính chất được quy định từ trước (mức xám, màu sắc), thì điểm đó sẽ được thêm vào vùng

của điểm hạt giống và quá trình sẽ tiếp tục lặp lại cho đến khi không còn điểm tương đồng nào.

VD: seed point = 8 (màu vàng), thresh < 3

1	3	6	3	2	1
2	1	1	6	2	5
7	1	9	8	7	7
1	3	2	7	8	1
6	2	9	6	1	2
3	1	1	2	3	4

1	3	6	3	2	1
2	1	1	6	2	5
7	1	9	8	7	7
1	3	2	7	8	1
6	2	9	6	1	2
3	1	1	2	3	4

1	3	6	3	2	1
2	1	1	6	2	5
7	1	9	8	7	7
1	3	2	7	8	1
6	2	9	6	1	2
3	1	1	2	3	4

Phần 4. Thực nghiệm

Nhóm em chọn ra 1 ảnh để nêu từng bước làm đối với 2 kỹ thuật này

4.1 Input và output

Input:

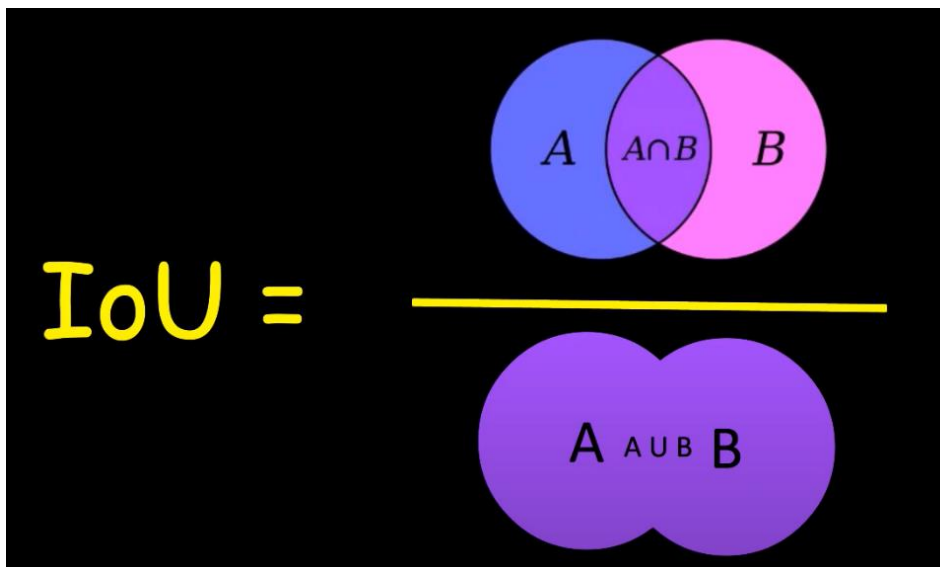
- Ảnh màu chứa 1 vật thể cần phân đoạn
- Ảnh có độ tương phản cao
- Background đơn giản, có màu khác với vật thể

Output: ảnh mask nhị phân, màu trắng là vật thể đã phân đoạn, màu đen là background.

4.2 Độ đo IoU

IoU (Intersection-Over-Union) là một thước đo đánh giá phổ biến để phân đoạn hình ảnh.

