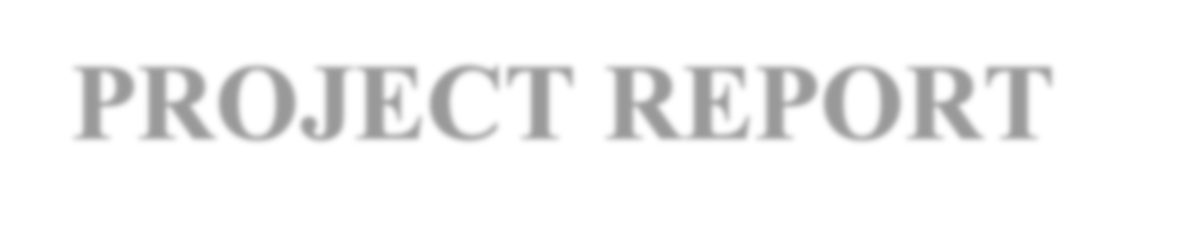
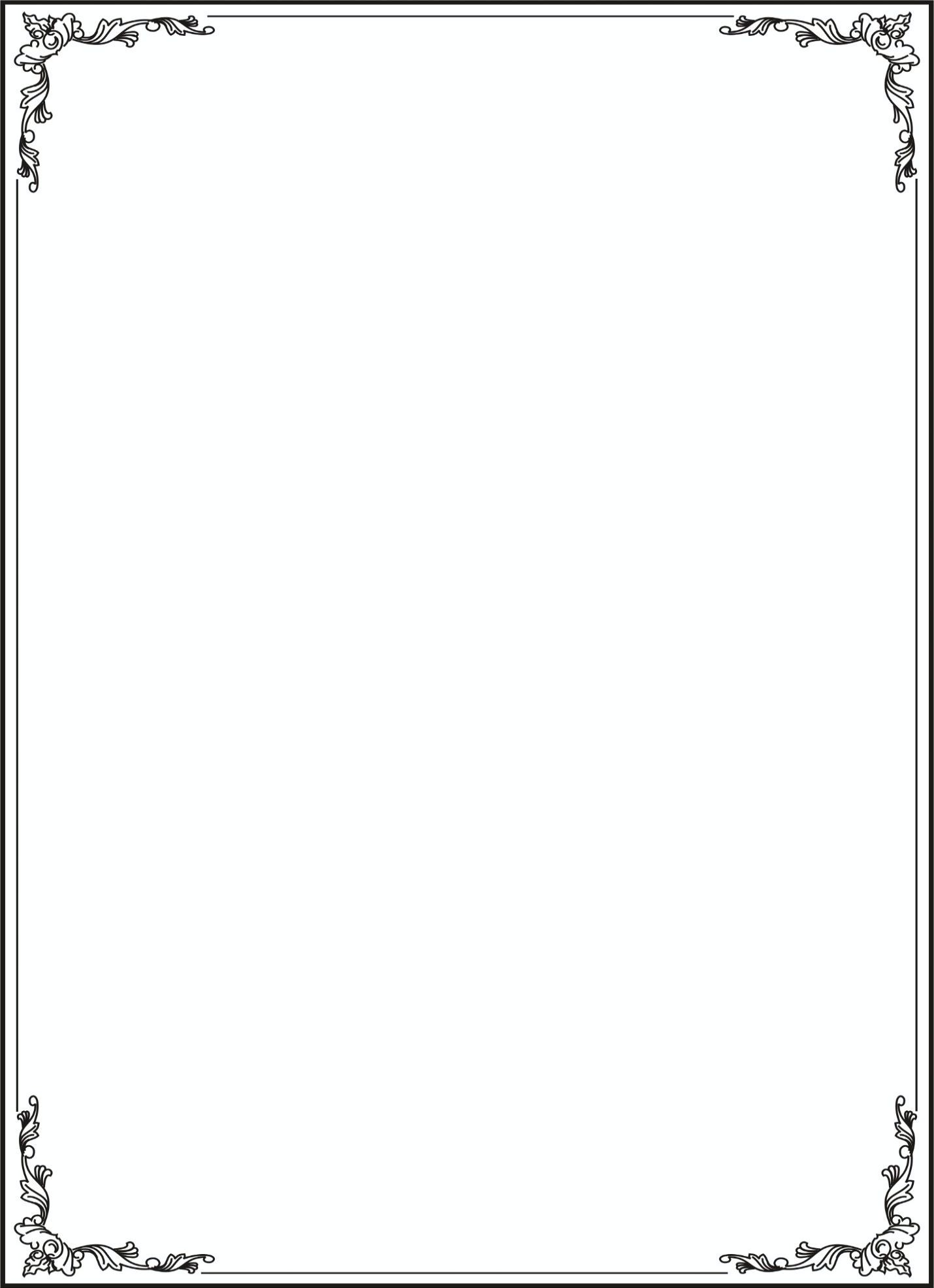
# VIET NAM NATIONAL UNIVERSITY HO CHI MINH CITY

