

Bài tập lớn 2:

Ứng dụng lý thuyết xác suất thống kê vào phân ngưỡng ảnh - Bài tập 4

Mục lục

Bảng phân công công việc của các thành viên.....	2
Bảng chấm chéo của các thành viên:	2
Lời mở đầu báo cáo.....	3
I. Kiến thức liên quan cơ bản	4
I.1. Kiến thức về phân đoạn ảnh.....	4
I.2. Phân đoạn ảnh xám	4
II. Tiếp cận – giải quyết yêu cầu bài toán.....	6
II.1. Cơ sở lý thuyết.....	6
II.2. Giải quyết yêu cầu 1 – Bài toán dẫn nhập.....	6
II.3. Giải quyết yêu cầu 2 – Lập trình.....	7
III. Hạn chế của phương pháp – Đề xuất phương án khắc phục	10
III.1. Hạn chế	10
III.2. Khắc phục	11

Bảng phân công công việc của các thành viên

- (1) Đào Quang Huy:
 - Soạn phần cơ sở lý thuyết của thuật toán tìm mức phân ngưỡng
 - Nhận xét – Nêu ra các điểm hạn chế của thuật toán – Tìm minh họa (Cùng Anh Kha)
- (2) Võ Dân Anh Kha:
 - Trình bày – giải quyết Yêu cầu 1 của BTL
 - Nêu ra các điểm hạn chế của thuật toán – Chạy thuật toán trên các minh họa hạn chế (Cùng Quang Huy)
- (3) Nguyễn Hoàng Khang:
 - Tìm hiểu yêu cầu 2 – Viết chương trình tìm ngưỡng và phân đoạn ảnh xám (cùng Nguyễn Khoa).
 - Đề xuất các giải pháp khắc phục các hạn chế của thuật toán (cùng Xuân Khoa).
- (4) Nguyễn Phan Vĩnh Khang:
 - Soạn phần “Kiến thức cơ bản liên quan” cho bài báo cáo (cùng Xuân Khoa)
 - Tìm hiểu yêu cầu 2 – Viết lưu đồ thuật toán để giải quyết yêu cầu
- (5) Nguyễn Bùi Nguyên Khoa:
 - Tìm hiểu yêu cầu 2 – Viết chương trình tìm ngưỡng và phân đoạn ảnh xám (cùng Hoàng Khang).
 - Tìm ví dụ hình ảnh để phân đoạn ảnh hiệu quả - Áp dụng thuật toán lên các hình ảnh đó.
 - Phân công công việc – tổng hợp nội dung – trình bày file word báo cáo.
- (6) Đoàn Xuân Khoa:
 - Soạn phần “Kiến thức cơ bản liên quan” cho bài báo cáo (cùng Vĩnh Khang).
 - Đề xuất giải pháp khắc phục hạn chế của thuật toán (cùng Hoàng Khang).

Bảng chấm chéo của các thành viên:

- Thang điểm là thang 4.
- Số hiệu của thành viên là số thứ tự của thành viên trong danh sách nhóm tương ứng
- Mỗi cột ứng với 1 thành viên, thành viên ấy cho điểm cho các thành viên khác trên từng hàng:

	1	2	3	4	5	6
1		4	4	4	4	4
2	4		4	4	4	4
3	4	4		4	4	4
4	4	4	4		4	4
5	4	4	4	4		4
6	4	4	4	4	4	

Lời mở đầu báo cáo

Được sự phân công bài tập 4 của thầy Đặng Nguyên Châu, nhóm 4 gồm 6 thành viên Khoa Điện – Điện tử thực hiện Báo cáo Bài tập lớn 2 – môn học Xác suất & Thống kê để trình bày nội dung mà nhóm được phân công. Bài tập 4 gồm có 2 yêu cầu là Yêu cầu 1 và Yêu cầu 2, trong đó nhóm trình bày lần lượt tại các mục :

- Yêu cầu 1: II.2. Giải quyết yêu cầu 1 – Bài toán dẫn nhập (trang 6)
- Yêu cầu 2: II.3. Giải quyết yêu cầu 2 – Lập trình (trang 7)

Tuy vậy nhóm cảm thấy đây là một đề tài thú vị và cần có thêm các phần giới thiệu lý thuyết và những hạn chế của phương pháp , vì thế bài báo cáo còn có thêm các phần về kiến thức cơ bản , cơ sở lý thuyết, những hạn chế của phương pháp , v.v...

