**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP HÀ NỘI**

KHOA CÔNG NGHỆ THÔNG TIN

**-------------------**

A picture containing screenshot, text, graphics, logo

Description automatically generated

**BÁO CÁO THỰC NGHIỆM HỌC PHẦN**

**XỬ LÝ ẢNH SỐ VÀ THỊ GIÁC MÁY TÍNH**

**CHỦ ĐỀ**

**TÌM HIỂU, CÀI ĐẶT, ỨNG DỤNG**

**CỦA PHƯƠNG PHÁP TRÍCH CHỌN ĐẶC TRƯNG LBP**

GVHD: ThS. Lê Thị Thuỷ

Sinh viên:

|  |  |
| --- | --- |
| Ngô Thế Tài |  |
| Nguyễn Văn Vũ | 2020601029 |
| Đỗ Trung Hiếu |  |
| Trần Thị Khánh Linh |  |

Lớp: 20222IT6072002 – Khoá 15

**Hà Nội - Năm 2023**

**MỤC LỤC**

[CHƯƠNG 1. TÌM HIỂU PHƯƠNG PHÁP TRÍCH CHỌN ĐẶC TRƯNG LOCAL BINARY PATTERN 4](#_heading=h.z337ya)

[1.1.](#_heading=h.3j2qqm3) Giới thiệu 4

[1.2.](#_heading=h.1y810tw) Local binary pattern 5

[1.3.](#_heading=h.4i7ojhp) Local binary pattern mở rộng 7

[1.4.](#_heading=h.2xcytpi) Local binary pattern với phép quay 8

[1.5.](#_heading=h.1ci93xb) Mẫu thống nhất (uniform pattern) 9

[1.6.](#_heading=h.3whwml4) Tìm hiểu hệ màu trong xử lý ảnh 11

[1.6.1.](#_heading=h.2bn6wsx) Hệ màu RGB 11

[1.6.2.](#_heading=h.qsh70q) Hệ màu HSV 11

[CHƯƠNG 2. CHƯƠNG TRÌNH 13](#_heading=h.3as4poj)

MỤC LỤC HÌNH ẢNH

[Hình 1. Dãy nhị phân và giá trị mới theo chiều kim đồng hồ từ vị trí trái trên cùng 6](#_heading=h.3znysh7)

[Hình 2. Ví dụ về cách lấy mẫu 8](#_heading=h.tyjcwt)

[Hình 3. Xoay điểm ảnh theo chiều kim đồng hồ 9](#_heading=h.1t3h5sf)

[Hình 4. Một số ví dụ về mẫu đồng nhất và không phải mẫu đồng nhất 10](#_heading=h.2s8eyo1)

# CHƯƠNG 1. TÌM HIỂU PHƯƠNG PHÁP TRÍCH CHỌN ĐẶC TRƯNG LOCAL BINARY PATTERN

## Giới thiệu

Phân tích cấu trúc hai chiều có nhiều ứng dụng tiềm năng, ví dụ ứng dụng cảm biến từ xa, kiểm tra bề mặt, phân tích hình ảnh trong ngành Y, … Nhưng có một số ví dụ hạn chế về việc triển khai các kết cấu đặc trưng của đối tượng. Một vấn đề xảy ra đó là kết cấu trong thế giới thực thường không đồng nhất, do sự thay đổi về hướng, tỉ lệ hoặc hình thức trực quan khác. Sự bất biến của ảnh xám thường rất quan trọng do độ sáng không đồng đều hoặc độ biến thiên của ảnh đó. Ngoài ra độ phức tạp tính toán của hầu hết các thuật toán được cho là quá cao.

Hầu hết các cách tiếp cận để phân loại các kết cấu đều giả định, rằng các mẫu chưa được biết được phân loại giống hệt với các mẫu đào tạo về quy mô không gian, hướng và các thuộc tính trong thang độ xám. Tuy nhiên, các kết cấu trong thế giới thực có thể xảy ra ở độ phân giải và góc quay không gian tuỳ ý và chúng có thể chịu các điều kiện chiếu sáng khác nhau.

