

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN
KHOA KHOA HỌC MÁY TÍNH



BÁO CÁO ĐỒ ÁN CUỐI KỲ

**MÔN HỌC: CS231.M22.KHCL – NHẬP MÔN THỊ GIÁC MÁY
TÍNH**

ĐỀ TÀI:

CHROMAKEY USING LOGISTIC REGRESSION

GIẢNG VIÊN HƯỚNG DẪN: MAI TIẾN DŨNG

SINH VIÊN THỰC HIỆN: LỮ THỊ THÚY QUỲNH - 20521826

BÙI VĂN THUẬN - 20521990

TP. Hồ Chí Minh, ngày 24 tháng 6 năm 2022

MỤC LỤC

1. Giới thiệu	3
2. Ý tưởng	4
2.1. Phương pháp đầu tiên thực hiện:	4
2.1.1. Nhận xét	5
2.1.2. Lý do không chọn phương pháp này làm chủ đề:	6
2.2. Phương pháp thứ hai thực hiện Logistic Regression:	6
2.2.1. Lý do chọn Logistic Regression	6
2.2.2. Chi tiết về ý tưởng:	6
3. Quá trình thực hiện	7
3.1. Bộ dữ liệu để train:	7
3.2. Train mô hình	8
3.3. Hiện thị kết quả	8
3.4. Thực nghiệm để chọn threshold	9
4. Kết quả	13
5. Kết luận	17
TÀI LIỆU THAM KHẢO	19

BÁO CÁO ĐỒ ÁN

1. Giới thiệu

Trước hết **Chroma Key** là một kỹ thuật hiệu ứng hình ảnh và hậu kỳ để kết hợp (xếp lớp) hai hình ảnh hoặc luồng video với nhau dựa trên sắc độ màu (phạm vi sắc độ). Kỹ thuật này đã được sử dụng trong nhiều lĩnh vực để xóa phông nền khỏi chủ đề của ảnh hoặc video - đặc biệt là các ngành công nghiệp truyền hình, ảnh chuyển động và trò chơi điện tử.

Chromaticity hay còn gọi là *Chroma*, cho chúng ta biết được độ 'tinh khiết' của một tông màu. Đặc tính này sẽ được xác định dựa vào **mức độ hiện diện của màu trắng, xám và đen trong một màu**. 12 tông màu cơ bản được nhắc tới ở trên có mức Chroma cao nhất do không có lẫn thêm màu nào khác cả. Những màu với mức Chroma cao thường khá đậm và trông rất sống động.

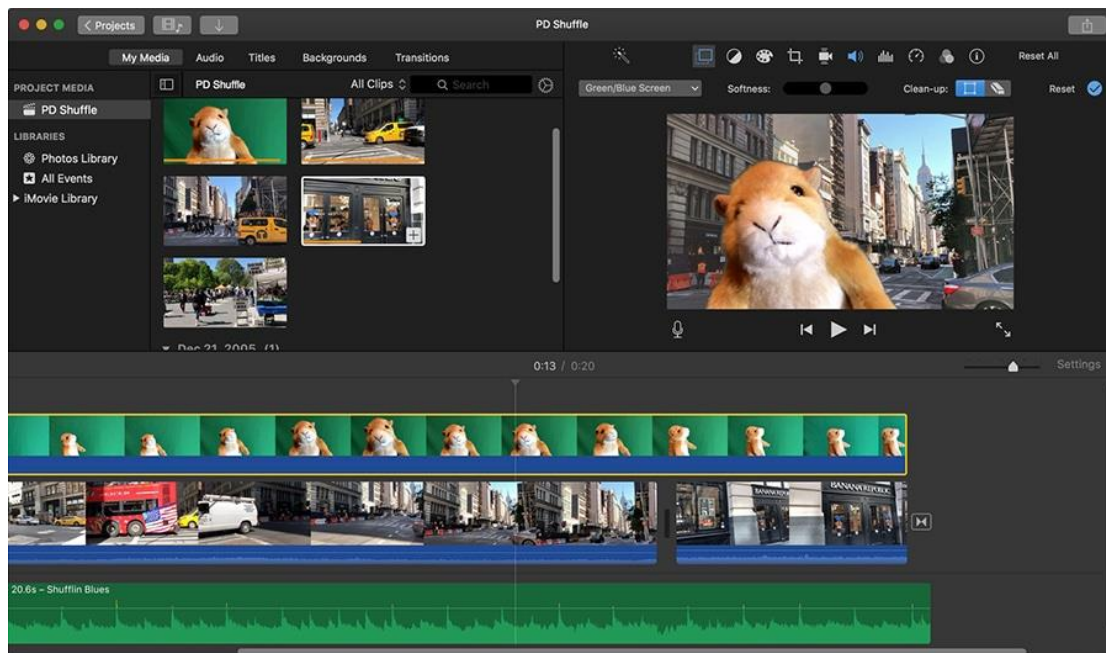
Keying: Thuật ngữ này được sử dụng để mô tả quá trình xóa nền màn hình xanh trong hậu kỳ bằng phần mềm chỉnh sửa video. Khi nền màn hình xanh đã được khóa, nó sẽ hoàn toàn trong suốt. Sau đó bạn có thể lấp khu vực trong suốt đó bằng một hình ảnh hoặc video khác. Mục tiêu là để có thể khiến khu vực trong suốt này trong suốt nhất có thể, có nghĩa là toàn bộ những gì màu xanh (màu nền) ta sẽ xóa chúng đi một cách sạch sẽ.



Chroma Key: Đó là kỹ thuật xếp lớp thực tế (layering), toàn bộ quá trình quay phim -> xóa phông nền -> làm mịn phông nền xử lý sự cố tràn... -> ghép phông nền mới -> tất cả quá trình đó gọi là kỹ thuật **Chroma Key**.

Chroma key sẽ dùng nền màu xanh lá cây hoặc xanh dương. Vì màu xanh lá cây và màu xanh dương sẽ không trùng với màu tóc màu da của người trong ảnh video. Để tránh trường hợp ảnh ghép ra bị mất một vài phần thì người ta sẽ tránh mặt quần áo phụ kiện trùng màu với phông nền.

Vì sự đơn giản và tính tiện dụng nên kỹ thuật được áp dụng vào gần như các bộ phim điện ảnh thời nay. Có luôn các máy quay hỗ trợ tách, ghép nền trực tiếp cũng như các phần mềm chỉnh sửa hậu kỳ cho hiệu ứng này.



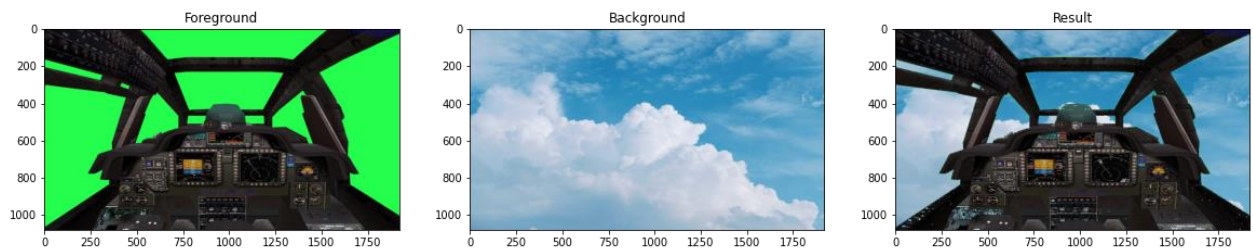
2. Ý tưởng

Trước khi nói về việc hình thành nên ý tưởng của Chroma key sử dụng Logistic Regression tại em sẽ nói qua về phương pháp đầu tiên mà tại em đã làm trước khi dùng Logistic Regression.