****

**PROJECT REPORT**



**SUBJECT: OBJECT SEGMENTATION**

**Instructor guide:** TS.Mai Tiến Dũng

**Students:** Phan Anh Lộc-19521766

Lê Đình Đức-19521372

**Class:** CS231.M12.KHCL

**Ho Chi Minh City,** 22nd **November, 2021**

Mục lục

[0](#_Toc516576709)

[1. Giới thiệu 2](#_Toc516576710)

[2. Contrast Enhancement using Partial Contrast Stretching. 2](#_Toc516576711)

[3. K-Means Clustering Algorithm. 3](#_Toc516576712)

[4. K-means++. 4](#_Toc516576713)

[5. Proposed Algorithm. 4](#_Toc516576714)

[6. Feature Space 6](#_Toc516576715)

[7. Đánh giá: 6](#_Toc516576716)

[8. Pros & Cons K-means++: 7](#_Toc516576717)

[9. Result. 8](#_Toc516576718)

[10. Ứng Dụng làm mờ ảnh nền để đổi bật nhân vật chính. 9](#_Toc516576719)

Trong Computer Vision, Image segmentation là quá trình phân chia một bức ảnh số thành nhiều phần(set of pixels – cũng được xem là super-pixels). Mục đích của Image segmentation là đơn giản hóa và/hoặc thay đổi cách biễu diễn của bức ảnh theo một cách có ngữ nghĩa hơn và đơn giản hơn để phân tích. Image segmentation thông thường được sử dụng để xác định một object và đường nét quanh object đó (đường thẳng, đường cong,…) trong bức ảnh. Chính xác hơn, image segmentation là quá trình gán nhãn cho mỗi pixel trong bức ảnh, những pixel với cùng một nhãn tương ứng là cùng thuộc một object. Hiện nay đã có nhiều nghiên cứu về việc segmentation một bức ảnh sử dụng kỹ thuật clustering. Trong những nghiên cứu đó đã đề xuất nhiều model, và model được sử dụng phổ biến nhất đó là K-means clustering. K-means clustering là một thuật toán học không giám sát (unsupervised) và nó được sử dụng để phân chia những vùng đặc biệt của bức ảnh.

1. **Giới thiệu:**

Image segmentation là một trong những phương pháp được sử dụng nhiều nhất để phân loại (classify) các pixel của một bức ảnh. Nó chia bức ảnh thành những vùng(region) riêng biệt như là những pixel có giá trị gần giống với nhau thành một vùng. Nó là một công cụ hữu ích trong nhiều lĩnh vực như chăm sóc sức khỏe (health care), image processing, traffic image, patern recognition,… Có nhiều kỹ thuật khác nhau được áp dụng cho việc segmentation một bức ảnh, nhưng một trong những kỹ thuật có hiệu quả, được sử dụng nhiều đó là kỹ thuật Clustering. Và trong đó có nhiều kỹ thuật clustering khác nhau như: K-means clustering, Fuzzy C-means clustering, mountain clustering method và subtractive clustering method.