File source của dự án luôn có tại Git repository sau:

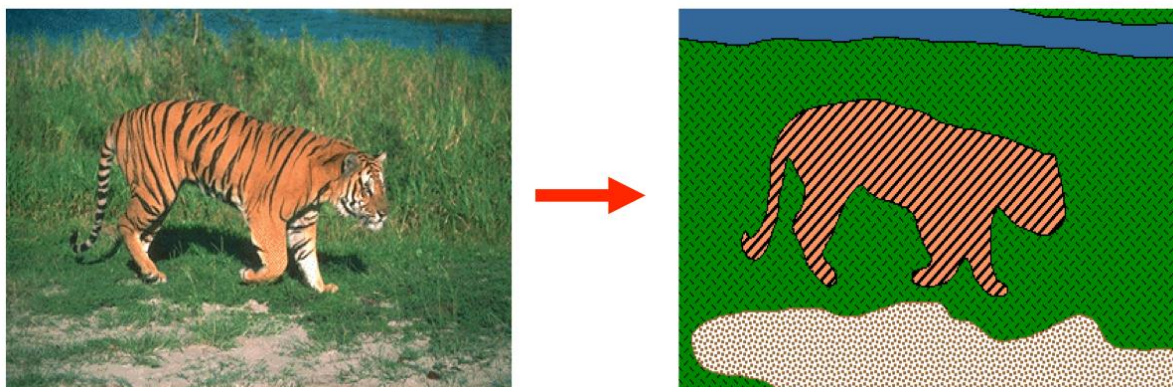
https://github.com/nguyenkhoa2001/Project_Findthreshold_Otsu

Và tiếp sau đây là phần nội dung báo cáo của nhóm.

I. Kiến thức liên quan cơ bản

I.1. Kiến thức về phân đoạn ảnh

Trong lĩnh vực xử lý ảnh và thị giác máy tính, thì phân ngưỡng ảnh là những kỹ thuật được áp dụng để chia một bức ảnh số thành các phần khác nhau. Mục tiêu của việc này là để đơn giản hoá bức ảnh hơn, khiến cho việc phân tích bức ảnh trở nên dễ dàng hơn. Phân ngưỡng ảnh được áp dụng nhiều trong việc xác định các đối tượng trong ảnh, đường biên, v.v....



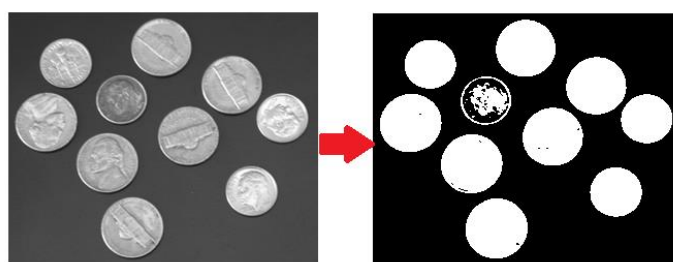
Hình minh họa 1: Kỹ thuật phân ngưỡng ảnh được áp dụng trên hình chia thành 4 đối tượng (Nguồn: stanford.edu)

Ở Hình minh họa 1 ta thấy từ một tấm ảnh có rất nhiều chi tiết và rất phức tạp, áp dụng kỹ thuật phân ngưỡng ảnh cho ta một hình ảnh mà trong đó 4 đối tượng (con hổ, vùng đất, vùng cỏ và vùng nước) được phân chia rất rõ ràng và có đường biên giữa chúng. Sự đơn giản hoá bức hình này có ý nghĩa rất lớn đối với máy tính (nhất là đối với lĩnh vực thị giác máy tính).

I.2. Phân đoạn ảnh xám

Phân đoạn ảnh xám là một trong những phương pháp dễ dàng nhất để phân ngưỡng ảnh xám. Từ một hình ảnh xám (tức có nhiều mức độ trắng đen khác nhau), áp dụng kỹ thuật, ta sẽ thu được một hình ảnh trắng đen (chỉ có mức trắng và mức đen).

Dĩ nhiên, việc phân đoạn ảnh sẽ đưa bức ảnh từ việc có quá nhiều chi tiết phức tạp trở nên đơn giản hơn, có ý nghĩa hơn đối với máy tính.