$\text{IoU} = \text{Diện tích phần GIAO} / \text{Diện tích phần HỢP}$



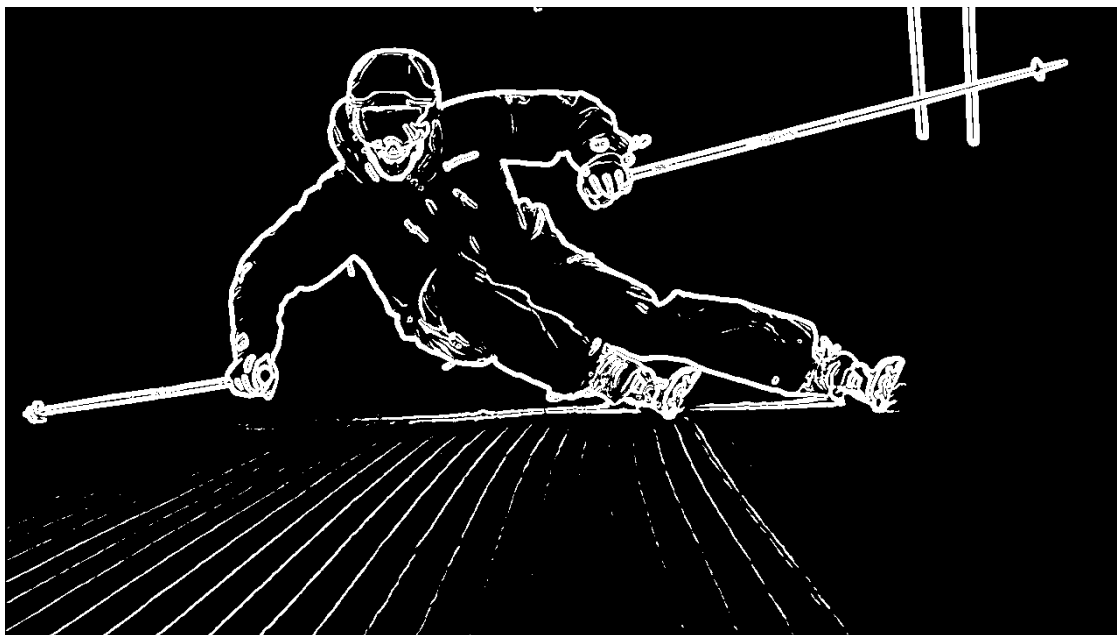
4.3 Thực nghiệm phân đoạn dựa trên cạnh

Trong phần báo cáo các bước xử lý để đưa ra một phân vùng ảnh cuối cùng như sau:

1. Tiền xử lý: làm mờ ảnh bằng Bộ lọc gaussian được khởi tạo thông qua hàm `cv2.GaussianBlur()`
2. Sử dụng toán tử sobel với `ddepth = cv2.CV_64F` thông qua hàm `sobel` với hai trục x và y



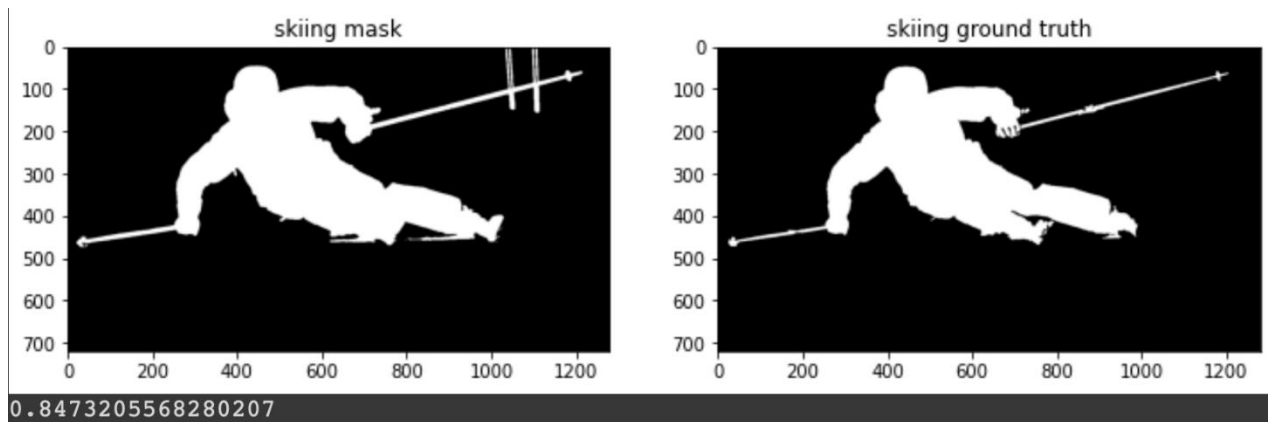
3. Ảnh sau khi tìm cạnh chưa thể hiện rõ các cạnh do vẫn còn nằm trong những mức xám khác nhau. Để làm rõ, ta sử dụng ngưỡng (threshold) để chia ảnh thành 2 màu trắng và đen, đối với các pixel có mức xám bé hơn ngưỡng ta cho thành màu đen và ngược lại là màu trắng



4. Sau khi có ảnh trắng đen, để có được ảnh phân vùng cuối, ta sử dụng hàm `cv2.findcontour()` để tìm các pixel liên tiếp với nhau. Sau đó, tìm contour có diện tích bao quát vật thể cần được mask.
5. Để mask các contour, sử dụng `cv2.drawcontour()`, ảnh đầu ra là ảnh phân vùng vật thể cần tìm



6. So sánh với ảnh ground truth bằng IoU.



(*) Đối với mỗi ảnh có màu sắc, độ tương phản, vật thể cần phân đoạn khác nhau, cần chỉnh các thông số phù hợp.