Local Binary Patterns (LBP – Mẫu nhị phân cục bộ) là một bộ mô tả nhằm mục đích tóm tắt hiệu quả các cấu trúc cục bộ của hình ảnh. Trong những năm gần đây, sự quan tâm ngày càng tăng trong nhiều lĩnh vực xử lý hình ảnh và thị giác máy tính và đã cho thấy tính hiệu quả của nó trong một số ứng dụng, đặc biệt là phần tích hình ảnh khuôn mặt, bao gồm các nhiệm vụ như phát hiện khuôn mặt, nhận dạng khuôn mặt, phân tích nét mặt, phân loại nhân khẩu học, …

Trong báo cáo này nhóm chúng em trình bày về phương pháp phân tích hình ảnh khuôn mặt dựa trên LBP và ứng dụng vào một số bài toán điển hình của LBP.

## Local binary pattern

Local Binary Pattern là một phương pháp rút trích đặc trưng và các đặc trưng sẽ được chọn lọc và thu về một vector đặc trưng. Khi xử lý ảnh, một điều cần thiết đó là một đối tượng bất biến với phép chiếu sáng và phép xoay. Do đó LBP được phát minh để giải quyết hai vấn đề

***Phép chiếu sáng***

Phép chiếu sáng được sử dụng để tăng cường đặc trưng của của các điểm ảnh trong quá trình tính toán.

***Phép xoay***

Tức là khi xoay hình ảnh theo một góc xác định, ví dụ quay hình với góc 90 độ

Vào năm 1996, Ojala đã đề xuất phương pháp LBP ban đầu của LBP chỉ bất biến với phép chiếu sáng. Phiên bản đầu tiên sử dụng một ma trận 3x3 trượt qua từng điểm ảnh, lấy 8 điểm xung quanh và so sánh với điểm trung tâm. Nếu điểm nào lớn hơn hoặc bằng điểm trung tâm thì đặt giá trị mới là 1 ngược lại thành 0. Sau khi tính toán, có thể bắt đầu từ vị trí bất kì tại 8 điểm xung quanh và có thể đi ngược kim đồng hồ và xuôi kim đồng hồ để kết hợp thành mã nhị phân, chuyển chúng thành số thập phân để lưu lại kết quả mới.

A picture containing text, screenshot, diagram, font

Description automatically generated

Hình 1. Dãy nhị phân và giá trị mới theo chiều kim đồng hồ từ vị trí trái trên cùng

Trong bức ảnh, lấy ra 9 pixel ứng với ma trận 3x3. Với những giá trị lớn hơn hoặc bằng giá trị trung tâm thì sẽ đặt là 1 ngược lại là 0. Xây dựng Binary Pattern – LBP theo hướng kim đồng hồ. Giá trị mới được tính như sau:

Nói cách khách, với một vị trí (), LBP được định nghĩa là một tập hợp các so sánh nhị phân có thứ tự về cường độ pixel giữa pixel trung tâm và các pixel xung quanh. Giá trị thập phân kết quả thu được từ 8-bit có thể được biểu diễn như sau:

Trong đó tương ứng với giá trị xám của pixel vị trí (), là các giá trị xám của 8 ô xung quanh, hàm s(k) được định nghĩa như sau:

***Ưu điểm***

* Tính toán đơn giản và nhanh chóng
* Khả năng khái quát hoá cao, có thể được áp dụng cho hầu hết các loại hình ảnh, bao gồm ảnh xám và ảnh màu, trích xuất đặc trưng quan trọng của hình ảnh như cạnh, góc, màu sắc, cấu trúc.
* Khả năng chống nhiễu tốt, do phụ thuộc vào các giá trị nhị phân của các điểm lân cận, các nhiễu nhỏ không ảnh hưởng nhiều đến giá trị pixel.
* Độ tin cậy cao, với phiên bản đầu tiên của LBP xét 8 điểm xung quanh thì thường cho kết quả rất tốt trong các nhiệm vụ nhận dạng và phân loại hình ảnh, đặc biệt trong những bộ dữ liệu lớn.

***Nhược điểm***

* Không phù hợp với một số loại hình ảnh cụ thể, ví dụ các hình ảnh có độ phân giải thấp, việc xử dụng LBP phiên bản đầu tiên có thể dẫn đến xự mất mát thông tin hoặc giảm độ chính xác của quá trình trích xuất đặc trưng.
* Biểu diễn cho các đặc trưng thì kích thước quá lớn, đặc trưng kém ổn định

## Local binary pattern mở rộng

Từ các nhược điểm ở trên Ojala đã đưa ra một phương pháp để cải tiến thuật toán này. Trong một bức hình, tại một điểm ảnh (), thay vì lấy 8 điểm lân cận ở phương pháp cũ, tác giả đã lấy tập hợp P điểm tạo thành hình tròn có bán kính R, và T là một phân phối đại số của P+1 điểm ảnh đang xét. (Tính cả điểm trung tâm).