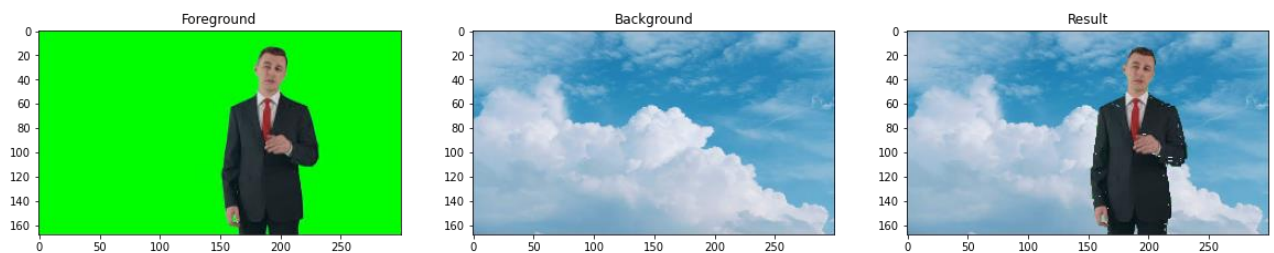
2.1. Phương pháp đầu tiên thực hiện:

Cách đơn giản nhất để thực hiện là xác định giá trị kênh màu RGB của nền xanh và tiến hành tách nền và đối tượng bằng cách phân loại từng pixel xem giá trị RGB của nó có thuộc trong khoảng RGB của nền xanh hay không nếu bằng điểm ảnh đó thuộc về nền, còn không thuộc về đối tượng. Phân loại đối tượng và nền rồi tạo mask để lấy

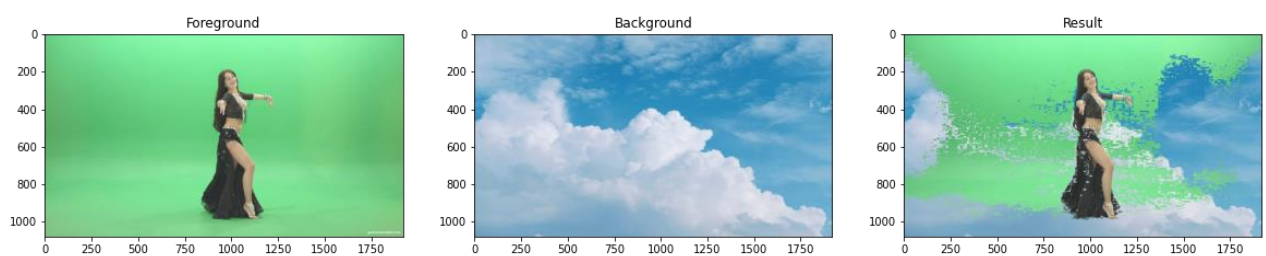
được phần đối tượng cần có. Sau đó ghép phần đối tượng mình đã tách ra được từ mask kép hợp với ảnh nền ta sẽ có ảnh mong muốn.



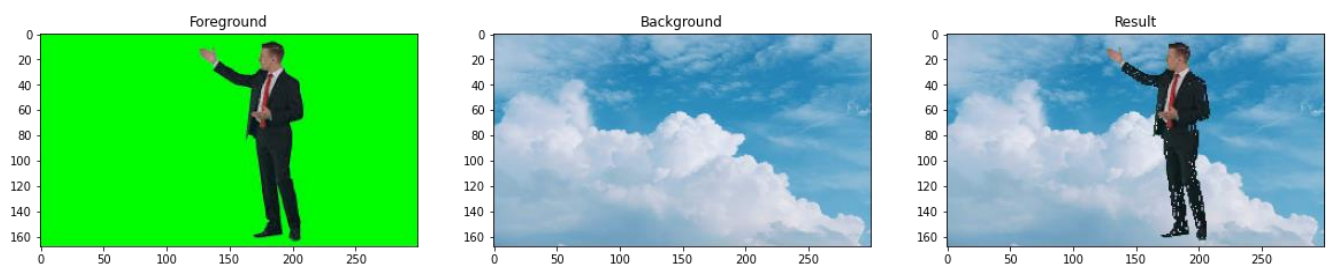
Hình 1. 1



Hình 1. 2



Hình 1. 3



Hình 1. 4

2.1.1. Nhận xét

Tụi em đã thực hiện cách này và cho ra kết quả đối tượng đã được tách khỏi nền khá rõ ràng như hình 1.1 và hình 1.2.

Tuy nhiên cũng có hình tách nền và đối tượng chưa được tốt như hình 1.3 do phông nền đằng sau dù màu có màu xanh nhưng không được làm mịn, chỗ sáng chỗ tối do hình ảnh này chưa đủ tiêu chuẩn của chroma key giống như những hình trên nên điều này cũng khó tránh được.

Riêng hình 1.4 thì đối tượng đã được tách nền và không còn phần nào của nền xanh tuy nhiên lại có những khoảng trắng trên đối tượng. Có nghĩa là một số điểm ảnh của đối tượng đã bị nhầm với nền nên cũng bị loại bỏ.

Phương pháp khá đơn giản dễ thực hiện và cho ra kết quả cũng chấp nhận được với những hình ảnh đủ tiêu chuẩn

2.1.2. Lý do không chọn phương pháp này làm chủ đề:

Vì cách này khá đơn giản không cần nghiên cứu và thực nghiệm nhiều. Với nhóm 2 người thì như vậy hơi ít việc để làm và nội dung để trình bày nên tụi em vẫn muốn thử một cách khác mới hơn để xem kết quả sẽ như thế nào, có cải thiện được những khuyết điểm của phương pháp trên không.

2.2. Phương pháp thứ hai thực hiện Logistic Regression:

2.2.1. Lý do chọn Logistic Regression

Tụi em nghiên cứu và đọc nhiều bài báo để xem cách người ta hay dùng để làm chroma key. Thì từ bài báo [FPGA Implementation for Real-time Chroma-key Effect Using Coarse and Fine Filter](#) trong phần II. Study of chroma key method thì tụi em tìm kiếm được ý tưởng chung của đa số các phương pháp Chroma key sẽ gán những pixel thuộc đối tượng là 1 và nền là 0. Sau khi tạo mask thì hình có được sẽ là ảnh kết hợp từ ảnh ban đầu đã tạo mask và phông nền muốn ghép đã tạo mask.

$$C = Fg.*mask + Bg.*(\sim mask)$$

Dựa trên điều này và qua sự thảo luận của nhóm thì tụi em quyết định chọn mô hình Logistic Regression để phân loại nền và ảnh.