Hình minh họa 2: Áp dụng kỹ thuật phân đoạn ảnh lên hình đồng xu

Ở Hình minh họa 2 cho ta thấy được việc áp dụng phân đoạn ảnh lên ảnh xám đã đưa ảnh của chúng ta về ảnh nhị phân, và giữa các đồng xu và miếng nền đã có sự phân chia rạch ròi với nhau qua các đường biên (trắng – đen).

Vậy làm thế nào để có thể làm được điều này? Điều chúng ta mong muốn là từ một bức ảnh xám thông thường (có 256 cấp độ xám) trở thành ảnh nhị phân (chỉ có 2 cấp độ, hoặc trắng hoặc đen). Phương pháp phân đoạn ảnh xám đơn giản nhất đó chính là chọn được một mức ngưỡng phù hợp T (với T nằm trong tầm giá trị cấp độ xám), sau đó những pixel nào có độ xám lớn hơn T sẽ được thay hoàn toàn

thành pixel đen , các pixel nào có độ xám nhỏ hơn hoặc bằng T sẽ được thay hoàn toàn thành pixel trắng.

Vậy có thể thấy rằng, ảnh sau khi phân đoạn có được rõ ràng, có thể hiện được ý nghĩa, chính xác hay không phụ thuộc rất lớn vào mức ngưỡng T này. Vì thế bài toán đặt ra phải tìm được một mức ngưỡng T sao cho phù hợp nhất. Đối với đề bài 4 được đặt ra lần này, nhóm 4 sẽ trình bày một kỹ thuật tìm mức ngưỡng T theo phương pháp cực đại phương sai giữa hai nhóm.

II. Tiếp cận – giải quyết yêu cầu bài toán

II.1. Cơ sở lý thuyết

Để tìm được mức ngưỡng T phù hợp, ta sẽ lần lượt quét hết các giá trị T , đối với từng mức ngưỡng T ấy, ảnh sẽ được phân thành hai nhóm: nhóm các pixel có độ xám lớn hơn T (nhóm 1) và nhóm các pixel có độ xám nhỏ hơn hoặc bằng T (nhóm 2). Gọi A, A_1, A_2 lần lượt là trung bình mức xám toàn ảnh, của nhóm 1, nhóm 2. Xác suất của nhóm 1, 2 lần lượt là P_1, P_2 . Từ đó đối với mỗi T , ta có thể tính toán được phương sai giữa hai nhóm theo công thức:

$$\sigma_B^2 = P_1(A_1 - A)^2 + P_2(A_2 - A)^2 \quad (1)$$

mà ta có $A = P_1A_1 + P_2A_2$

Theo đó, đơn giản biểu thức (1) ta thu được:

$$\sigma_B^2 = P_1P_2(A_1 - A_2)^2$$

Cứ ứng với mỗi giá trị T , ta sẽ tìm ra được một giá trị σ_B^2 tương ứng. Mức ngưỡng T phù hợp được chọn để tiến hành phân đoạn ảnh chính là mức T làm cho giá trị σ_B^2 đạt cực đại.

II.2. Giải quyết yêu cầu 1 – Bài toán dẫn nhập

Theo yêu cầu của đề bài BTL, ta có đề bài sau đây:

Cho các mức xám có số lượng pixel tương ứng theo bảng sau đây:

M. Xám	0	1	2	3	4	5	6	7
Số lần	100	150	20	50	200	180	170	300

Giả sử $T = 4$, hãy tìm giá trị σ_B^2 giữa hai nhóm này ?

Giải:

Tổng số lượng pixel : $100 + 150 + 20 + 50 + 200 + 180 + 170 + 300 = 1170$

Với $T = 4$, ta thu được hai nhóm với thông tin sau đây

Nhóm 1:

M. Xám	5	6	7
Số lần	180	170	300

Nhóm 2:

M. Xám	0	1	2	3	4
Số lần	100	150	20	50	200

Ta thu được hai nhóm với P_1, P_2, A_1, A_2 lần lượt có giá trị là

$$P_1 = \frac{180 + 170 + 300}{1170} = \frac{5}{9}$$

$$P_2 = \frac{100 + 150 + 20 + 50 + 200}{1170} = \frac{4}{9}$$

$$A_1 = \frac{5 \times 180 + 6 \times 170 + 7 \times 300}{180 + 170 + 300} \approx 6.185$$

$$A_2 = \frac{0 \times 100 + 1 \times 150 + 2 \times 20 + 3 \times 50 + 4 \times 200}{100 + 150 + 20 + 50 + 200} \approx 2.192$$