Một trong số những kỹ thuật clustering được sử dụng nhiều nhất là kỹ thuật K-means clustering. Nó đơn giản và tính toán nhanh hơn. Và nó có thể thực hiện được với số lượng lớn các biến (variable). Nó yêu cầu nhập vào số lượng các nhóm cluster.

### **Contrast Enhancement using Partial Contrast Stretching.**

Trên thực tế, có một số ảnh bị mờ hoặc có độ tương phản thấp, chính vì vậy để cải thiện chất lượng hình ảnh và độ tương phản ta sử dụng kỹ thuật tang cường độ tương phản như **Starial Spatial Starching (PCS).** Bằng cách kéo dài và nén process , nhờ áp dụng kĩ thuật này, phạm vi pixel của giá trị ngưỡng thấp hơn và giá trị ngưỡng trên sẽ ánh xạ tới một phạm vi pixel mới và kéo dài tuyến tính đến phạm vi pixel nơi mà các giá trị kéo dài thấp hơn và các giá trị còn lại bị nén.

### **K-Means Clustering Algorithm.**

**Clustering** là một phương thức để chia một tập hợp dữ liệu thành một số nhóm cụ thể. Trong đó, phương pháp phổ biến là **K-means**. Phương pháp này phân tách tập dữ liệu đã cho thành K cụm. Thuật toán **K-means** gồm 2 giai đoạn riêng biệt:

Giai đoạn 1: Tính toán K *centroids* (trọng tâm).

Giai đoạn 2: Lấy mỗi điểm trong tập dữ liệu đã cho, lấy từng điểm đến cụm có *centroid* gần nhất.

Có nhiều phương pháp khác nhau để xác định khoảng cách của centroid gần nhất và một trong những phương pháp được sử dụng nhiều nhất là Euclidean distance. Khi nhóm được thực hiện xong, ta tính toán lại *centroid* mới của mỗi *cluster,* dựa trên *centroid* đó, một khoảng cách Euclidean mới được tính toán giữa mổi trung tâm và mổi điểm dữ liệu và gán các điểm trong *cluster* có khoảng cách Euclidean tối thiểu.

Mổi *cluster* trong phân vùng được xác định bởi *centroid* và các đổi tượng thành viên của nó.Trong đó, c*entroid* của mổi *cluster* là điểm mà tổng khoảng cách từ nó tới tất cá các đối tượng thành viên trong cụm đó nhỏ nhất. Chính vì vậy K-means là một thuật toán lặp đi lặp lại, trong đó nó giảm thiểu tổng khoảng cách từ mỗi đối tượng đến trung tâm cụm của nó, trên tất cả các *cluster*.

Xét một hình ảnh X x Y và hình ảnh phải nhóm thành K *cluster.* Cho P(X,Y) là một input pixels và ck là trung tâm cụm. Thuật toán **K-means clustering** như sau:

Bước 1: Khởi tạo số cụm k và trung tâm

Bước 2: Đổi với mổi pixel của hình ảnh, tính khoảng cách Euclidean d giữa trung tâm và mổi pixel của ảnh sử dụng mổi quan hệ dưới đây:



Bước 3: Gán tất cả các điểm ảnh cho trung tâm gần nhất dựa trên khoảng cách d vừa tình được ở trên

Bước 4: Sau khi tất cả các điểm ảnh đã được gán, tính lại vị trí mới của trung tâm bằng cách sử dụng mối quan hệ sau:



Bước 5: Lặp lại quy trình cho đến khi nó thỏa mãn đúng sai

Bước 6: Định hình lại cluster pixels vào hình ảnh ban đầu.