Trong đó, và , , … ứng với cấp độ xám của điểm ảnh trung tâm và các điểm lân cận. Cách lấy mẫu được ví dụ như hình 2

A picture containing diagram, line

Description automatically generated

Hình 2. Ví dụ về cách lấy mẫu

Khi tăng bán kính R thì mối tương quan giữa các điểm ảnh giảm dần, cho nên thường sẽ thực hiện trên R nhỏ cụ thể là 1 hoặc 2. Thực hiện tính toán trên mỗi điểm ảnh không làm thay đổi khoảng cách giữa các điểm ảnh. Do đó có thể trừ đi một lượng . Thay thế lượng thành hàm s(k) đã được đề cập ở trên sẽ loại bỏ được việc cùng tăng, cùng giảm độ sáng.

Sau đó tính toán các giá trị biểu diễn với là trọng số tương ứng với công thức

***Ưu điểm***

Bất biến với phép chiếu sáng. Có chi phí tính toán thấp, ổn định khi cường độ thay đổi đơn điệu và dễ mở rộng lên không gian nhiều chiều như RGB, HSV.

***Nhược điểm***

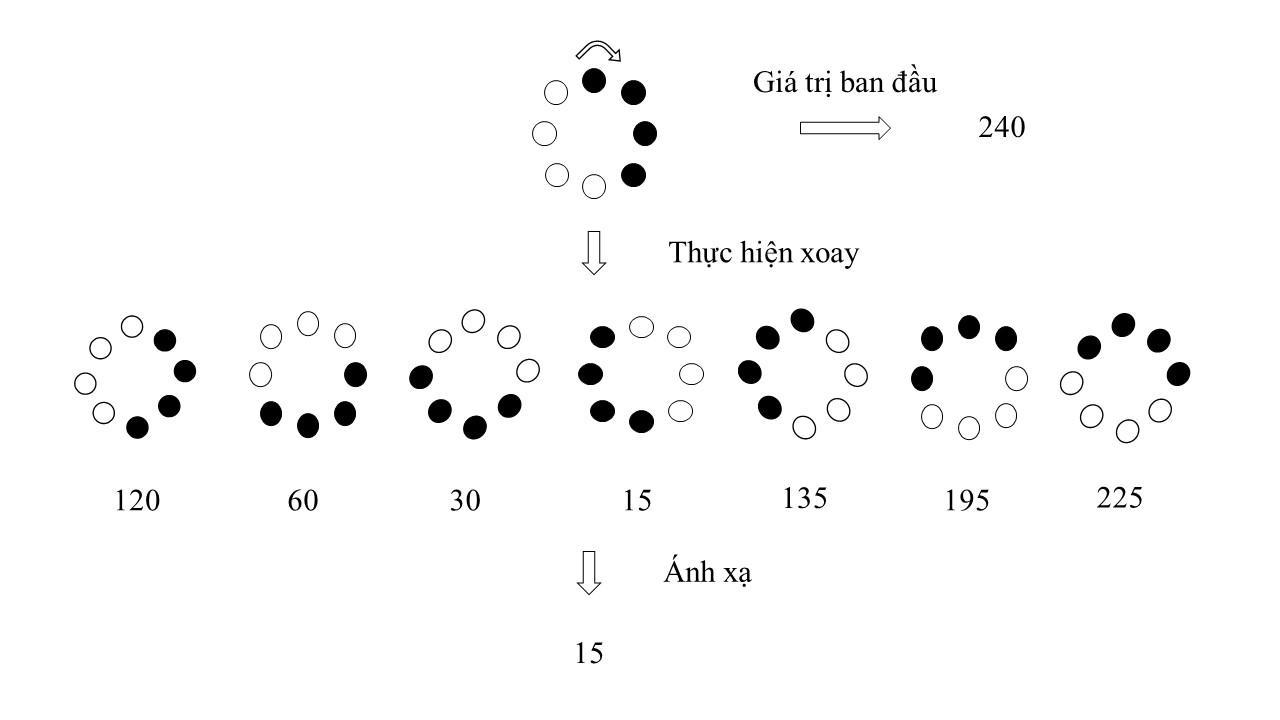
Nếu số điểm lân cận được xét quá lớn, giá trị đặc trưng cho kết cấu sẽ tăng quá cao. Không rõ ràng vì hai biểu diễn cục bộ giống nhau có thể nằm ở hai đối tượng khác nhau

