

TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA CÔNG NGHỆ THÔNG TIN



Mai Hoàng Hiệp
Trần Trung Hiếu
Nguyễn Duy Đạt

BÁO CÁO
TÌM HIỂU VÀ NGHIÊN CỨU VỀ
THƯ VIỆN OPENCV

CHƯƠNG TRÌNH CHÍNH QUY

Tp. Thủ Đức, tháng 11/2024

TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA CÔNG NGHỆ THÔNG TIN

Mai Hoàng Hiệp - 24122015
Trần Trung Hiếu - 24122033
Nguyễn Duy Đạt - 24122014

BÁO CÁO
TÌM HIỂU VÀ NGHIÊN CỨU VỀ
THƯ VIỆN OPENCV

DỒ ÁN MÔN LÝ THUYẾT GIỚI THIỆU
NGÀNH TRÍ TUỆ NHÂN TẠO
CHƯƠNG TRÌNH CHÍNH QUY

GIÁO VIÊN HƯỚNG DẪN
ThS. Nguyễn Trần Duy Minh

Tp. Thủ Đức, tháng 11/2024

Lời cảm ơn

Chúng em xin chân thành cảm ơn đến thầy Phạm Trọng Nghĩa vì những kiến thức bổ ích và sự tận tâm mà thầy đã dành cho chúng tôi trong suốt quá trình học môn Giới thiệu ngành Trí tuệ nhân tạo. Nhờ những bài giảng sinh động và ví dụ thực tế của thầy, nhóm chúng tôi đã có cái nhìn sâu sắc hơn về các thuật toán học máy, mạng thần kinh nhân tạo và cách chúng được ứng dụng trong thực tế. Đặc biệt, phần về thị giác máy tính và đồ án về OpenCV đã giúp khám phá ra tiềm năng của AI trong việc giải quyết các bài toán thực tế,...

Mục lục

Lời cảm ơn	i
Đề cương chi tiết	ii
Mục lục	ii
Tóm tắt	iv
1 Giới thiệu về thư viện OpenCV	1
1.1 Thư viện OpenCV	1
1.2 Mục đích sử dụng của OpenCV là gì?	2
1.2.1 Khả năng xử lý hình ảnh và video	2
1.2.2 Hỗ trợ học máy	2
1.2.3 Tối ưu hóa hiệu suất	2
1.3 Khái niệm về Computer Vision	3
1.3.1 Định nghĩa	3
1.3.2 Cách thức hoạt động	4
1.3.3 Vai trò của thị giác máy tính	4
1.4 Đối tượng sử dụng	5
1.5 Ứng dụng của OpenCV	6
2 Hướng dẫn sử dụng thư viện OpenCV	8
2.1 Yêu cầu cơ bản	8
2.2 Sơ lược về các hàm cơ bản trong OpenCV	8
2.3 Sử dụng OpenCV không mã hóa	9
2.4 Vision capabilities of OpenCV	11
2.4.1 Real-time Object Detection	11
2.4.2 Image Segmentation: Phân loại IMG	12
2.4.3 Movement and Gesture Recognition: nhận diện hành động cử chỉ	12

2.4.4	Face Recognition	13
2.4.5	Thực tế tăng cường	14
3	Các kỹ thuật thị giác máy tính nâng cao	15
3.1	Phát hiện và mô tả tính năng chính	15
3.2	Kỹ thuật phát hiện đối tượng	15
3.3	Phân đoạn hình ảnh và phân tích chuyển động	16
3.3.1	Phân đoạn hình ảnh	16
3.3.2	Phân tích chuyển động	16
3.4	Hình thái học và tăng cường hình ảnh	17
3.4.1	Khái niệm các phép toán hình thái học	17
3.4.2	Các phép toán hình thái học cơ bản:	17
3.4.3	Ứng dụng của các phép toán hình thái học:	18
3.4.4	Khái niệm lọc hình ảnh	18
3.4.5	Một số loại bộ lọc phổ biến:	18
3.4.6	Ứng dụng của lọc hình ảnh:	18
3.5	Hiệu chuẩn máy ảnh và tầm nhìn lập thể	19
3.5.1	Khái niệm Thị giác Stereo	19
3.5.2	Tại sao cần thị giác stereo?	19
3.5.3	Các bước cơ bản trong thị giác stereo:	19
3.6	Trừ nền và ghép ảnh	20
3.7	Kết luận	20
Tài liệu tham khảo		21
Đánh giá thành viên		22