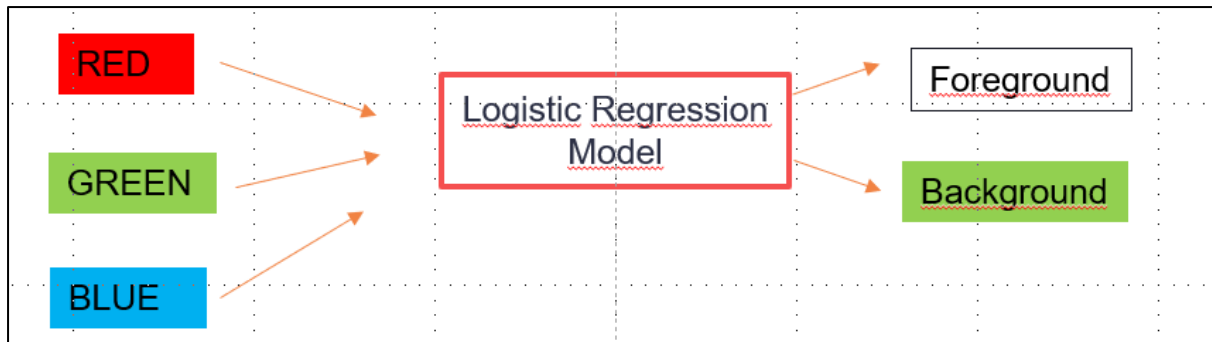
Bởi vì Logistic Regression là 1 thuật toán phân loại được dùng để gán các đối tượng cho 1 tập hợp giá trị rời rạc (như 0, 1, 2, ...). và thường được sử dụng cho bài toán phân lớp nhị phân

→ Chọn Logistic Regression làm phương pháp dùng để phân lớp background và foreground

2.2.2. Chi tiết về ý tưởng:

Nhận xét: Vì hình ảnh sử dụng hiệu ứng chroma key đã có nền màu xanh sẵn tách biệt với màu của đối tượng ở trong nên ta có thể dễ dàng dựa vào đặc trưng về màu sắc để phân tách nền và đối tượng.

Chọn kênh màu của ảnh làm đặc trưng để phân lớp.



Chúng ta cần tìm ra bộ trọng số thỏa hàm dưới đây:

$$y = W_1 \times R + W_2 \times G + W_3 \times B + W_0 \times 1 \quad (1)$$

$$\text{mask} = s(y) = \frac{1}{1 + e^{-y}} \quad (2)$$

Mô hình Logistic Regression mình sẽ train bằng những đặc trưng đã rút trích được từ bộ data. Sau quá trình train thì mình đã tìm kiếm được bộ tham số phù hợp, khi đó muốn ghép được ảnh mong muốn ta cần lấy kênh màu của của ảnh đó đưa qua phép biến đổi (1) và (2) ta sẽ có được 1 ma trận mask có kích thước chiều cao, chiều rộng bằng với ảnh chroma key. Ma trận mask có nhiệm vụ lưu lại xác suất thuộc lớp nền của từng pixel, tương ứng với vị trí của nó trong ma trận ảnh ban đầu.

Thông thường ngưỡng xác suất để phân một đối tượng thuộc lớp nào trong Logistic Regression là 0.5. Tuy nhiên ở đây tụi em không sử dụng ngưỡng vì kết quả cho ra không tốt. Phần này tụi em sẽ nói trình bày chi tiết ở quá trình thực hiện

3. Quá trình thực hiện

3.1. Bộ dữ liệu để train:

Tập dữ liệu để huấn luyện sẽ gồm hai thư mục một là những ảnh có phần foreground và thư mục thứ hai gồm những bức ảnh chỉ chứa phông nền.

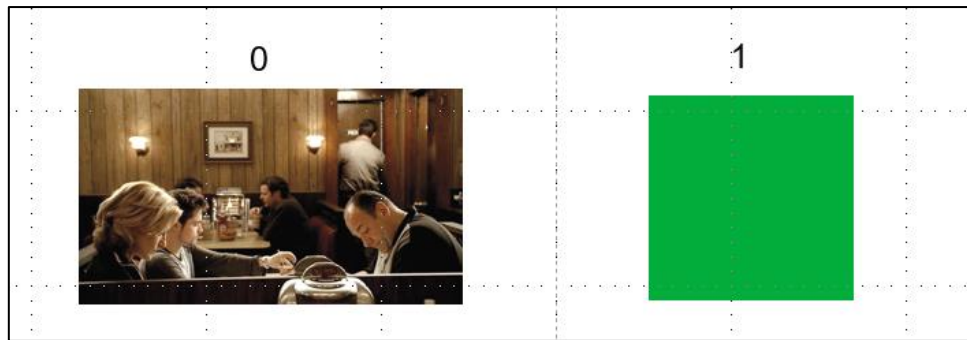
Tập foreground: Những hình ảnh ở trên mạng tụi em tự chọn. Tụi em sẽ chọn những tấm không có màu xanh lá để khi phân lớp thì mô hình sẽ không học nhầm màu của nền.

Thầy có thể xem kĩ hơn trên thư mục tụi em tải lên.

Tập background: Thì những hình ảnh màu xanh lá có sắc độ khác nhau

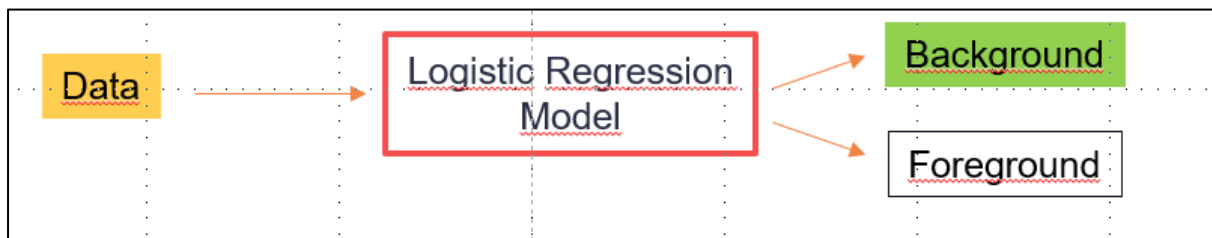
Trên từng tấm ảnh sẽ có thứ tự của ảnh và nhãn. VD: 1-1, 1-0,..

Nhãn 0 tương ứng với đối tượng, nhãn 1 là phần nền.



Tập foreground và background mỗi tập sẽ có 5 hình.

Về phần rút trích đặt trưng từ tập dữ liệu đã trình bày ở phần trên. Sau khi rút trích đặc trưng thì data của mình sẽ là một list gồm có bộ kênh màu của từng pixel và nhãn tương ứng của chúng. Ở đây nói đơn giản là ta sẽ cho máy học về bộ kênh màu thuộc nền và bộ kênh màu thuộc đối tượng.



3.2. Train mô hình

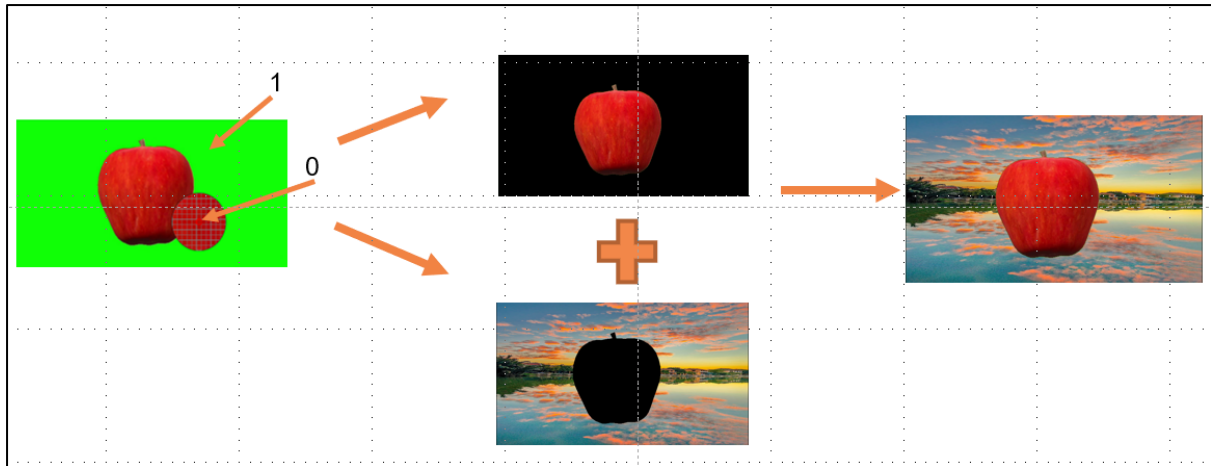
Dùng `train_test_split` của thư viện `sklearn` để tách ra tập train và tập test. Test size= 0.25 như mặc định

Dùng mô hình Logistic Regression của thư viện `sklearn` để train mô hình

Ở đây tham số của mô hình tụi em tụi em không thay đổi gì so với mặc định vì kết quả khi train đã đạt được 1.0. Do dữ liệu của tụi em khá ít nên việc đạt kết quả cao thế này cũng là điều dễ hiểu.