Mặc dù **K-means** có lợi thế lớn vì thực hiện dễ dàng nhưng nó có một số nhược điểm. Chất lượng của kết quả phân cụm cuối cùng phụ thuộc vào lựa chọn những *centroids* ban đầu. Ví vậy, nếu *centroids* ban đầu là ngẫu nhiên được chọn, nó sẽ nhận được kết quả khác nhau cho các trung tâm ban đầu khúc nhau. Một nhược điểm của thuật toán K-mean là nó nhạy cảm với việc khởi tạo các trọng tâm hoặc các điểm trung bình. Vì vậy, nếu một centroid được khởi tạo là một điểm “xa”, nó có thể chỉ kết thúc không có điểm nào được liên kết với nó và đồng thời, nhiều hơn một cụm có thể kết thúc với một centroid duy nhất. Tương tự, nhiều hơn một trung tâm có thể được khởi tạo vào cùng một cụm dẫn đến phân cụm kém. Ví dụ, hãy xem xét các hình ảnh hiển thị bên dưới. Chính vì vậy muốn được phân tách mong muốn, chúng ta nên lựa chọn *centroids* ban đầu một cách cẩn thận. Và phức tạp trong tính toán cũng là một vấn đề chúng ta cần cân nắc khi thiết kế cụm K-means. Nó dựa vào số lượng các phần tử dữ liệu, số lượng cụm và số lần lặp lại.

### **K-means++.**

K-means được sử dụng bởi tốc độ và sự đơn giản của nó nhưng K-means theo thực nghiệm, có rất nhiều thực nghiệm tạo ra bad clusterings bởi sự lựa chọn những *centroids* ban đầu thiếu cẩn thận. Chính vì vậy, K-means++ ra đời để cải thiện K-means.

Thuật toán K-means++ đề xuất một biến thể chọn các *centroids* ngẫu nhiên nhưng cân nhắc các điểm dữ liệu theo khoảng cách bình phương của chúng bình phương từ trung tâm gần nhất đã được chọn trước đó.

Thuật Thoán K-means++ như sau :

* Bước 1: Lấy một *centroid* ban đầu c1, được chọn ngẫu nhiên từ tập dữ liệu X.
* Bước 2: Lấy một *centroid* mới ci, chọn x X , với xác suất: 

*( D(x) biểu hiện khoản cách ngắn nhất từ điểm dữ liệu x cho đến trung tâm đã chọn ).*

* Bước 3: Lặp lại bước 2 cho đến khi tìm được K *centroid*.
* Bước 4: Tiến hành như K-means đã trình bày ở trên(Mục 3).

Trực giác:

* Bằng cách thực hiện theo quy trình khởi tạo ở trên, chúng ta nhận ra các centroid cách xa nhau. Điều này làm tăng cơ hội ban đầu nhặt được các trung tâm nằm trong các cụm khác nhau. Ngoài ra, vì các centroid được chọn từ các điểm dữ liệu, mỗi centroid có một số điểm dữ liệu được liên kết với nó ở cuối.

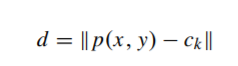
### **Proposed Algorithm.**

Thuật toán được đề xuất bao gồm: **partial contrast stretching, k-means clustering++.**

Trên thực tế vẫn có một số hình ảnh có độ tương phản thấp, chính vì vậy **contrast stretching** được sử dụng để cải thiện chất jg hình ảnh. Sau khi cải thiện chất lượng hình ảnh. Số lượng trung tâm được tạo ra dựa trên số cụm K ta mong muốn man đầu**.** Sử dụng thuật toán **K-means++**, hình ảnh được phân đoàn thành K cụm. Thuật toán này được đề xuất như sau:

* Bước 1: Tải hình ảnh để phân đoạn.
* Bước 2: Áp dụng **partial contrast stretching** và khởi tạo số cụm K.
* Bước 3: Lấy một *centroid* ban đầu c1, được chọn ngẫu nhiên từ tập dữ liệu X.
* Bước 4:Chọn x X làm một *centroid* mới ci với xác suất: 