Danh sách hình

1.1	Thư viện OpenCV	1
1.2	Khám phá về thị giác máy tính	3
1.3	Nhận diện khuôn mặt bằng OpenCV	5
1.4	Xe tự lái sử dụng công nghệ OpenCV	5
1.5	OpenCV chính là nền tảng của Google Lens.	5
1.6	Giải pháp giám sát đường bộ thông minh với OpenCV . .	6
1.7	Ứng dụng thị giác máy tính nhằm chẩn đoán bệnh trong y học	7
2.1	Chương trình đếm số ngón tay	9
2.2	Dễ dàng dự đoán hình ảnh với Teachable Machine	10
2.3	Phát hiện nhiều đối tượng dựa trên ML với OpenCV . . .	11
2.4	Phân đoạn hình ảnh	12
2.5	Phân tích chuyển động với ước tính tư thế bằng cách sử dụng các điểm chính	13
2.6	Nhận diện khuôn mặt	13
2.7	Thực tế tăng cường (AR)	14
3.1	Dòng quang học trong thực tế tăng cường với OpenCV . .	17

Chương 1

Giới thiệu về thư viện OpenCV

1.1 Thư viện OpenCV

OpenCV (Open Source Computer Vision Library) [3] là một thư viện mã nguồn mở phổ biến, được sử dụng rộng rãi trong các ứng dụng tầm nhìn máy tính thời gian thực. OpenCV cung cấp một bộ các hàm phong phú để xử lý hình ảnh và video, phát hiện vật thể, và học máy.

Được phát triển bởi Intel, OpenCV là một thư viện thị giác máy tính đa nền tảng miễn phí để xử lý hình ảnh theo thời gian thực. Phần mềm OpenCV đã trở thành một công cụ chuẩn thực tế cho mọi thử liên quan đến Thị giác máy tính.



Hình 1.1: Thư viện OpenCV

Thư viện OpenCV chứa hơn 2500 thuật toán, tài liệu mở rộng, mã nguồn và mã mẫu cho thị giác máy tính thời gian thực. Các nhà phát triển sử dụng gói Python và thư viện Python có thể tích hợp OpenCV vào

các dự án của họ bằng các lệnh như “import cv2”. Trình quản lý gói tạo điều kiện thuận lợi cho việc tích hợp này, cung cấp quy trình đơn giản để cài đặt và kiểm soát phiên bản.

1.2 Mục đích sử dụng của OpenCV là gì?

Mục tiêu chính của OpenCV là cung cấp một nền tảng thị giác máy tính đơn giản, giúp mọi người, kể cả những người mới bắt đầu, có thể dễ dàng xây dựng các ứng dụng thị giác tinh vi. OpenCV cung cấp hơn 500 hàm hỗ trợ, bao phủ hầu hết các lĩnh vực trong thị giác máy tính, từ kiểm tra sản phẩm, y tế, an ninh đến robot.

Các tác giả đã xây dựng OpenCV để đạt được hiệu quả và hiệu suất tối đa cho các tác vụ thị giác đòi hỏi nhiều tính toán. Do đó, nó tập trung mạnh vào các ứng dụng thị giác AI thời gian thực.

Sau đây là một số lĩnh vực chính mà OpenCV được ứng dụng:

1.2.1 Khả năng xử lý hình ảnh và video

OpenCV sở hữu khả năng xử lý hình ảnh toàn diện, hỗ trợ xử lý luồng video, ghép ảnh, hiệu chuẩn camera và nhiều tác vụ xử lý trước hình ảnh khác nhau. Điều này giúp OpenCV trở thành một công cụ hữu ích cho các ứng dụng đòi hỏi xử lý video thời gian thực như giám sát, phân tích hành vi, và thực tế ảo.

1.2.2 Hỗ trợ học máy

Để nâng cao khả năng nhận dạng và phân loại các đối tượng trong hình ảnh, thư viện tích hợp một thư viện học máy mạnh mẽ. Thư viện này tập trung vào các thuật toán học máy truyền thống như nhận dạng mẫu thông kê và phân cụm, giúp các nhà phát triển xây dựng các hệ thống thị giác thông minh hơn.

1.2.3 Tối ưu hóa hiệu suất

Để đáp ứng nhu cầu xử lý dữ liệu hình ảnh lớn và phức tạp trong thời gian thực, OpenCV đã được tối ưu hóa để tận dụng sức mạnh của GPU.

Kể từ năm 2011, nó đã hỗ trợ NVIDIA CUDA, giúp tăng tốc đáng kể quá trình xử lý hình ảnh. Điều này đặc biệt quan trọng đối với các ứng dụng đòi hỏi thời gian phản hồi nhanh như các hệ thống tự lái xe hoặc robot.

1.3 Khái niệm về Computer Vision

1.3.1 Định nghĩa

Computer Vision hay thị giác máy tính là một dạng AI trong đó máy tính thực hiện các hành động cụ thể dựa trên thông tin được trích xuất từ phương tiện kỹ thuật số, thế giới thực. Mặc dù có nhiều định nghĩa khác nhau về Trí tuệ nhân tạo (AI), nhưng tất cả các khái niệm thường tập trung vào việc sao chép trí thông minh của con người bằng máy móc. Nói cách khác, chúng ta đang trao cho máy tính khả năng "nhìn thấy".



Hình 1.2: Khám phá về thị giác máy tính

1.3.2 Cách thức hoạt động

Hệ thống thị giác máy có mục đích mô phỏng thị giác của con người và diễn giải hình ảnh hoặc video kỹ thuật số để hiểu và nhận dạng các vật thể và cảnh vật.