3.3. Hiện thị kết quả

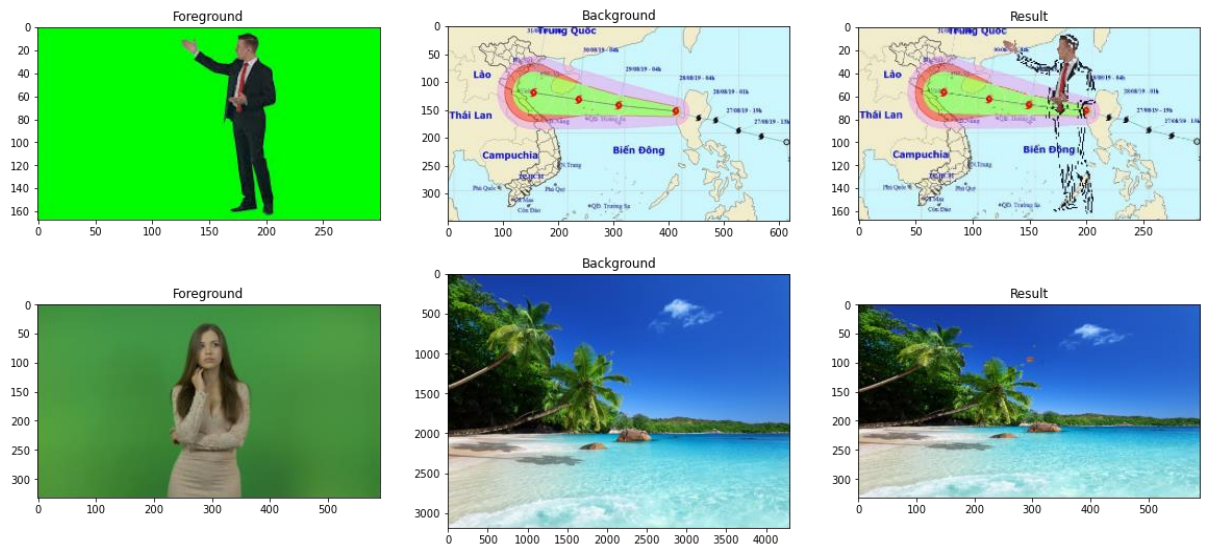
Về phần áp dụng bộ tham số của mô hình đã train được để tạo mask thì cũng giống như trong phần ý tưởng.



Hình ảnh được ghép sẽ tạo từ mask của hình ảnh phong xanh ban đầu kết hợp với mask của phần background muốn ghép.

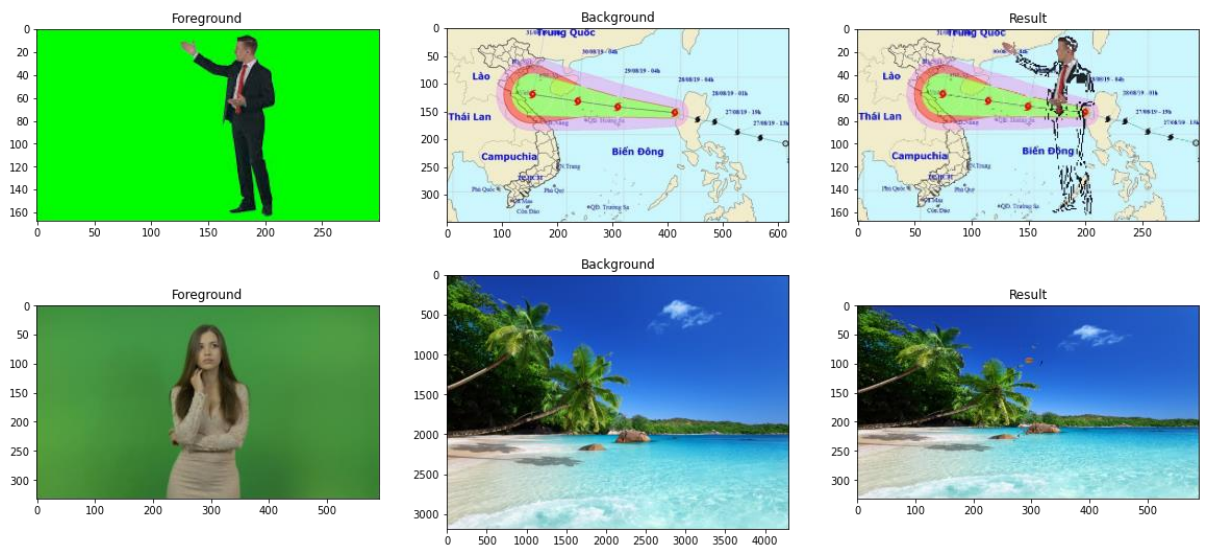
3.4. Thử nghiệm để chọn threshold

1. *Threshold* = 0.5



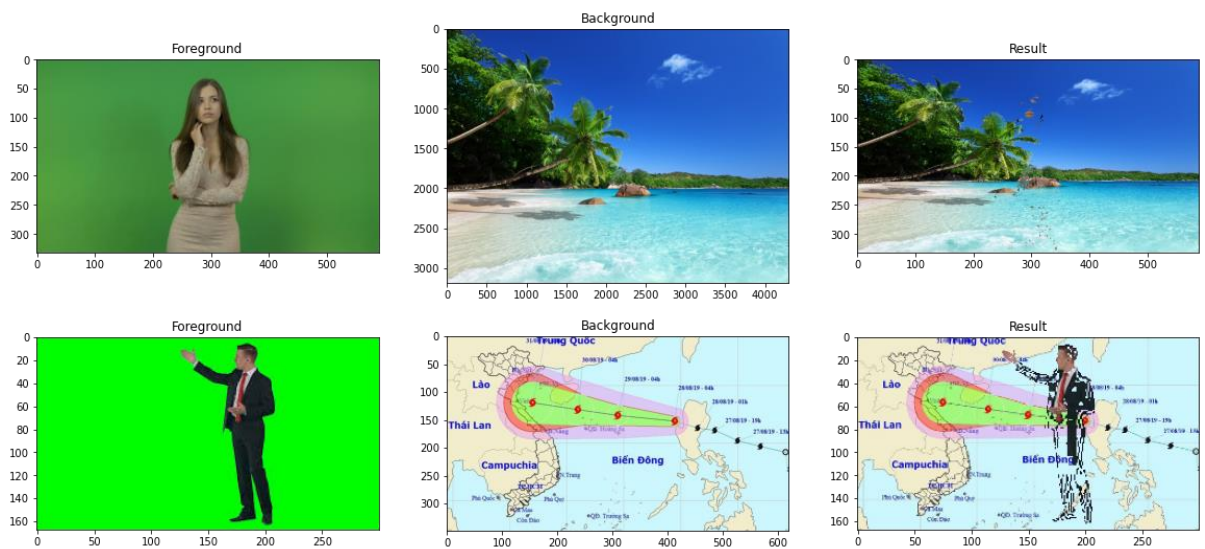
Kết quả tệ, ở ảnh đầu phần đối tượng gần như bị nhầm thành phần nền hết.
Ở hình thứ hai gần như không nhận diện được đối tượng