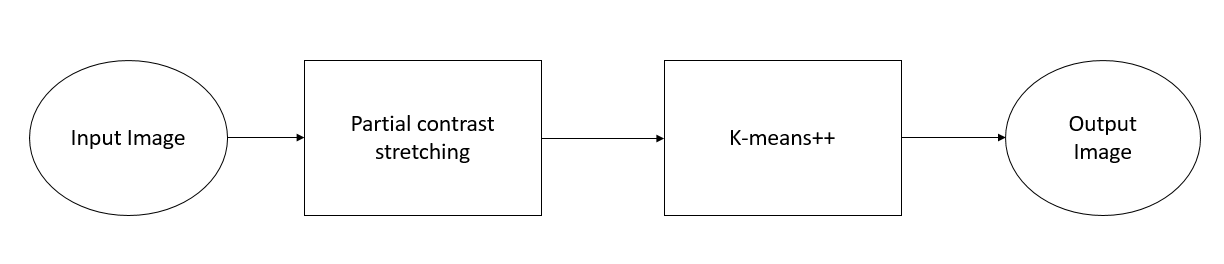
*( D(x) biểu hiện khoản cách ngắn nhất từ điểm dữ liệu x cho đến trung tâm đã chọn ).*

* Bước 5: Lặp lại bước 4 cho đến khi tìm được K *centroids* ban đầu .
* Bước 6: Tìm khoảng cách Euclidean của mổi *centroid* từ mọi điểm ảnh của ảnh bằng quan hệ sau:
* 
* Bước 6: Gôm điểm ảnh có khoảng cách tối thiểu với *centroid* đến cụm tương ứng với *centroid* của nó.
* Bước 7: Tính toán lại vị trí trung tâm mới bằng cách sử dụng phương trình:



* Bước 8: Lặp lại các bước 6-7 cho đến khi nó thỏa mãn dung sai.
* Bước 9: Định hình lại cụm vào hình ảnh.

***Sơ Đồ Khối***:



1. **Feature Space:**

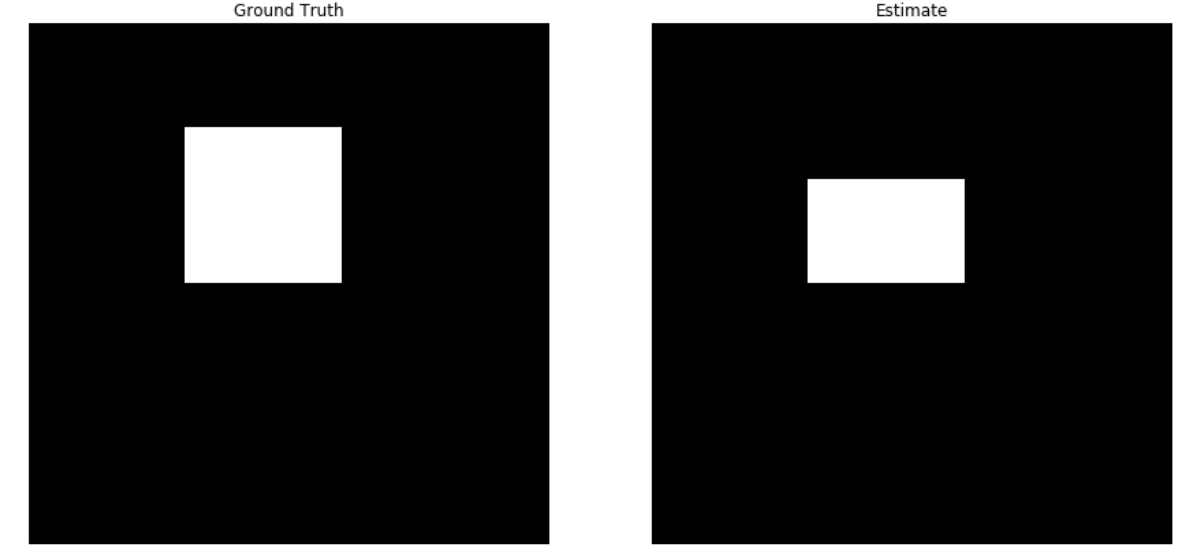
Để ứng dụng được thuật toán K-means++ Clustering chúng ta cần biễu diễn những pixel trên ảnh thành các data points. Tập hợp các data points này là feature Space. Chúng ta có thể chọn nhiều cách khác nhau để biễu diễn.