Từ các số liệu tính toán ở trên, thu được σ_B^2 tại $T = 4$:

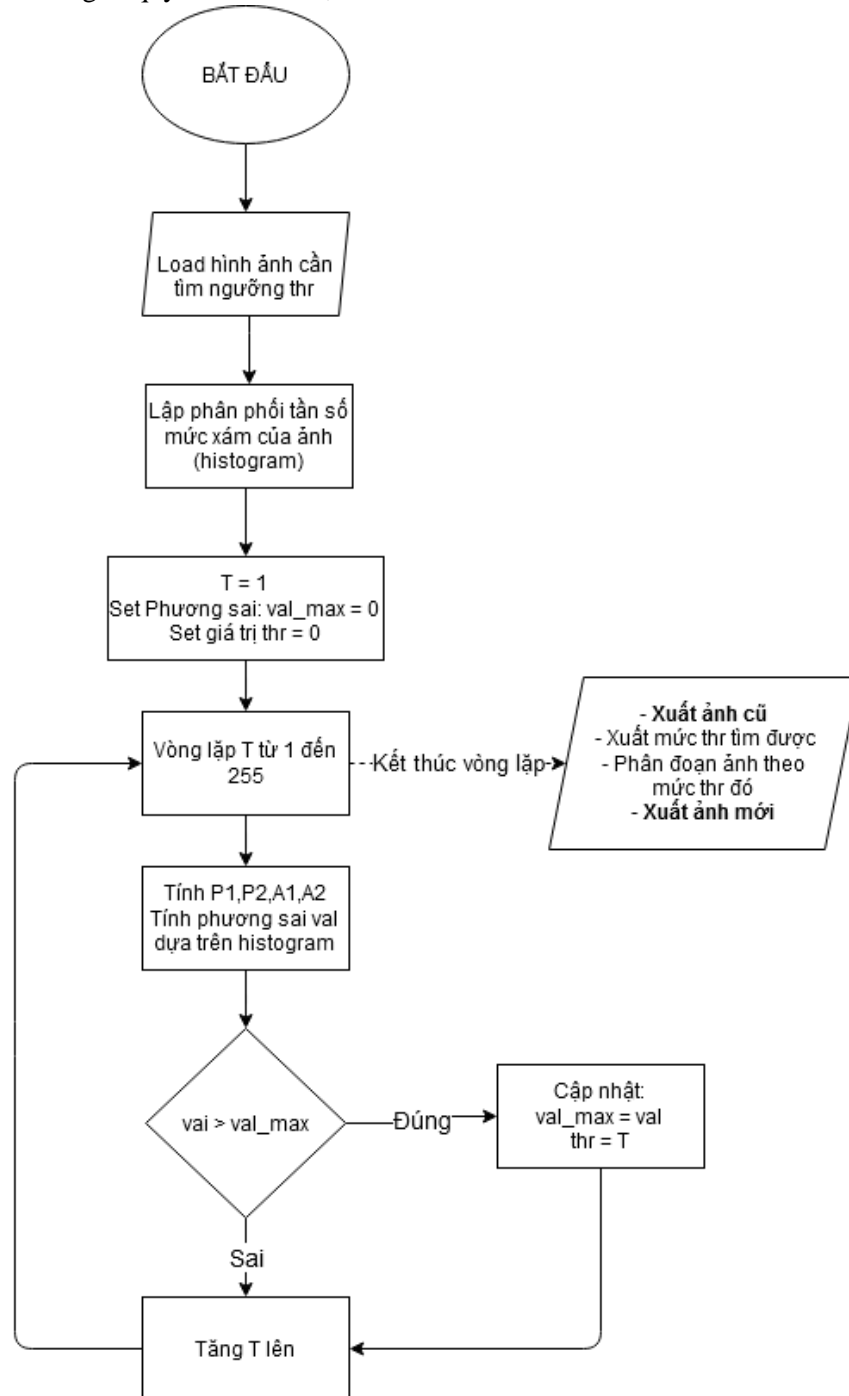
$$\sigma_B^2 = P_1 P_2 (A_1 - A_2)^2 = \frac{5}{9} \cdot \frac{4}{9} \cdot (6.185 - 2.192)^2 \approx 3.935$$

II.3. Giải quyết yêu cầu 2 – Lập trình

Đề bài: Viết một đoạn code (bằng bất kỳ ngôn ngữ lập trình nào) để đọc và hiển thị ảnh xám về nắm như trên. Với một mức ngưỡng T, ảnh nắm cũng sẽ được chia thành 2 nhóm. Tính phương sai giữa hai nhóm với các ngưỡng T khác nhau (T thay đổi từ 1 đến 255). Mức ngưỡng T được chọn là mức ngưỡng làm sai giá trị của phương sai đạt cực đại. Với mức ngưỡng T được chọn, thực hiện việc phân đoạn ảnh và hiển thị kết quả.