2. $Threshold = 0.6$



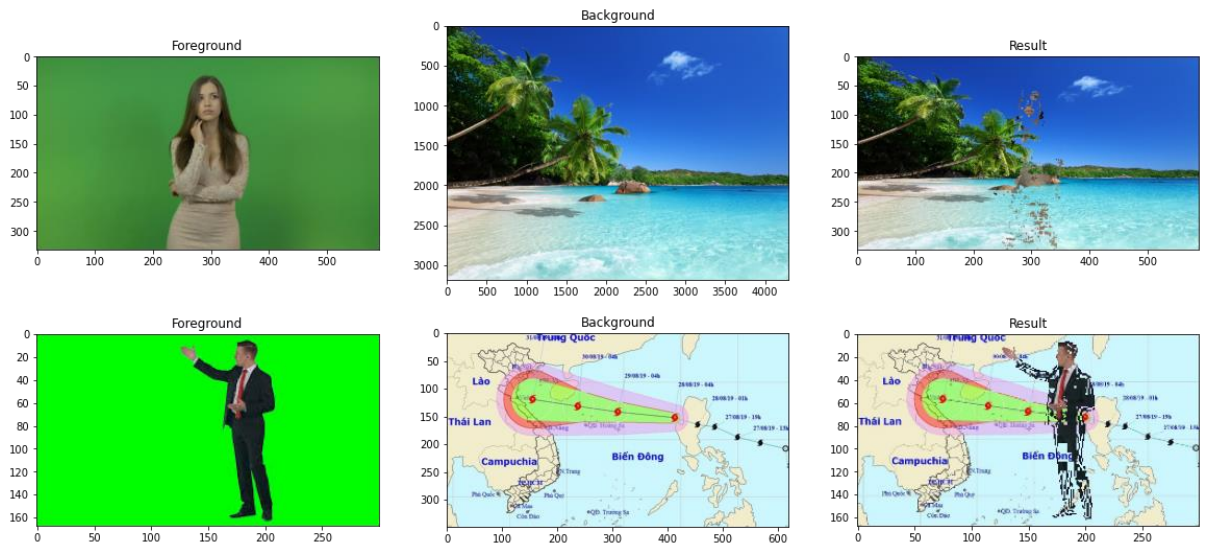
Kết quả vẫn không khác gì mấy so với $Threshold=0.5$

3. $Threshold = 0.7$



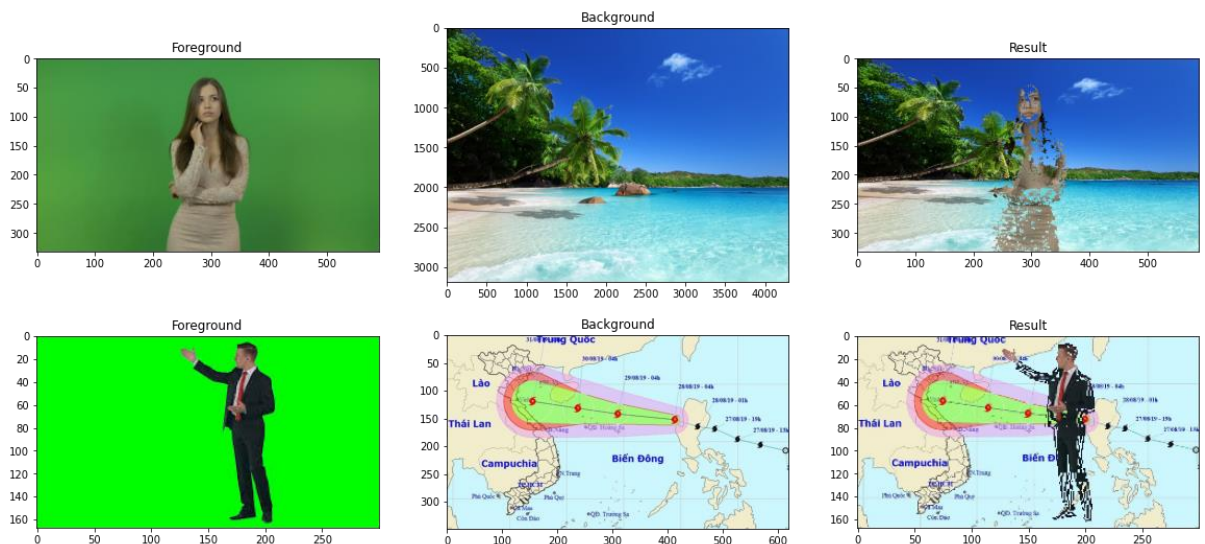
Vẫn chưa tốt được hơn nhiều

4. $Threshold = 0.8$



Vẫn giống ở trên nhưng có cải thiện hơn.

5. *Threshold = 0.9*

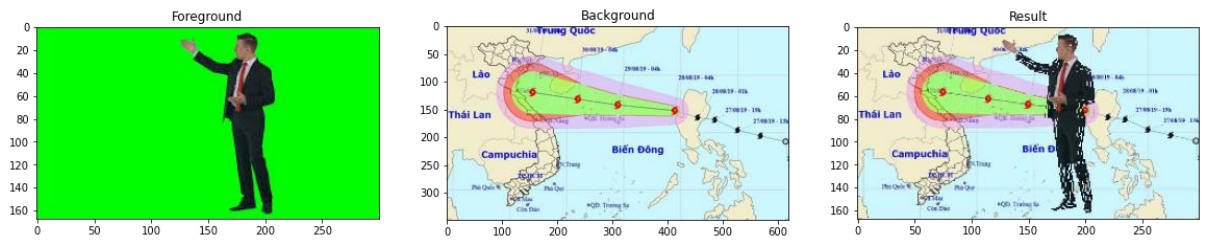


Có cải thiện hơn nhiều so với Threshold=0.5. Tuy nhiên vẫn chưa thể thấy rõ hết đối tượng.

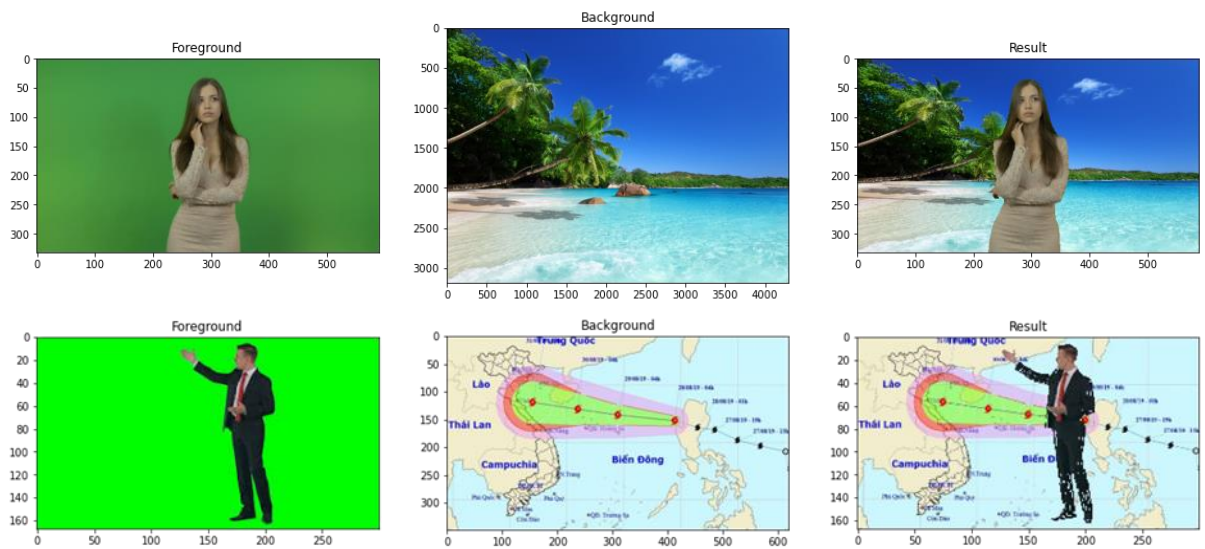
6. *Threshold = 0.99*

- Kết quả được cải thiện hơn ta đã thấy đối tượng nhưng một vài chỗ của đối tượng vẫn bị đang bị nhận diện nhầm thành nền.