Các hệ thống này bao gồm sự kết hợp giữa phần mềm và phần cứng mô phỏng hoạt động của hệ thống thị giác của con người. Một số phương pháp phổ biến được sử dụng trong thị giác máy là nhận dạng mẫu AI , trích xuất đặc điểm và xử lý hình ảnh.

Máy tính hoạt động như “bộ não” thực hiện các tác vụ tính toán nâng cao và áp dụng các thuật toán phức tạp. “Bộ não” này phân tích phương tiện trực quan do các cảm biến quang học hoặc máy ảnh cung cấp. Thuật toán thị giác AI có thể sử dụng hầu như bất kỳ hình ảnh nào của bất kỳ máy ảnh nào .

1.3.3 Vai trò của thị giác máy tính

Từ việc đơn thuần nhận dạng hình ảnh, thị giác máy tính đã phát triển thành một công cụ mạnh mẽ, thúc đẩy quá trình tự động hóa và số hóa sản xuất. Nhờ những đột phá trong học máy, máy tính ngày càng "thông minh" hơn trong việc hiểu và giải thích thông tin hình ảnh, mở ra những tiềm năng ứng dụng vô hạn.

Sự kết hợp giữa AI và IoT, cùng với sự phát triển của điện toán đám mây và Edge AI, đã tạo ra một cuộc cách mạng trong lĩnh vực tầm nhìn máy tính. Nhờ đó, máy tính có thể thực hiện nhiều tác vụ phức tạp, từ việc nhận dạng khuôn mặt, đọc chữ viết tay, đến việc kiểm tra sản phẩm tự động và phân tích hành vi khách hàng. Điều này mở ra những tiềm năng ứng dụng vô hạn trong nhiều lĩnh vực, từ an ninh, sản xuất đến y tế.



Hình 1.3: Nhận diện khuôn mặt bằng OpenCV

1.4 Đối tượng sử dụng

OpenCV đã trở thành nền tảng cho nhiều ứng dụng thị giác máy tính đột phá. Các tập đoàn công nghệ hàng đầu như Google, Microsoft và các tổ chức nghiên cứu danh tiếng như Stanford đều tin tưởng sử dụng OpenCV.



Hình 1.4: Xe tự lái sử dụng công nghệ OpenCV



Hình 1.5: OpenCV chính là nền tảng của Google Lens.

Từ việc nhận dạng khuôn mặt trên các nền tảng mạng xã hội, đến việc điều khiển robot trong các nhà máy sản xuất, OpenCV đã chứng minh được khả năng ứng dụng rộng rãi và đa dạng. Nhờ OpenCV, các doanh nghiệp có thể tự động hóa các quy trình, nâng cao hiệu quả sản xuất và tạo ra những sản phẩm thông minh hơn, góp phần thay đổi cuộc sống của chúng ta.

Các trường hợp sử dụng của thị giác máy tính rất rộng. Từ an ninh và giám sát video hoặc xe tự lái đến các trường hợp cụ thể trong ngành như sản xuất công nghiệp, nhà hàng hoặc phân tích bán lẻ,...

1.5 Ứng dụng của OpenCV

Hầu hết mọi người đều biết thị giác máy tính được sử dụng rộng rãi trong lĩnh vực an ninh và giám sát. Một số người cũng biết đến các ứng dụng nhỏ hơn như giám sát an toàn, máy bay không người lái (drone) hoặc phân tích sinh học. Tuy nhiên, ít người nhận ra tầm quan trọng của thị giác máy tính trong sản xuất công nghiệp. Hệ thống thị giác máy tính đã kiểm tra gần như mọi thứ được sản xuất hàng loạt.



Hình 1.6: Giải pháp giám sát đường bộ thông minh với OpenCV

Một số lĩnh vực tiêu biểu có thể kể đến như [2]:

- Phân tích hình ảnh y tế để hỗ trợ chẩn đoán bệnh ở người (Medical Imaging)
- Nhận dạng quảng cáo trong cảnh quay TV hoặc nhận dạng logo bằng thị giác AI (AI Vision)
- Tìm kiếm hình ảnh trên nền tảng kỹ thuật số, trong các ứng dụng dựa trên web
- Phát hiện lỗi hoặc sự cố trong quá trình sản xuất (Smart Factory)
- Đếm số người ở những nơi công cộng như sân bay (Phân tích đám đông),...



Hình 1.7: Ứng dụng thị giác máy tính nhằm chẩn đoán bệnh trong y học

Chương 2

Hướng dẫn sử dụng thư viện OpenCV

2.1 Yêu cầu cơ bản

Cài đặt OpenCV là bước khởi đầu không thể thiếu nếu ta muốn bắt tay vào các dự án liên quan đến xử lý hình ảnh, video bằng Python. Để bắt đầu, ta cần một lệnh đơn giản trong Command Prompt (cmd): "pip install opencv-python", ta đã có thể sở hữu một công cụ mạnh mẽ để thực hiện các tác vụ như nhận diện khuôn mặt, theo dõi vật thể, phân tích hình ảnh,...