4.4 Thực nghiệm phân đoạn dựa trên khu vực

Trong phần báo cáo các bước xử lý để đưa ra một phân vùng ảnh cuối cùng như sau:

1. Tiền xử lý: một số ảnh mà vật thể cần phân đoạn có nhiều thì phải làm mờ để thuật toán xử lý hiệu quả hơn.
2. Chọn pixel hạt giống (seed point) cho vật thể cần phân đoạn là chấm màu đỏ trên ảnh bằng cách chọn tọa độ x và y tương đối là từ 0 đến 1. Lí do: giúp việc chọn vị trí seed point dễ dàng hơn, cũng như mỗi ảnh có nhiều kích thước khác nhau tránh việc resize ảnh, sau khi chọn x và y tương đối sẽ nhân với kích thước dài, rộng của ảnh để ra được tọa độ tuyệt đối của seed point.



3. Tạo 1 mask có giá trị toàn bộ là 0
4. Kiểm tra các pixel nằm lân cận seed point có hợp lệ (nằm trong ảnh), trích xuất ra 3 giá trị màu r,g,b để so sánh với seed point, nếu chênh lệch các giá trị $< \text{thresh}$ thì cho chúng vào tập hợp các điểm tương đồng với seed point, đồng thời đánh dấu những pixel tương đồng đó giá trị là 255 trên mask. *Lưu ý: chọn thresh vừa phải, nếu quá nhỏ sẽ không phân vùng được hết vật thể, còn quá lớn sẽ phân vùng vào background.*
5. Tiếp tục xét các pixel lân cận của tập hợp các điểm tương đồng với seed point và lặp lại các bước tương tự.

Mask

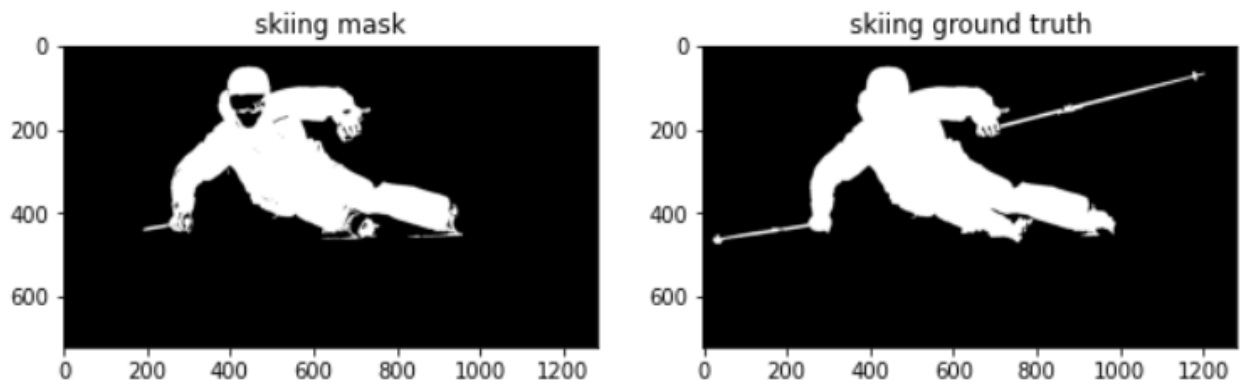


Các vùng màu vàng là tập hợp các điểm tương đồng với seed point.



6. So sánh với ảnh ground truth bằng IoU.

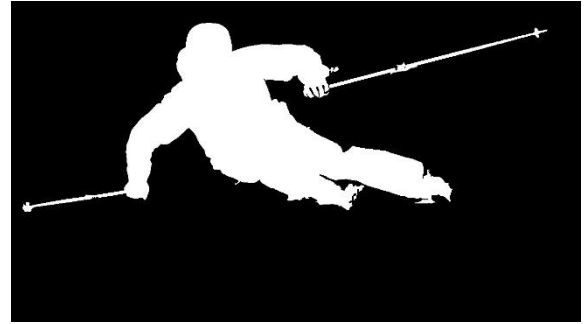
IoU of skiing = 0.8730130290825175



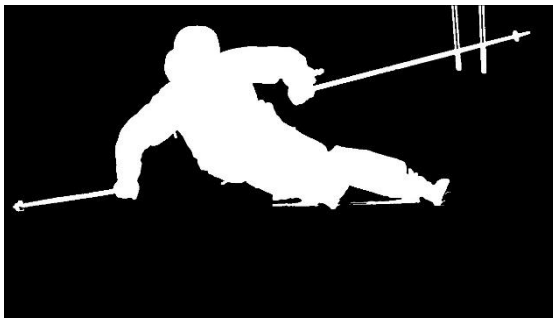
(*) Đối với mỗi ảnh có màu sắc, độ tương phản, vật thể cần phân đoạn khác nhau, cần chỉnh các thông số phù hợp.