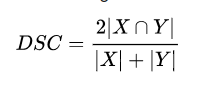
Biễu diễn các pixel dựa trên cường độ màu sắc của pixel hoặc cường độ kết hợp với vị trí,…

### **Đánh giá:**

Nhìn vào một hình ảnh đã được segments luôn là cách đánh giá hiệu quả để xem thuật toán chạy chính xác như thế nào, nhưng cách tốt nhất để đánh giá một thuật toán là cần một đại lượng đo lường cho độ chính xác của thuật toán đó.

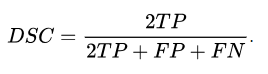
Dice coefficient là một kỹ thuật đánh giá dựa trên sự so sánh 2 mẫu. Để đánh giá hình ảnh đã segmentation của chúng ta – là hình ảnh được segment một cách tự động, có độ chính xác bao nhiêu chúng ta cần so sánh nó với hình ảnh ground truth – là ảnh đã segment một cách chính xác nhất do người dùng thực hiện thủ công.





Trong đó |X| , |Y| là tương ứng là số lượng các phần tử trong tập X và Y.

Chúng ta áp dụng lên trên dạng dữ liệu Boolean, sử dụng các định nghĩa true positive (TP), false positive(FP) và false negative (FN), chúng ta có công thức:



Dựa vào lí thuyết trên, chúng ta có thể đánh giá xem kết quả của thuật toán chính xác như thế nào. Chúng ta có 2 ảnh gồm: ảnh ground truth và ảnh được segment bằng thuật toán K-means++ clustering. Vì chúng ta có nhiều cụm k, nên sẽ lấy từng cụm k so sánh với ảnh ground truth và cụm k cho độ chính xác cao nhất. chúng ta sẽ chia bức ảnh thành 2 phần foreground (positive) và background(negative).

TP: là phần giao nhau giữa phần foreground của ảnh ground truth và phần foreground của ảnh được segment

FP: là phần foreground của ảnh ground truth không trùng với phần foreground của ảnh được segment

FN: là phần backgournd của ảnh ground truth không trùng với phần background của ảnh được segment

### **Pros & Cons K-means++:**

**Pros**:

-Đơn giản và dễ dàng để triển khai

-Nhanh với những dữ liệu có số chiều nhỏ

-Dễ dàng biểu diễn dữ liệu

**Cons**:

-Không có nhận biết những outliers

-Cần phải xác định số lượng clusters k

-Giới hạn trên dữ liệu mà có khái niệm theo các cụm có các trung tâm

1. **Result.**

**Test 1:**

**Ảnh Gốc Khi Phân Đoạn(k=2)**

**Test 2:**

### **Ứng Dụng làm mờ ảnh nền để đổi bật nhân vật chính.**

**Test 1:**

Ảnh gốc Ảnh Áp Dụng

**Test 2:**

**Ảnh Gốc Ảnh Áp Dụng** 

Tài liệu tham khảo

<http://ilpubs.stanford.edu:8090/778/1/2006-13.pdf>

<https://www.sciencedirect.com/science/article/pii/S1877050915014143>

<https://en.wikipedia.org/wiki/Image_segmentation>

<https://docs.opencv.org/3.0-beta/doc/py_tutorials/py_imgproc/py_morphological_ops/py_morphological_ops.html>

<https://en.wikipedia.org/wiki/S%C3%B8rensen%E2%80%93Dice_coefficient>