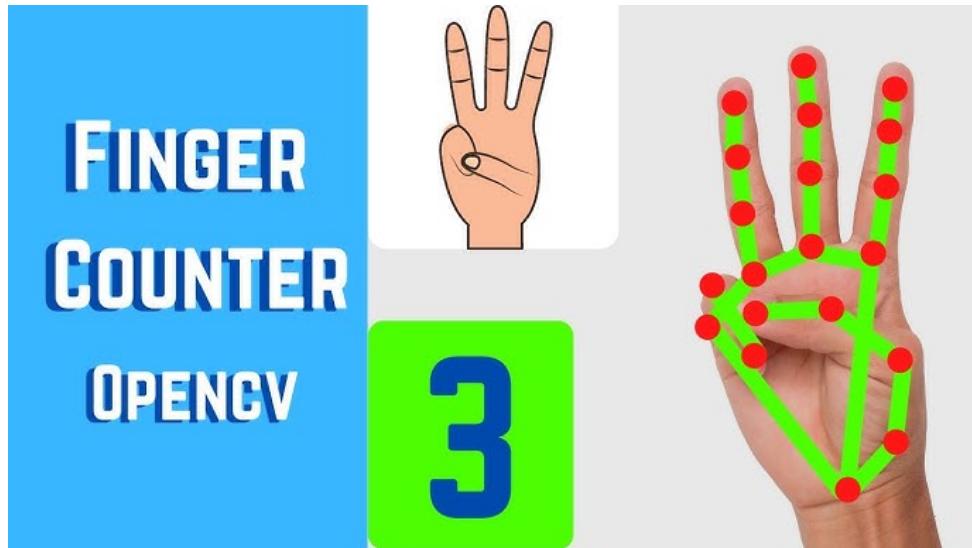
2.2 Sơ lược về các hàm cơ bản trong OpenCV

Để có cái nhìn tổng quan về các câu lệnh thường dùng, chúng ta sẽ xem xét đoạn code dưới đây:

```
1 import cv2 # Khai bao thu vien OpenCV
2
3 # Doc hinh anh
4 img = cv2.imread('anh_cua_ban.jpg')
5 # Giam kich thuoc anh xuong con mot nua
6 resized_img = cv2.resize(img, None, fx=0.5, fy=0.5)
7 # Xoay hinh anh
8 rotate_img = cv2.rotate(img, cv2.ROTATE_90_CLOCKWISE) #Xoay theo kim dong ho
9 # Hien thi hinh anh
10 cv2.imshow('Anh cua ban', img)
11 # Nhan phim bat ki de dong cua so
12 cv2.waitKey(0)
13 # Giai phong bo nho
14 cv2.destroyAllWindows()
```

Listing 2.1: Các câu lệnh cơ bản của thư viện OpenCV

Để có cái nhìn sâu hơn về các ứng dụng thực tế của OpenCV và khám phá thêm nhiều ví dụ mã nguồn thú vị, chúng ta hoàn toàn có thể tham khảo những tài liệu, chương trình mẫu mà bạn đã chia sẻ ở tài liệu tham khảo [5]. Việc nghiên cứu các dự án cụ thể không chỉ giúp chúng ta hiểu rõ hơn về cách các thuật toán được áp dụng trong thực tế mà còn cung cấp một kho tàng ý tưởng để phát triển các ứng dụng thị giác máy tính của riêng mình [4].



Hình 2.1: Chương trình đếm số ngón tay

2.3 Sử dụng OpenCV không mã hóa

Ta có thể dễ dàng thực hiện các tác vụ như nhận diện khuôn mặt, theo dõi vật thể, hay xử lý ảnh đơn giản chỉ bằng cách sử dụng các công cụ trực quan hoặc các nền tảng đã tích hợp sẵn các chức năng của OpenCV. Chẳng hạn như:

- Sử dụng các mô hình đã được huấn luyện sẵn:
 - Các mô hình của OpenCV: thư viện cung cấp sẵn nhiều mô hình đã được huấn luyện để thực hiện các tác vụ như nhận diện khuôn mặt, phát hiện vật thể, và phân loại hình ảnh. Ta có thể dễ dàng tích hợp những mô hình này vào ứng dụng của mình.
 - Các thư viện và khung khác: Các thư viện như TensorFlow, PyTorch, và Keras cũng cung cấp rất nhiều mô hình đã được