Giải:

Lưu đồ thuật toán để giải quyết bài toán đặt ra:



Phần code chương trình (sử dụng ngôn ngữ lập trình Python3)

```
import numpy as np
from cv2 import cv2
from matplotlib import pyplot as plt

im = cv2.imread("mushrooms0_gray.jpg", 0)

[hist, _] = np.histogram(im, bins=256, range=(0, 255))
val_max = 0
thr = 0
for t in range(1, 255):
    if ((np.sum(hist[t:]) == 0) or (np.sum(hist[:t]) == 0)):
        continue
    P1 = np.sum(hist[t:]) / np.sum(hist)
    P2 = np.sum(hist[:t]) / np.sum(hist)
    A1 = np.sum(np.array([i for i in range(t, 256)]) * hist[t:]) / np.sum(hist[t:])
    A2 = np.sum(np.array([i for i in range(t)]) * hist[:t]) / np.sum(hist[:t])
    val = P2 * P1 * np.power(A2-A1, 2)
    if val_max < val:
        val_max = val
        thr = t

print("Muc nguong T: {}".format(thr))

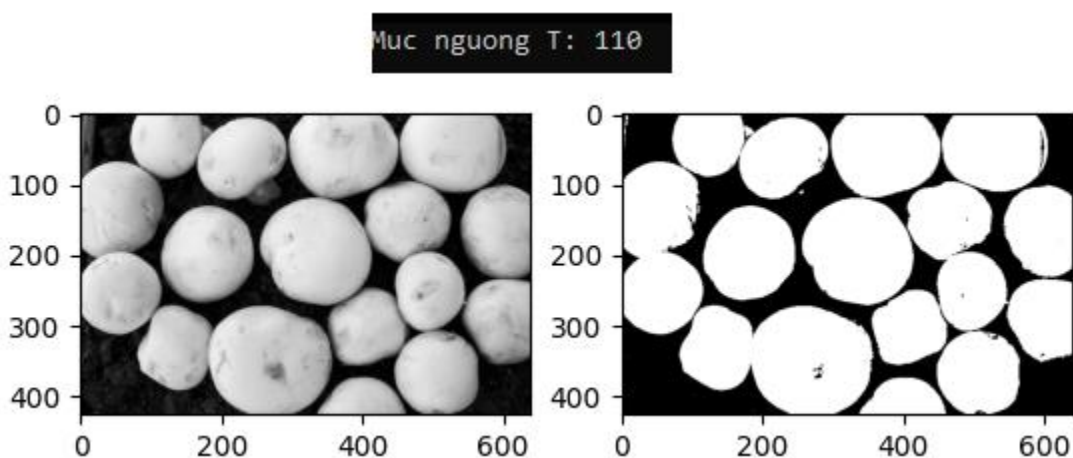
plt.subplot(121)
plt.imshow(im, cmap='gray')

im = im > thr
im = np.uint8(im)

plt.subplot(122)
plt.imshow(im, cmap='gray')

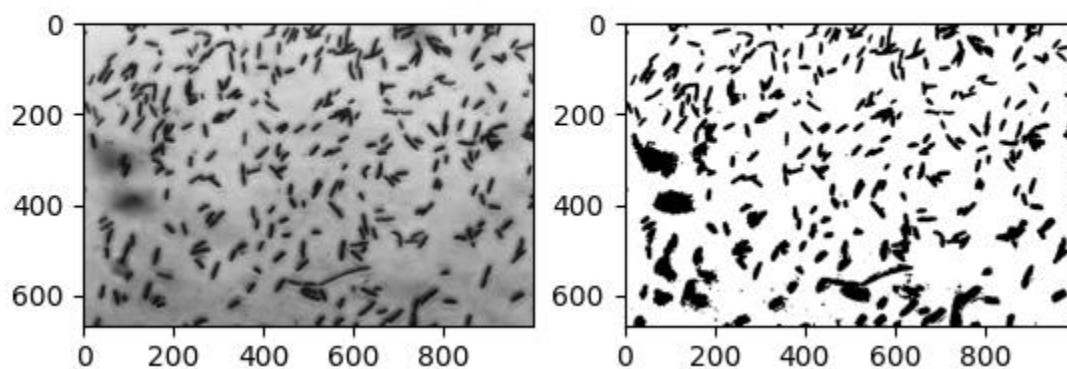
plt.show()
```