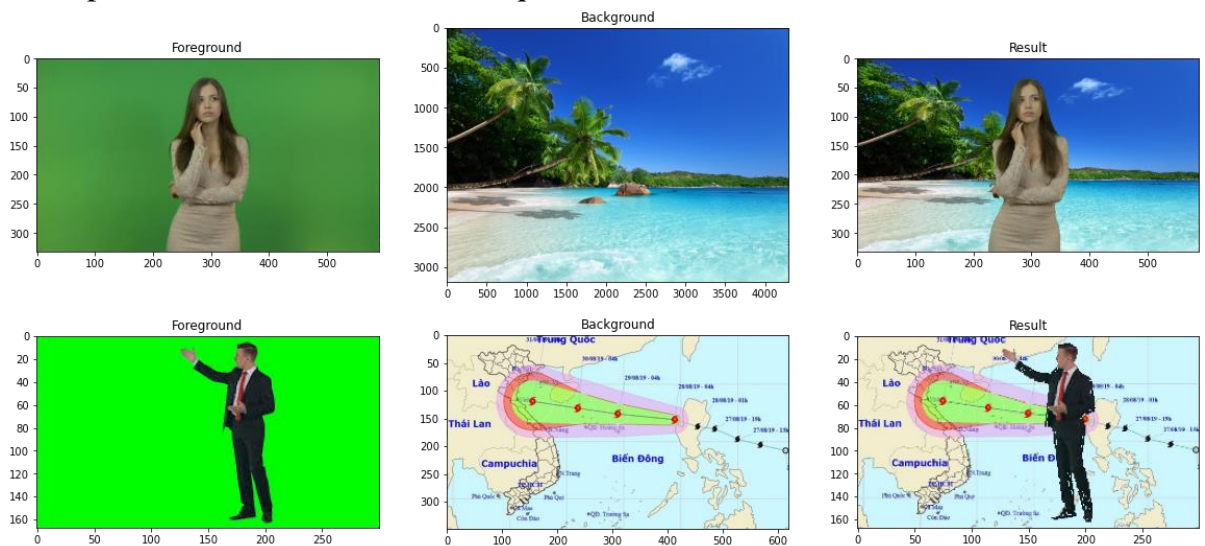
7. $Threshold = 0.999$



Vẫn bị lốm đốm trắng ở hình thứ hai giống ở trên.

8. $Threshold = 0.9999$

Ảnh thứ hai bên trong đối tượng vẫn còn một vài chỗ nhỏ bị thiếu đi. Nên tụi em tiếp tục thử thêm để có được kết quả hoàn thiện hơn



9. $Threshold = 0.99999$

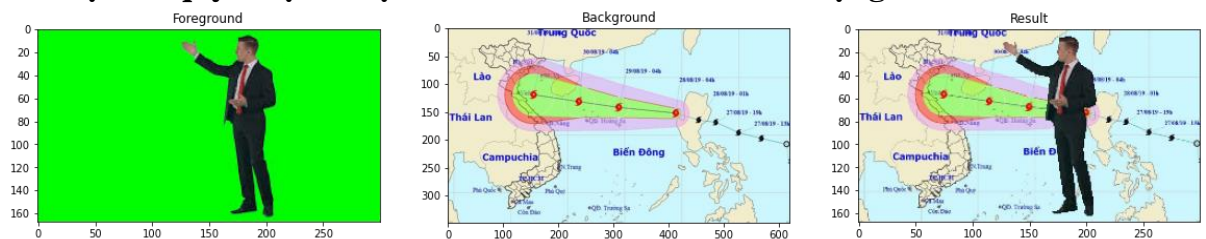


Vẫn bị lốm đốm trắng nhỏ.

10. $Threshold = 0.999999$

Không còn bị lốm đốm trắng bên trong chủ thể chỉ có đường viền hơi răng cưa do hình ảnh này chưa được sắc nét.

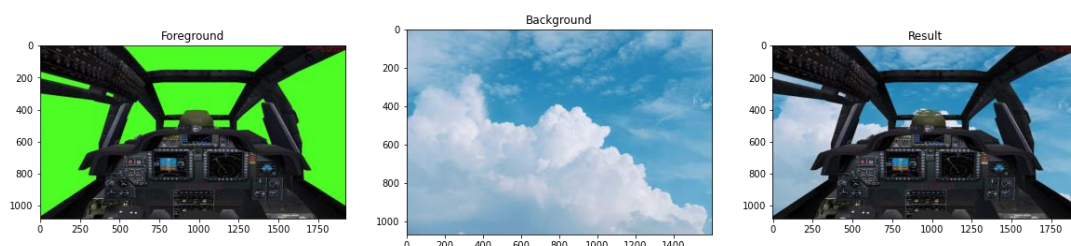
→ Nên tụi em quyết định chọn $threshold = 0.999999$ để sử dụng



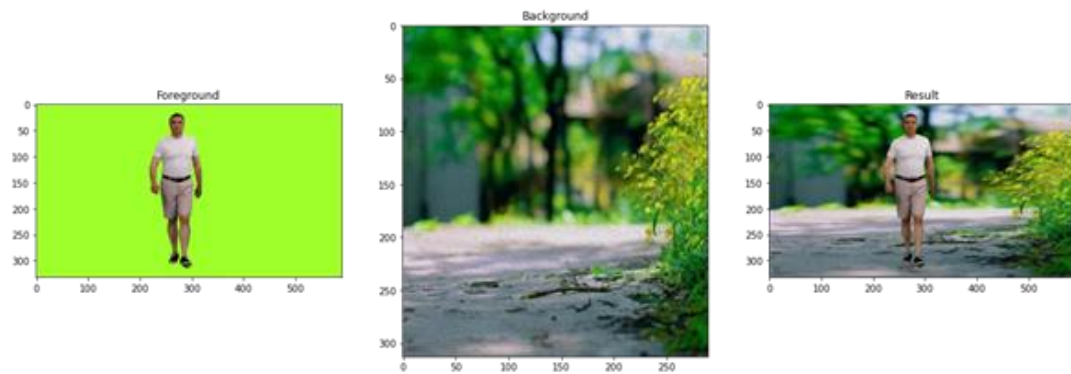
4. Kết quả

Sau khi chạy và quan sát với nhiều kết quả thử nghiệm nhóm chúng em đã rút ra được một số nhận xét về phương pháp này như sau.