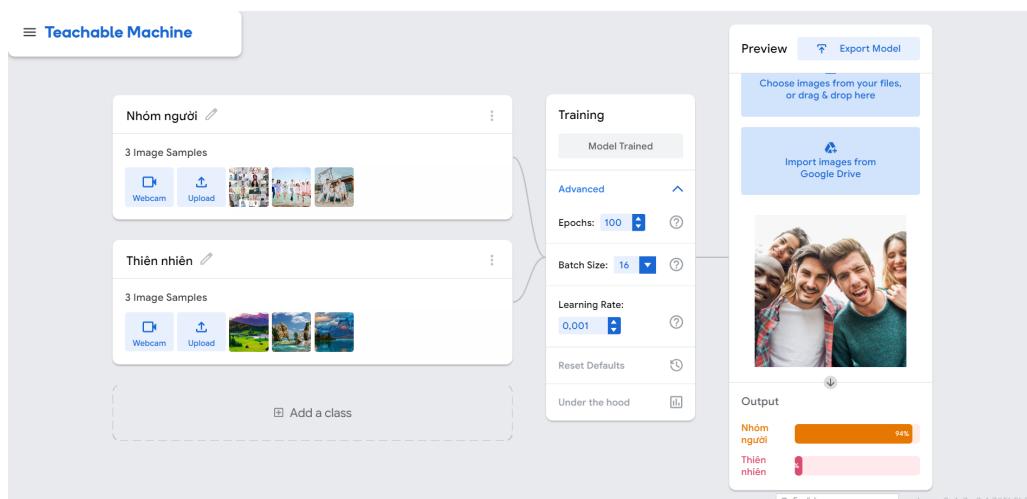
huấn luyện sẵn cho các nhiệm vụ thị giác máy tính. Ta có thể kết hợp chúng với OpenCV để mở rộng khả năng.

- Các nền tảng không cần hoặc ít cần code:

- Google Teachable Machine: Nền tảng này cho phép bạn huấn luyện các mô hình phân loại hình ảnh tùy chỉnh mà không cần viết code. Bạn có thể sử dụng các mô hình này trong ứng dụng của mình, và thậm chí tích hợp với OpenCV để xử lý thêm.
- Các nền tảng khác: Có nhiều nền tảng khác cung cấp giao diện kéo thả để tạo các ứng dụng thị giác máy tính, thường sử dụng OpenCV làm nền tảng.

- Các dịch vụ đám mây:

- Google Cloud Vision API: Cung cấp các mô hình đã được huấn luyện sẵn cho nhiều tác vụ như dán nhãn hình ảnh, nhận diện khuôn mặt, nhận dạng ký tự quang học (OCR),...
- Amazon Rekognition: Cung cấp các dịch vụ tương tự như Google Cloud Vision API, bao gồm phân tích hình ảnh và video.
- Microsoft Azure Computer Vision: Cung cấp một loạt các dịch vụ thị giác máy tính, bao gồm phân tích hình ảnh, phát hiện vật thể, và nhận diện khuôn mặt.



Hình 2.2: Dễ dàng dự đoán hình ảnh với Teachable Machine

Mặc dù không cần viết code, nhưng để sử dụng tốt và hiệu quả các công cụ và dịch vụ này vẫn đòi hỏi ta hiểu rõ về các khái niệm cơ bản của thị giác máy tính. Điều này sẽ giúp mình lựa chọn công cụ phù hợp và giải quyết các vấn đề phát sinh.

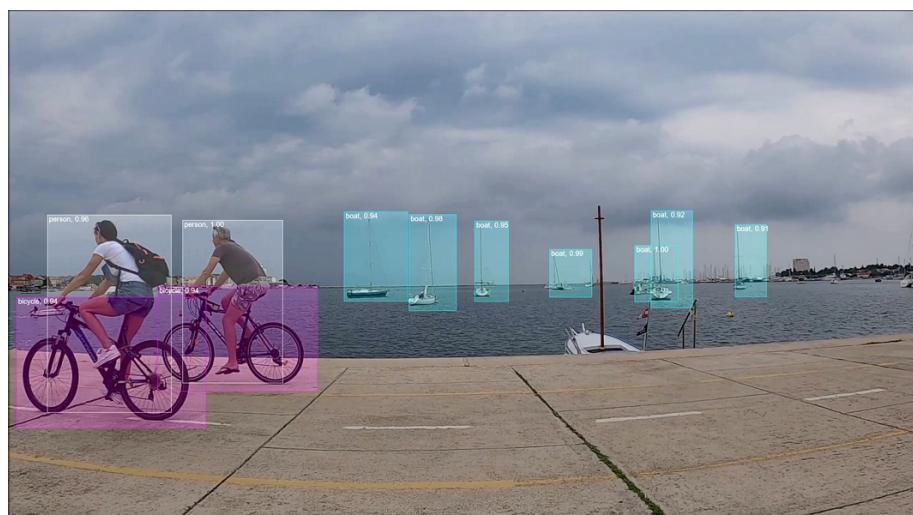
OpenCV vốn được thiết kế để linh hoạt và mạnh mẽ khi kết hợp với việc viết code. Tuy nhiên, với sự phát triển của công nghệ, ngày càng có nhiều công cụ và dịch vụ cho phép chúng ta tận dụng sức mạnh của OpenCV mà không cần phải dùng đến một dòng code nào, nhưng phạm vi ứng dụng sẽ hạn chế hơn so với việc viết code trực tiếp.