## Local binary pattern với phép quay

Ngoài việc bất biến với phép chiếu sáng như đã được đề cập ở phần trước thì phương pháp này cần phải bất biến với phép xuay. Kí hiệu là

Với là viết tắt của rotation invatiant (Phép quay không đổi, hoặc bất biến với phép quay)

Để bất biến với phép xoay, gọi là ảnh được xoay một góc . Sau khi xoay tại vị trí sẽ được chuyển đổi, ánh xạ sang .



Hình 3. Xoay điểm ảnh theo chiều kim đồng hồ

Sau khi được các kết quả trên, lấy kết quả nhỏ nhất đại diện cho tất cả phép quay. Do đó công thức bất biến của phép chiếu sáng và phép xoay là:

Trong đó là phép dịch vòng tròn số nhị phân lần. Phương pháp này kèm theo bất biến với phép xoay nên có nhiều ứng dụng hơn với Local binary pattern mở rộng.

## Mẫu thống nhất (uniform pattern)

Một bộ LBP được gọi là thống nhất nếu mẫu nhị phân hình tròn theo chiều kim đồng hồ có chứa nhiều nhất 2 lần chuyển đổi bit từ 0 sang 1 và ngược lại.

A picture containing text, screenshot, number, font

Description automatically generated

Hình 4. Một số ví dụ về mẫu đồng nhất và không phải mẫu đồng nhất

Mỗi mẫu trong số này có một bin – (Đại diện cho khoảng dữ liệu, hoặc mức xám trong biểu đồ Histogram). Các mẫu còn lại với hơn 2 lần chuyển đổi bit sẽ được tính vào một bin riêng.

Dưới đây là công thức để kiểm tra mẫu nhị phân đó co phải là mẫu nhị phân nhỏ hơn hoặc bằng 2 lần chuyển đổi bit hay không

Khi thì mẫu nhị phân đó là UNIFORM. Các mẫu còn lại được tính chung giá trị chuyển thành 0. Như vậy công thức cuối cùng của UNIFORM PATTERN là:

Trong đó v(x) là hàm ánh xạ các giá trị của mẫu nhị phân sang khoảng giá trị mới từ 1 đến 58. Như vậy với việc biểu diễn như này thì sẽ có 59 mẫu, giảm từ 255 mẫu. Việc sử dụng UNIFORM PATTERN (Mẫu nhị phân thống nhất) giúp giảm số lượng đặc trưng và cải thiện tính ổn định của các đặc trưng được tạo ra trong quá trình nhận dạng. Phương pháp này chỉ có tối đa 2 lần chuyển đổi bit nên ít bị ảnh hưởng bởi nhiễu và thay đổi độ sáng. Do đó việc sử dụng UNIFORM PATTERN có thể giúp tăng độ chính xác của quá trình nhận dạng, giảm thiểu các sai sót không mong muốn.

## Tìm hiểu hệ màu trong xử lý ảnh

Không gian màu (Colour Space) được hiểu là các mô hình toán để miêu tả màu sắc. Mỗi không gian màu đều có một tác dụng và ứng dụng trong các bài toán khác nhau.

***Các hệ màu cơ bản***

### Hệ màu RGB

A picture containing colorfulness, circle, graphics, creativity

Description automatically generated

Hình 5. Hệ màu RGB

RGB là không gian màu phổ biến dùng trong máy tính, máy ảnh, điện thoại, và nhiều thiết bị kĩ thuật số khác nhau. Không gian này khá gần với cách mắt người tổng hợp màu sắc. Nguyên lý cơ bản sử dụng 3 màu sắc cơ bản R(red – đỏ), G(green – xanh lục), B(blue – xanh lam) để biểu diễn tất cả màu xắc.

Hệ màu RGB sử dụng 8bit để biểu diễn cho mỗi kênh màu, tức là giá trị R, G, B nằm trong khoảng 0 đến 255. Bộ 3 số này biểu diễn cho từng điểm ảnh, mỗi số biểu diễn cho cường độ của một màu. Với mô hình này thì tối đa có thể tạo ra 255x255x255=16581375 màu. Với thư viện matplotlib, Pillow của Python sẽ đọc ảnh đầu vào là RGB. Tuy nhiên với Opencv thì sẽ đọc ảnh ở dạng BGR.