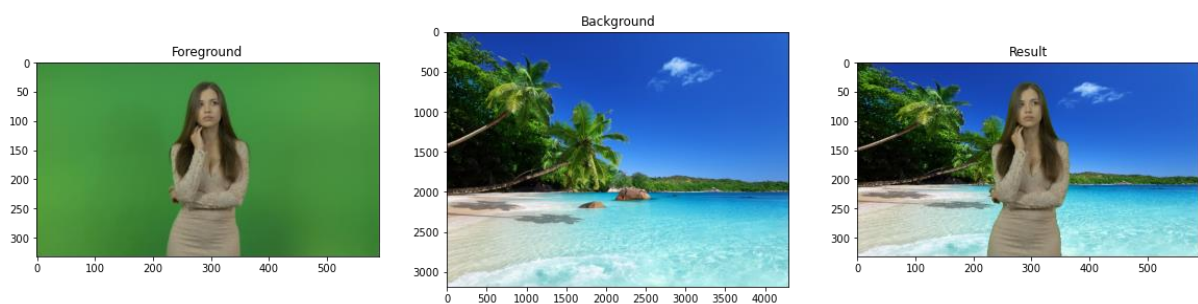
Tám ảnh có màu nền thỏa mãn điều kiện của Chroma key, chất lượng ảnh tốt thì đối tượng và nền được tách ra khá hoàn chỉnh.



Hình 4. 1

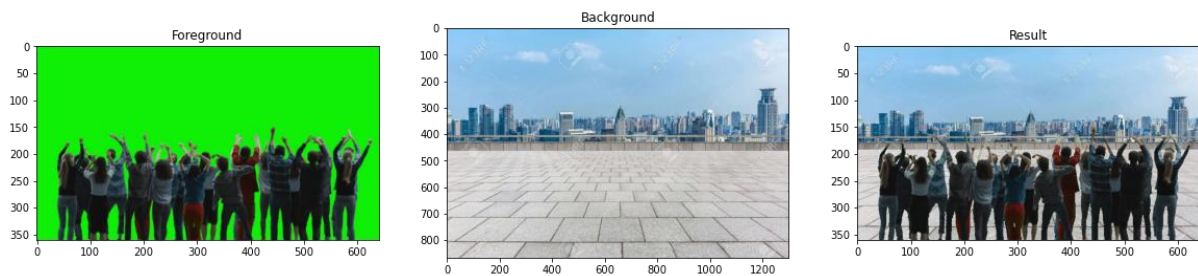


Hình 4. 2



Hình 4. 3

Những tấm ảnh có nhiều đối tượng cũng tách ra tốt.

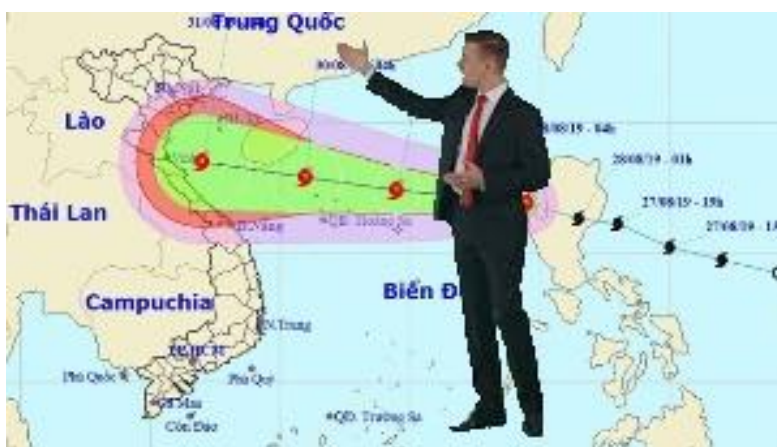


Hình 4. 4

Đối với những tấm ảnh có độ phân giải thấp thì phần chủ thể lúc cắt ra sẽ không được tốt -> phần chủ thể sẽ không được rõ nét (bị răng cưa). Tuy nhiên so với phương pháp đầu các phần bên trong của đối tượng không bị đốm trắng.

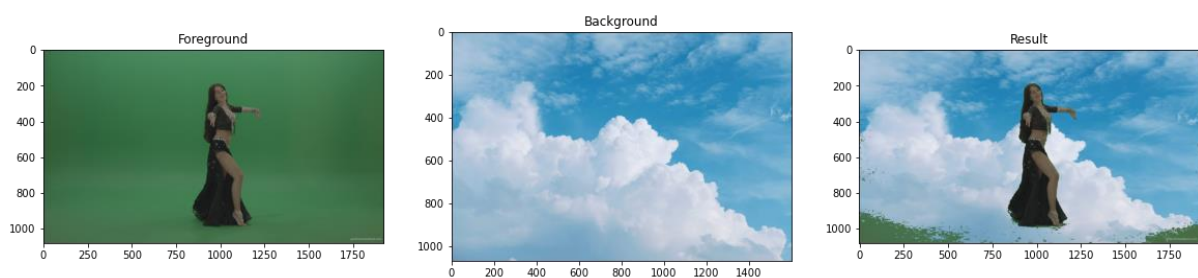


Hình 4. 5



Hình 4. 6

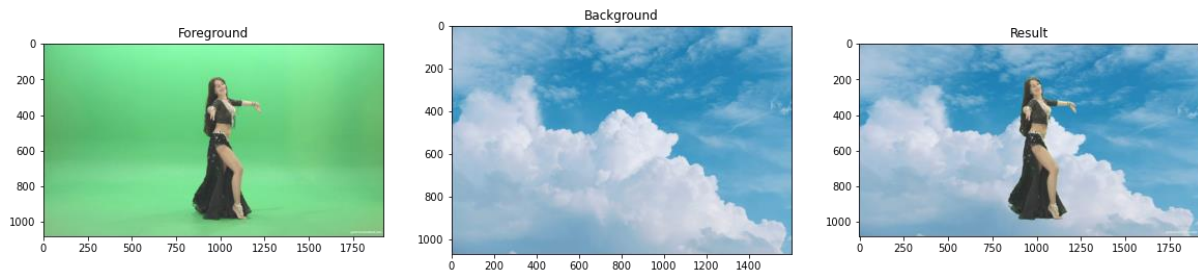
Nhưng tấm ảnh quá tối và có màu nền không đều màu sẽ mang lại kết quả không như mong muốn



Hình 4. 7

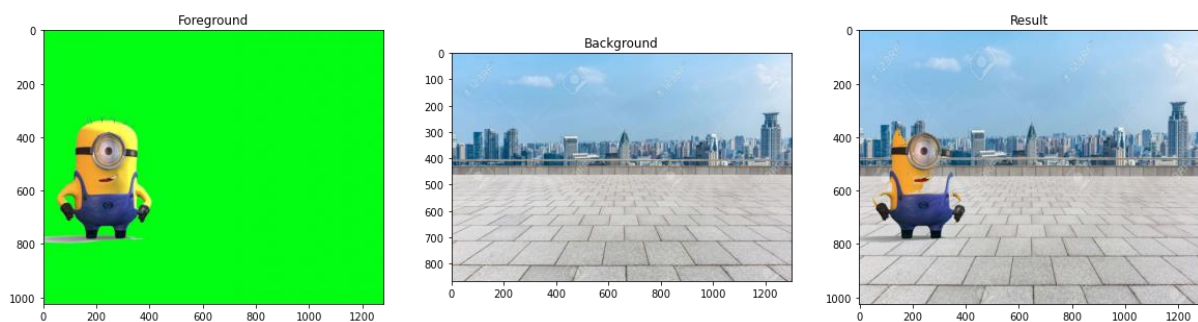
→ Do background quá tối nên model không phân biệt được hết background.