2.4 Vision capabilities of OpenCV

OpenCV [1] có lẽ là công cụ thị giác máy tính linh hoạt nhất được sử dụng trong nhiều lĩnh vực nhiệm vụ thị giác máy tính. Các nhiệm vụ này bao gồm từ nhận dạng hình ảnh và phân tích 2D hoặc 3D đến theo dõi chuyển động, nhận dạng khuôn mặt, v.v. Những khả năng nổi bật nhất:

2.4.1 Real-time Object Detection

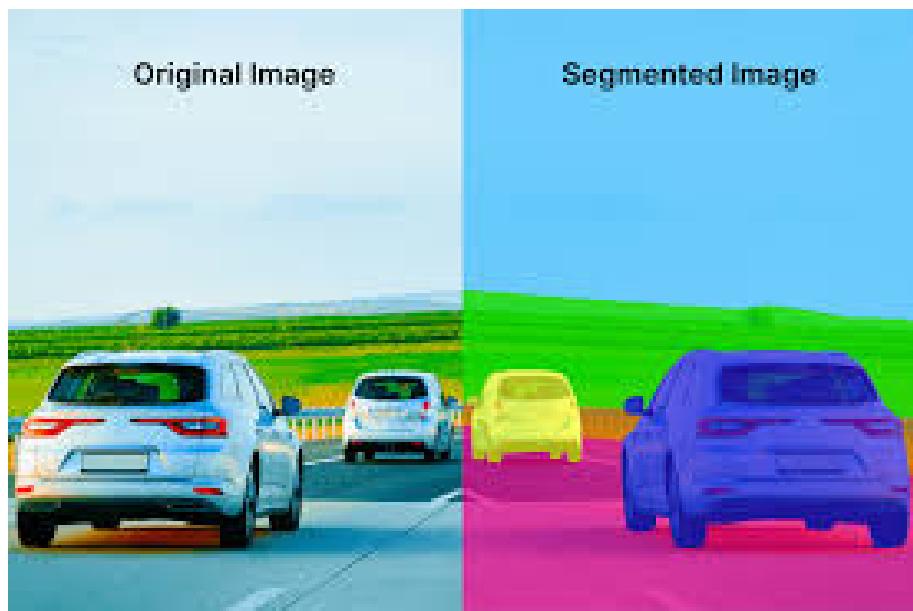
Công nghệ phát hiện đối tượng áp dụng nhận dạng hình ảnh và định vị các đối tượng cụ thể trong dữ liệu video hoặc hình ảnh. Các đối tượng này có thể là bất kỳ thứ gì: ô tô, con người và động vật hoặc thiết bị trong sản xuất công nghiệp.



Hình 2.3: Phát hiện nhiều đối tượng dựa trên ML với OpenCV

2.4.2 Image Segmentation: Phân loại IMG

Phân đoạn hình ảnh áp dụng các thuật toán xử lý hình ảnh để chia hình ảnh thành các phân đoạn khác nhau. Phân đoạn thường đơn giản hóa, thay đổi hoặc tăng cường hình ảnh, thường kết hợp với các tác vụ thị giác máy tính tiếp theo. Một ví dụ là lái xe tự động, sử dụng phân đoạn hình ảnh để xác định đường.



Hình 2.4: Phân đoạn hình ảnh

2.4.3 Movement and Gesture Recognition: nhận diện hành động cử chỉ

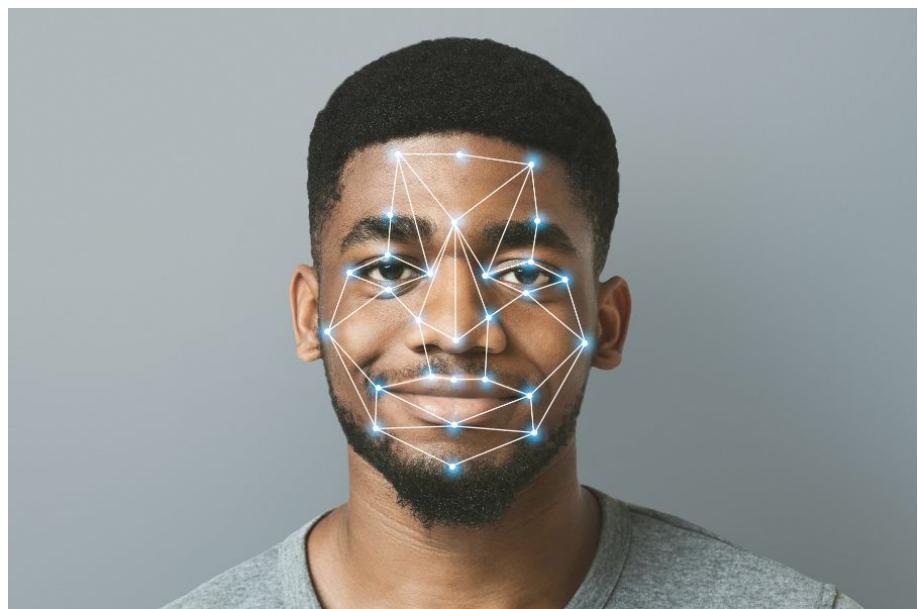
Nhận dạng tư thế và cử chỉ của con người diễn giải và hiểu được cử chỉ của con người thông qua phân tích video. Nó nhận dạng và phân loại các chuyển động của cơ thể, bàn tay hoặc khuôn mặt để gán cho mỗi chuyển động một danh mục được xác định trước. Phân tích chuyển động thường là một phần của ước tính tư thế để phân tích các chuyển động của cơ thể với các điểm chính tham chiếu (khớp, chân tay). Tính toán tư thế của vật thể truyền tải cách định hướng của vật thể trong không gian 3D