### Hệ màu HSV

Không gian màu HSV là một cách để mô tả màu sắc dựa trên 3 số liệu:

* H: (Hue) vùng màu

Hue là phần màu của mô hình màu và được biểu thị dưới dạng một số từ 0 đến 360 độ:

A picture containing text, screenshot, font, number

Description automatically generated

Hình 6. Phân bố kênh màu H trong HSV

* S: (Saturation) Độ bão hoà màu

Độ bão hoà là lượng màu xám trong màu, từ 0 đến 100%. Một hiệu ứng mờ nhạt có thể có được từ việc giảm độ bão hoà về 0 để thể hiện nhiều màu xám hơn.

* V: (Value) Độ sáng hoặc giá trị

Giá trị kết hợp với độ bão hoà và mô tả sáng hoặc cường độ màu sắc, từ 0 đến 100%. Trong đó 0% là hoàn toàn đen, 100% là sáng nhất.

* + 1. Ảnh xám và ảnh nhị phân

Ảnh xám (GrayScale) đơn giản là một hình ảnh trong đó các màu sắc là các sắc thái của màu xám. Ảnh xám cung cấp ít thông tin hơn cho mỗi pixel trong khi các hệ màu khác thường là 3 kênh màu thì ảnh xám chỉ có 1.

Ảnh nhị phân (Binary) là kết quả của việc giảm thông tin của ảnh xám về 2 giá trị 0 hoặc 255. Tuy thông tin bị giảm khá nhiều nhưng ảnh nhị phân lại thể hiện được rất rõ các yếu tố về góc cạnh và hình dạng của vật thể, điều này có lợi trong những bài toán như lọc nhiễu, nhận dạng vật thể, …

# CHƯƠNG 2. CHƯƠNG TRÌNH

* 1. **Lớp đọc hình ảnh**

A picture containing text, screenshot, font, line

Description automatically generated

Hình 7. Khởi tạo lớp BaseImage đọc hình ảnh

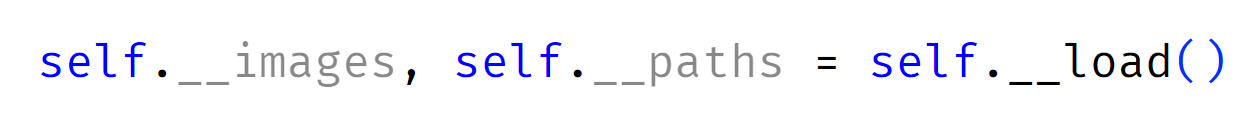
Lớp BaseImage khởi tạo gồm 3 tham số đầu vào:

* root: đường dẫn đến thư mục chứa các ảnh, mặc định là “data/”
* types: kiểu đuôi file có thể là “.png”, “.jpeg”, nếu đọc hết thì sẽ là “\*”
* size: kích cỡ của toàn bộ ảnh khi đọc vào mặc định là (128, 128)

A picture containing text, screenshot, font, line

Description automatically generated

Hình 8. Hàm đọc ảnh và trả về danh sách ảnh, đường dẫn



Hình 9. Lưu lại kết quả tại hàm \_\_init\_\_ sau khi thực hiện đọc ảnh

Hàm \_\_load thực hiện đọc ảnh từ đường dẫn ảnh đã được khởi tạo ở trên. Sử dụng thư viện glob để liệt kê tất cả các file có trong đường dẫn self.\_\_root, có đuôi file là self.\_\_types. Thực hiện đọc ảnh bằng thư viện opencv-python và chỉnh lại kích thước về kích thước đã được định nghĩa sẵn ở trên. Cuối cùng hàm trả về 2 giá trị đó là \_\_images và \_\_paths lần lượt là 1 danh sách ảnh được đọc ở hệ màu BGR và 1 danh sách đường dẫn của mỗi ảnh.