Phần 5: Kết quả

Sau khi đã thử nghiệm trên rất nhiều ảnh thì nhóm em chọn ra 3 ảnh có kết quả tương đối tốt để báo cáo:



Ground truth



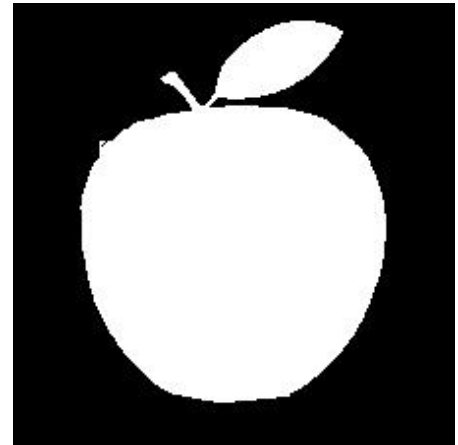
Edge-based



Region-based

Ảnh “skiing”:

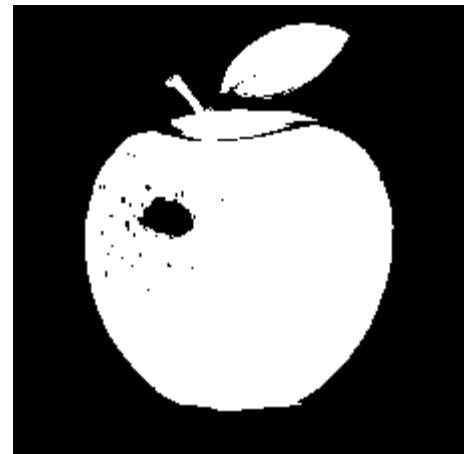
- Edge-based: vật thể người được tách rõ ràng, tuy nhiên phân đoạn dư do background tồn tại vật thể khác bị nhận diện là cạnh quan trọng.
- Region-based: phân đoạn khá ổn trừ khuôn mặt và đôi giày trượt tuyết có màu khác màu đen.



Ground truth



Edge-based



Region-based

Ảnh “apple”:

- Edge-based: đưa ra ảnh mask rất tốt, nhưng các phân đoạn hơi thừa ở lá và cuống
- Region-based: sau khi chọn 2 seed point ở lá và quả thì thu được kết quả có 1 vùng nhỏ trên quả táo không phân đoạn được do màu sắc quá khác biệt



Ground truth



Edge-based



Region-based

Ảnh “cat”:

- Edge-based: phân đoạn tốt, chỉ có phần râu là chưa tốt lắm.
- Region-based: do vật thể con mèo có màu lông tương đối phức tạp (vùng sáng, vùng tối) và background có 1 chấm mờ màu vàng trùng màu lông nên đã phân đoạn bị dư.

Phần 6. So sánh

Sau khi thực nghiệm các ảnh nhóm em có một số so sánh về 2 kỹ thuật phân đoạn dựa trên cạnh và phân đoạn dựa trên khu vực:

Edge-based	Region-based (region growing)
Phù hợp với ảnh mà vật thể cần phân đoạn có cạnh rõ ràng, background ít chi tiết cạnh. Vật thể càng ít chi tiết càng tốt, có các cạnh quan trọng là viền ngoài	Phù hợp với ảnh mà vật thể cần phân đoạn có màu gần giống nhau và tách biệt với màu của background. Vật thể nên có màu gần giống nhau, vùng màu liên tục không bị chia cắt bởi các màu khác.
Tốt hơn region-based khi vật thể cần phân đoạn có nhiều màu khác nhau, nhiều màu sắc gây nhiễu.	Tốt hơn edge với các ảnh có background nhiều cạnh, vật thể phát hiện cạnh khó.

Tài liệu tham khảo

- [1] “Image segmentation,” *Wikipedia*. Apr. 01, 2022. Accessed: Apr. 02, 2022. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Image_segmentation&oldid=1080452824
- [2] “What is Computer Vision? | IBM.” <https://www.ibm.com/topics/computer-vision> (accessed May 24, 2022).
- [3] “Phân vùng ảnh là gì? Một số kỹ thuật phân vùng phổ biến,” *VinBigdata Product*, Sep. 01, 2021. <https://product.vinbigdata.org/phan-vung-anh-la-gi-mot-so-ky-thuat-phan-vung-pho-bien/> (accessed May 24, 2022).
- [4] M. Tyagi, “Image Segmentation : Part 1,” *Medium*, Jul. 19, 2021. <https://towardsdatascience.com/image-segmentation-part-1-9f3db1ac1c50> (accessed May 25, 2022).
- [5] M. Tyagi, “Image Segmentation: Part 2,” *Medium*, Jul. 24, 2021. <https://towardsdatascience.com/image-segmentation-part-2-8959b609d268> (accessed May 25, 2022).