Hình 2.5: Phân tích chuyển động với ước tính tư thế bằng cách sử dụng các điểm chính

2.4.4 Face Recognition

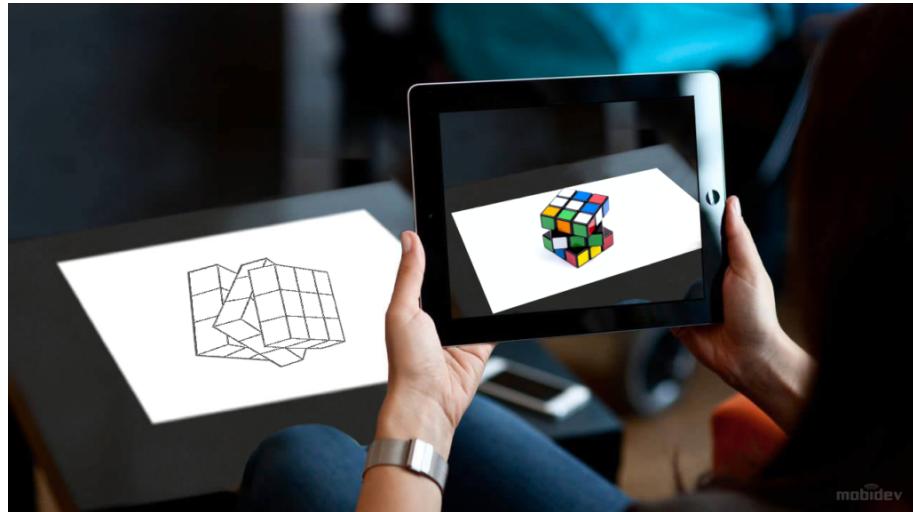
Nhận dạng khuôn mặt tự động xác định khuôn mặt người và khớp với khuôn mặt đó trong cơ sở dữ liệu đặc điểm khuôn mặt. FaceRecognizer của OpenCV cung cấp một bộ thuật toán nhận dạng khuôn mặt phổ biến để sử dụng trong các ứng dụng thực tế.



Hình 2.6: Nhận diện khuôn mặt

2.4.5 Thực tế tăng cường

Thực tế tăng cường (AR) cho phép tương tác thời gian thực giữa thế giới thực và thế giới ảo. Do đó, thực tế tăng cường nhằm mục đích tăng cường thế giới vật lý xung quanh chúng ta bằng thông tin nhận thức do máy tính tạo ra.



Hình 2.7: Thực tế tăng cường (AR)

Chương 3

Các kỹ thuật thị giác máy tính nâng cao

OpenCV cung cấp những kỹ thuật phức tạp hơn, vượt xa những chức năng cơ bản.

3.1 Phát hiện và mô tả tính năng chính

- SIFT (Scale-Invariant Feature Transform). SIFT là một thuật toán OpenCV để phát hiện và mô tả các tính năng chính trong hình ảnh. Bản chất scale-invariant của SIFT có của đối tượng. Do đó, nó có hiệu quả trong việc nhận dạng đối tượng và ghép ảnh.
- SURF (Tính năng mạnh mẽ được tăng tốc). SURF là một thuật toán khác để khả năng phục hồi trước các biến thể về kích thước và hướng phát hiện điểm chính và mô tả tính năng hình ảnh. Dựa trên SIFT, SURF cung cấp tốc độ tính toán tăng lên, hữu ích cho các ứng dụng thời gian thực.

3.2 Kỹ thuật phát hiện đối tượng

OpenCV cung cấp các mô hình được đào tạo trước để phát hiện đối tượng. Các mô hình được xây dựng trước đơn giản hóa việc triển khai mô hình, cho phép người dùng bắt đầu mà không cần cấu hình thủ công nhiều.

Haar Cascades, được triển khai thông qua `cv2.CascadeClassifier()`, hữu ích trong việc phát hiện các mẫu cụ thể như khuôn mặt và mắt. Phương pháp dựa trên học sâu. OpenCV tích hợp với các mô hình học sâu

phổ biến: Single Shot Multibox Detector (SSD) và You Only Look Once (YOLO).