***Ví dụ cụ thể***

A picture containing text, font, line, diagram

Description automatically generated

Hình 10. Ví dụ về đường dẫn ảnh

Ví dụ trên trong thư mục data có 3 loại ảnh là png, bmp và jpg. Khởi tạo hàm BaseImage như sau

A picture containing text, font, line, screenshot

Description automatically generated

Hình 11. Ví dụ khởi tạo hàm BaseImage

Lớp BaseImage nhận vào thư mục root là “data/”, gồm 3 kiểu đuôi file là png, bmp và jpg. Cuối cùng là kích cỡ ảnh (224, 224). Sau đó, hàm \_\_load được gọi và kích thước của 2 giá trị trong hình 7 như sau

* self.\_\_images: kích cỡ là (3, 224, 224, 3). Do có 3 ảnh và kích thước là 224x224 với 3 kênh màu B, G, R
* self.\_\_paths: Kích cỡ là (3, ) do có 3 ảnh và với mỗi giá trị là đường dẫn của mỗi ảnh: ['data/obj.png' 'data/obj1.bmp' 'data/test.jpg']

A screenshot of a computer code

Description automatically generated with low confidence

Hình 12. Các hàm get, set để thực hiện gán và lấy giá trị

* 1. **Lớp LBP**

**A screen shot of a computer code

Description automatically generated with low confidence**

Hình 13. Khởi tạo lớp LBP

Lớp LBP kế thừa từ lớp BaseImage

* root: đường dẫn ảnh, mặc định “data/”
* types: đuôi file có thể là “\*.png”, … Mặc định là “\*”
* size: kích cỡ của ảnh khi đọc vào
* is\_json: có thể đọc ảnh từ file json, nếu đọc ảnh từ file json thì giá trị này bằng True và root để rỗng. ngược lại là False
* json\_data: đường dẫn file json chứa dữ liệu ảnh
* json\_path: đường dẫn file chứa các đường dẫn ảnh
* save: Nếu giá trị bằng True thì hàm sẽ thực hiện lưu lại giá trị đặc trưng của ảnh vào file json

**A picture containing text, screenshot, font, number

Description automatically generated**

Hình 14. Tạo hàm ánh xạ từ số nhị phân sang khoảng giá trị từ 1 đến 58

Hàm \_\_genBIT trên tạo một ánh xạ kiểm tra tất cả các số nhị phân nằm trong khoảng 0 đến 255. Kiểm tra xem số nhị phân đó có phải là Uniform pattern hay không. Nếu đúng thì sẽ lưu lại vào một dictionary (từ điển) để có thể dễ dàng truy cập khi sử dụng hàm LBP.

A picture containing text, screenshot, font, number

Description automatically generated

Hình 15. Hàm chuyển ảnh xám thành LBP

Hàm LBP trên sử dụng phương pháp Uniform pattern để chuyển đổi một ảnh xám thành LBP gồm 59 giá trị khác nhau.

A picture containing text, font, screenshot, line

Description automatically generated

Hình 16. Hàm tính Histogram của một hình ảnh

Do tính toán bằng Uniform pattern nên giá trị nằm trong khoảng 0 đến 58. Vì vậy giá trị bins ở hàm np.histogram là 59. Nếu đặt là 256 như các ảnh RGB, BGR, thì sẽ bị thừa. Kết hợp các hàm ở trên để tạo thành một hàm trích xuất đặc trưng hoàn chỉnh như hình sau.

A picture containing text, screenshot, font, number

Description automatically generated

Hình 17. Hàm trích xuất đặc trưng

Hàm trên xử lý ảnh trên hệ màu HSV và chuyển mỗi kênh màu H, S, V sang LBP được và chuyển chúng thành Histogram, kết quả là tạo thành 3 vector có kích cỡ là 59. Nối chúng lại với nhau để đặc trưng cuối cùng là một vector có kích cỡ là 177. Như vậy với ví dụ 3 hình ảnh ở trên thì đặc trưng thu được sau khi hàm \_\_get chạy có kích thước là (3, 177).

Khi được các hàm ở trên, ở hàm \_\_init\_\_ thực hiện các câu lệnh sau câu lệnh super như sau

A picture containing text, screenshot, font, line

Description automatically generated

Hình 18. Câu lệnh thực hiện các chức năng của LBP

Đầu tiên khởi tạo hàm ánh xạ các mẫu nhị phân đã được đề cập ở trên và lưu vào biến self.\_map. Nếu is\_json là True thì sẽ thực hiện đọc các đặc trưng, đường dẫn ảnh từ file json đã được truyền ở biến json\_data và json\_path. Ngược lại đọc dữ liệu bằng hàm self.\_\_get. Cuối cùng biến save nếu bằng True sẽ lưu các đặc trưng vào file json.

* 1. Ứng dụng