Kết quả thu được:

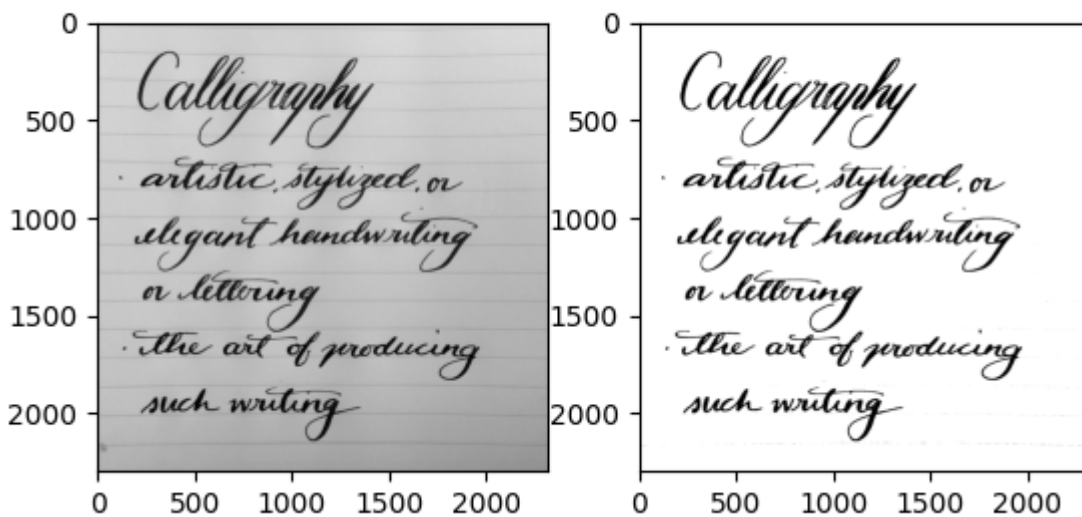


Thử áp dụng đối với một số ảnh xám cần được phân đoạn khác

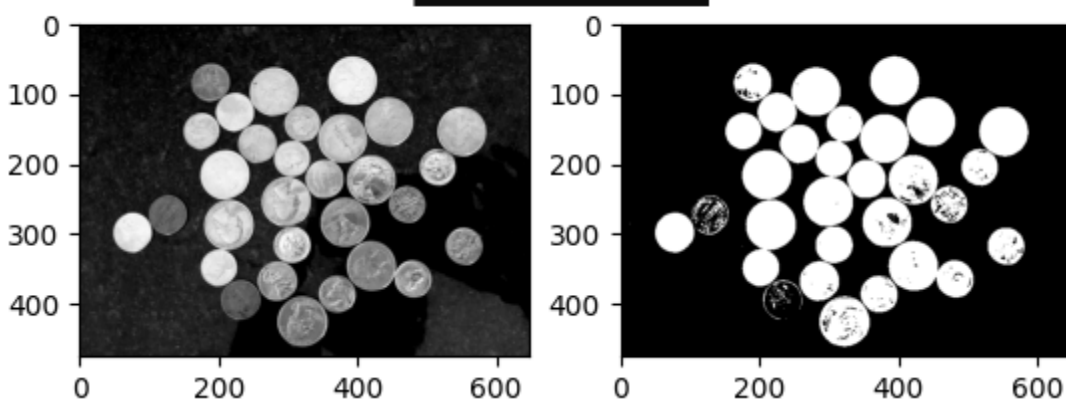
Mức ngưỡng T: 124



Mức ngưỡng T: 110



Mức ngưỡng T: 103



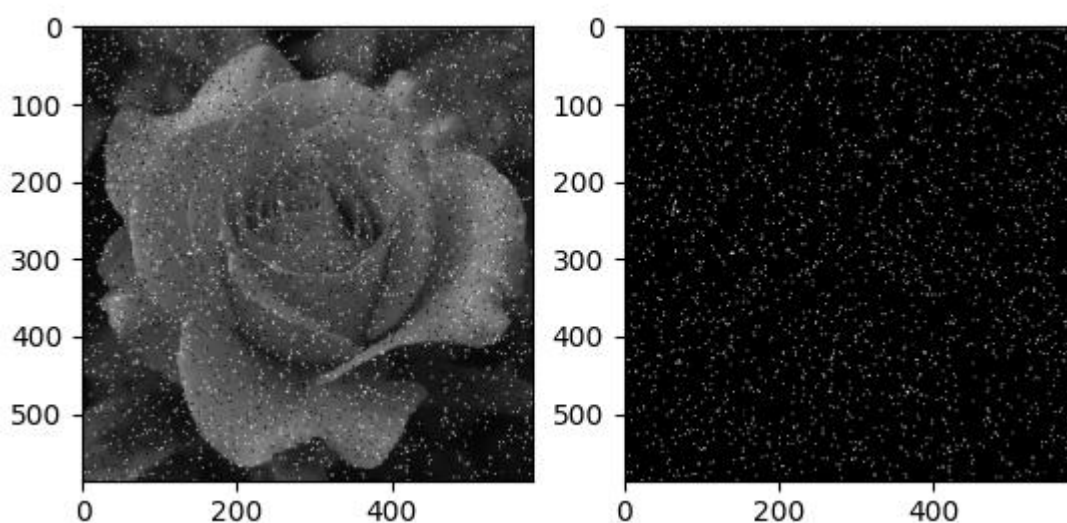
III. Hạn chế của phương pháp – Đề xuất phương án khắc phục

III.1. Hạn chế

Chính từ nền cơ sở lý thuyết của thuật toán này cũng đã làm lộ rõ ra những hạn chế của thuật toán. Nền cơ sở lý thuyết này đòi hỏi cực đại phương sai tính được phải là một giá trị có “ý nghĩa”, tức là phân phối thống kê phải có dạng tương tự phân phối nhị thức. Qua đó, nhìn nhận về góc độ lý thuyết, phương pháp phân ngưỡng ảnh xám dựa trên cực đại phương sai không thực sự hiệu quả khi :