+ Cách khắc phục: Nhóm tụi em đã tiến hành chỉnh sáng bức ảnh lên và nhận thấy kết quả mang lại khá tốt so với ảnh gốc

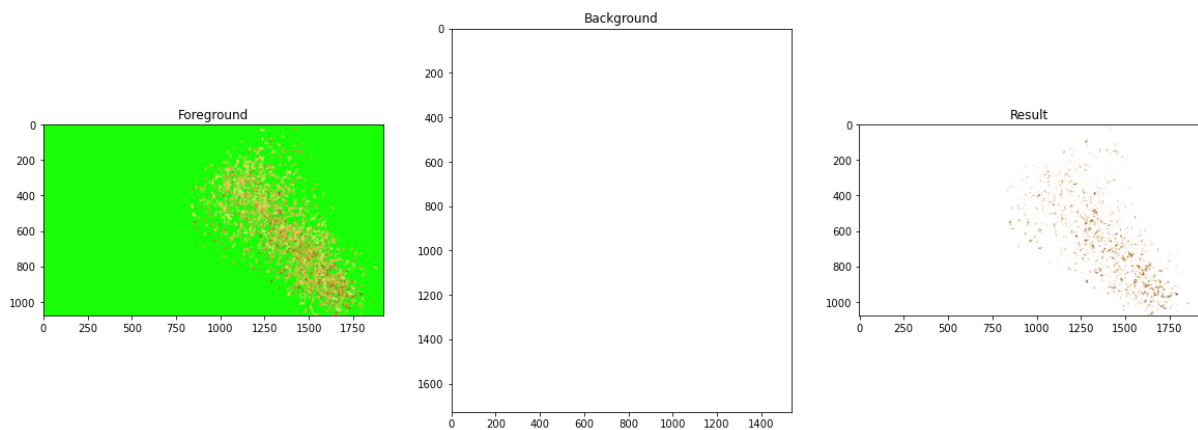


Hình 4. 8

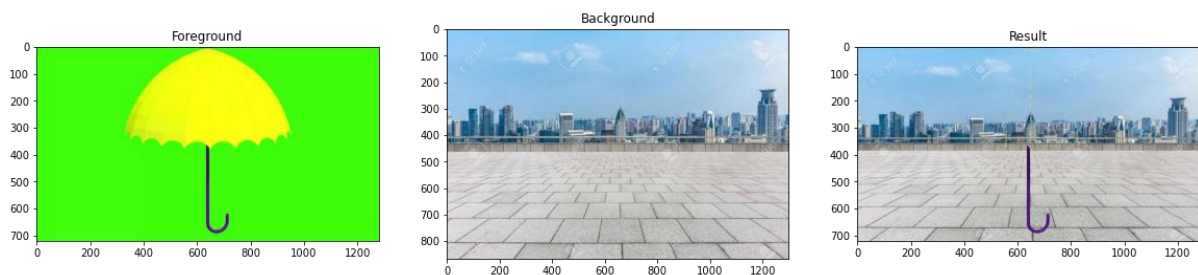
Chưa nhận diện tốt các chủ thể có màu vàng sáng. Bởi vì màu vàng sáng (255,255,102) còn màu xanh sáng (102,255,102). Kênh màu green, blue gần như giống nhau chỉ có khác ở kênh màu red.



Hình 4. 9



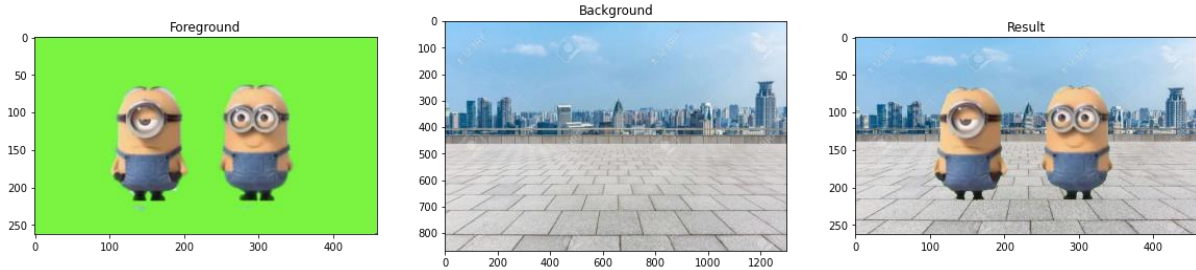
Hình 4. 10



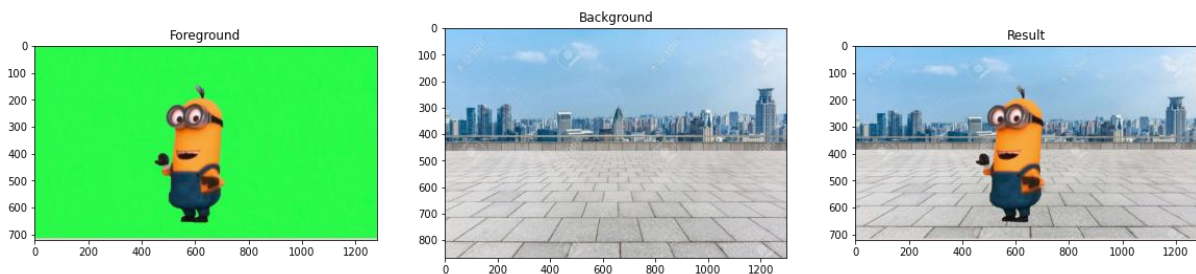
Hình 4. 11

➔ Phần chủ đề bị nhận diện sai khiến chủ thể bị mất, mang lại kết quả không mong muốn

- Tuy nhiên với những chủ thể có màu vàng đậm hơn thì mang lại kết quả có thể chấp nhận được



Hình 4. 12



Hình 4. 13



Hình 4. 14

Về xử lý video cũng mang lại kết quả có thể chấp nhận được. Tuy nhiên tốc độ đọc các frame còn chậm khiến cho video như bị delay

(gần video đó đây)

5. Kết luận

Ưu điểm:

- + Đây là một phương pháp cũng khá đơn giản dễ thực hiện và triển khai.
- + (Phương pháp này không bị bó buộc vào một màu nền cụ thể.

- + Chúng ta có thể dễ dàng thay đổi model bằng cách train lại model với màu nền mong muốn.
- + Và thời gian xử lý cũng rất là nhanh.
- + Việc tách phân đối tượng ra khỏi nền khá ổn tách được hết tất cả mọi phần của đối tượng

Khuyết điểm:

Đối với vài trường hợp kết quả thu được không được như mong muốn. Chẳng hạn như tấm ảnh có màu nền thỏa mãn điều kiện về chroma key nhưng màu sắc quá tối phong xanh chưa được làm mịn (4.7) hay những bức ảnh được chụp bằng máy ảnh chất lượng thấp (hình 4.5) những tấm ảnh có chủ thể với màu vàng quá sáng (hình 4.9,4.10,4.11).

Hướng phát triển:

Có thể train thêm nhiều màu phong xanh hơn và thêm vô tập foreground nhiều chủ thể có màu vàng sáng để khắc phục được việc nhận dạng sai chủ thể có màu vàng sáng.

Thêm vô phần xử lý ảnh để những tấm ảnh có chất lượng kém như 4.5 hoặc tấm ảnh quá tối nền không đều (hình 4.7) được xử lý làm nét hơn và chỉnh sáng lên cũng như làm đều màu nền phía sau trước khi vào công đoạn tách nền ghép background.

TÀI LIỆU THAM KHẢO

- [1] [Chroma key là gì](#)
- [2] [Blue screen vs green screen](#)
- [3] [FPGA implementation for real-time Chroma-key effect using Coarse and Fine Filter](#)
- [4] [High Quality Chroma Key](#)
- [5] [Image Segmentation Using Color Spaces in OpenCV + Python](#)
- [6] [Chroma key wiki](#)
- [7] [Logistic regression](#)
- [8] [DECISION BOUNDARY FOR CLASSIFIERS: AN INTRODUCTION](#)