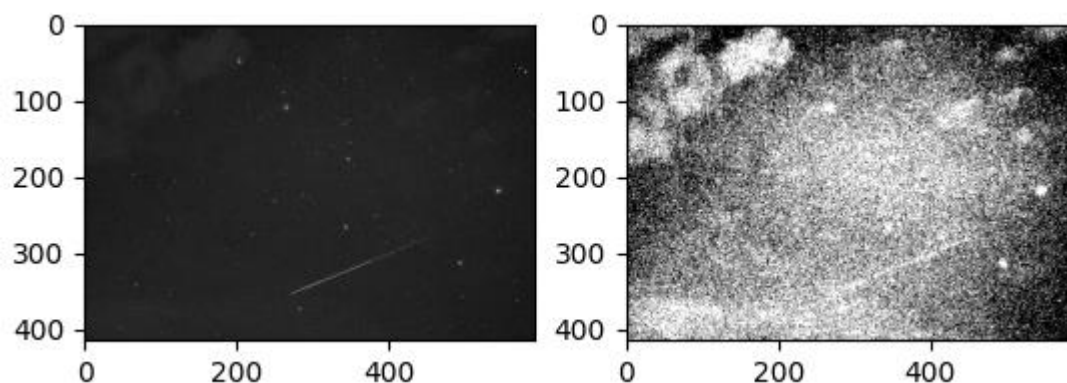
- Có nhiễu trên hình ảnh: nhiễu làm cho các giá trị mức xám của ảnh tương đối gần nhau, tác động trực tiếp lên phân phối mức xám. Lẽ đó khi quét các giá trị T để tính phương sai giữa hai nhóm, các giá trị phương sai sẽ tương tự nhau, làm cho giá trị cực đại phương sai trở nên vô nghĩa về mặt thống kê, việc lấy T từ đó cũng không còn chính xác.

Ví dụ: Ví dụ trên tấm ảnh hoa hồng bị nhiễu rất nặng này, thuật toán đã trả cho chúng ta 1 tấm ảnh trắng đen vô nghĩa, không có bất cứ liên quan gì đến ảnh gốc.



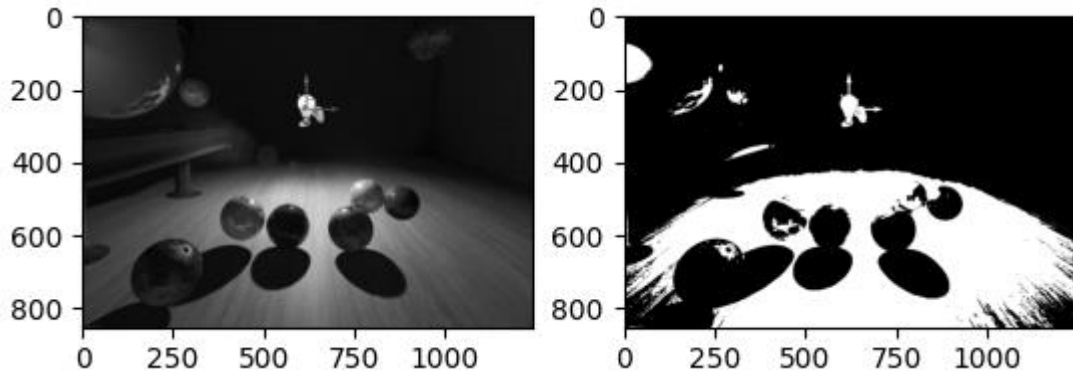
- Vật thể cần tách chiết từ ảnh quá nhỏ bé so với phong nền phức tạp của bức ảnh : điều này cũng làm thay đổi phân phối mức xám của ảnh, mất đi tính cân bằng phân phối nhị thức, khiến mức T được chọn có thể phân ngưỡng sai, mất luôn cả vật thể.

Ví dụ: Hình ảnh chụp sao băng và chòm sao Bắc Giải, nhưng do background của ảnh quá phức tạp, và vật thể được chụp quá nhỏ so với background, nên ảnh sau khi phân đoạn vô nghĩa !



- Vật thể được chụp 3D dưới điều kiện ánh sáng không đồng nhất: điều này gây ra trên bề mặt vật thể tồn tại nhiều mức xám khác nhau, nếu vật thể không đủ lớn so với nền thì thuật toán khả năng cao sẽ xóa đi 1 phần vật thể, làm mất đi ý nghĩa của việc phân đoạn.

Ví dụ: Những trái bóng được dựng 3D nhưng với vị trí chiếu sáng ngược hướng nên gây nên bề mặt vật thể nhiều mức xám, thuật toán đã không thể phân biệt được nên đã làm mất 1 phần các vật thể.



III.2. Khắc phục

Dựa trên từng nhược điểm mà ta sẽ đề ra cách xử lý riêng. Chủ yếu cũng dựa trên cơ sở lý thuyết đưa phân phối mức xám của ảnh về lại mức phân phối nhị thức.

- Có nhiều trên hình ảnh: không chỉ đối với phương pháp cực đại phương sai mà nhiều hầu như tác động xấu đến các phương pháp phân ngưỡng khác. Phương án khắc phục ở đây đề xuất rằng nên áp dụng các bộ lọc mịn như Gaussian trước để giảm nhiễu đến mức hợp lý rồi bắt đầu phân đoạn ảnh.
- Vật thể cần tách chiết từ ảnh quá nhỏ bé so với phong nền phức tạp của bức ảnh : có thể thủ công giảm nhỏ phân vùng nền, giảm nhỏ lại kích thước của background. Hoặc nếu vật thể đặc biệt (sáng nhất, tối nhất) như sao băng trong ví dụ ở trên có thể tăng/giảm mức xám trên toàn bộ pixel ảnh để khử được các yếu tố phức tạp của ảnh (đưa phân phối về lại phân phối nhị thức), rồi mới áp dụng kỹ thuật phân ngưỡng.
- Điều kiện ánh sáng không đồng nhất: ta sẽ quy ước các mức xám có trên vật thể sẽ đi chung với nhau, từ đó T được chọn phải thỏa mãn điều kiện “đi chung” trên , việc này sẽ giảm được tình trạng mất 1 phần vật thể , nhưng hiệu quả giảm đi nhiều.

Có thể thấy những cách khắc phục trên chỉ mang tính tạm chấp nhận, bất khả kháng, một số giải pháp mang tính thủ công quá cao, gây phiền toái, vì thế, trong thực tế, người ta sẽ sử dụng nhiều bộ lọc và nhiều kỹ thuật phân ngưỡng để có thể cho ra được một tấm ảnh được phân ngưỡng tốt nhất.