3.3 Phân đoạn hình ảnh và phân tích chuyển động

3.3.1 Phân đoạn hình ảnh

Quá trình chia một hình ảnh thành các vùng đồng nhất, mỗi vùng đại diện cho một đối tượng hoặc một phần của đối tượng trong hình ảnh. Mục tiêu của phân đoạn là đơn giản hóa và làm rõ hình ảnh, giúp cho các quá trình xử lý hình ảnh tiếp theo trở nên dễ dàng hơn.

3.3.2 Phân tích chuyển động

Quá trình tìm hiểu và mô tả các thay đổi trong một chuỗi hình ảnh theo thời gian. Nó bao gồm việc phát hiện các đối tượng đang chuyển động, theo dõi chuyển động của chúng và ước tính tốc độ, hướng di chuyển.

- **Kỹ thuật phân đoạn.** Bạn có thể tìm hiểu sâu hơn về phân đoạn hình ảnh bằng các kỹ thuật như GrabCut và MeanShift. Chúng tôi dựa trên màu sắc, kết cấu và cường độ của các hình ảnh được phân tách này.
- **Dòng quang học .** Thuật toán Lucas-Kanade và Farneback theo dõi chuyển động giữa các khung hình video liên tiếp bằng cách tính toán độ dịch chuyển pixel theo thời gian. Từ đó có thể ứng dụng vào việc theo dõi chuyển động của người và vật thể hoặc sửa chữa, khôi phục các video bị hư hỏng



Hình 3.1: Dòng quang học trong thực tế tăng cường với OpenCV

3.4 Hình thái học và tăng cường hình ảnh

3.4.1 Khái niệm các phép toán hình thái học

Một tập hợp các kỹ thuật xử lý ảnh phi tuyến tính được sử dụng để thay đổi hình dạng và cấu trúc của các đối tượng trong một hình ảnh nhị phân. Những phép toán này được sử dụng rộng rãi để loại bỏ nhiễu, làm mịn hình ảnh, trích xuất các đặc trưng hình thái của đối tượng và phân tích hình ảnh.

3.4.2 Các phép toán hình thái học cơ bản:

- **Dãn (Dilation):** Mở rộng các vùng sáng trong ảnh, thường được sử dụng để lấp đầy các lỗ hổng nhỏ.
- **Co (Erosion):** Thu hẹp các vùng sáng trong ảnh, thường được sử dụng để loại bỏ các điểm nhiễu nhỏ.
- **Mở (Opening):** Kết hợp phép co và phép dãn, giúp loại bỏ nhiễu và làm mịn các đường viền.

- **Đóng (Closing):** Kết hợp phép dãn và phép co, giúp lấp đầy các lỗ hổng trong các đối tượng.

3.4.3 Ứng dụng của các phép toán hình thái học:

- **Xử lý ảnh:** Loại bỏ nhiễu, làm mịn hình ảnh, trích xuất đặc trưng.
- **Phân tích hình ảnh:** Phát hiện cạnh, phân đoạn hình ảnh.
- **Xử lý văn bản:** Nhận dạng ký tự quang học.

3.4.4 Khái niệm lọc hình ảnh

một kỹ thuật xử lý ảnh cơ bản, được sử dụng để cải thiện chất lượng hình ảnh và trích xuất các đặc trưng quan trọng. Các bộ lọc khác nhau được áp dụng để đạt được các mục đích khác nhau.

3.4.5 Một số loại bộ lọc phổ biến:

- **Bộ lọc Gaussian:** Làm mờ hình ảnh bằng cách giảm nhiễu và làm mịn các chi tiết nhỏ.
- **Bộ lọc Sobel:** Nhận diện các cạnh trong hình ảnh bằng cách tính toán độ dốc.
- **Bộ lọc Laplacian:** Phát hiện các vùng thay đổi cường độ nhanh, nhấn mạnh các chi tiết nhỏ.

3.4.6 Ứng dụng của lọc hình ảnh:

- **Xử lý ảnh:** Loại bỏ nhiễu, tăng cường độ tương phản, làm mịn hình ảnh.
- **Phân tích hình ảnh:** Phát hiện cạnh, phân đoạn hình ảnh, nhận dạng đối tượng.
- **Xử lý video:** Tăng cường chất lượng video, theo dõi chuyển động.

3.5 Hiệu chuẩn máy ảnh và tầm nhìn lập thể

Hiệu chuẩn máy ảnh ước tính các thông số máy ảnh bên trong và bên ngoài để sửa lỗi méo và cải thiện độ chính xác của phép đo hình ảnh. `cv2.calibrateCamera()`. Chức năng này cho phép hiệu chuẩn bằng cách chụp ảnh của một mẫu hiệu chuẩn từ các góc nhìn khác nhau. Các thông số, chẳng hạn như tiêu cự và hệ số méo, cho phép tái tạo 3D chính xác và hiệu chỉnh phối cảnh chính xác. Điều này hữu ích trong các ứng dụng như thực tế tăng cường.

3.5.1 Khái niệm Thị giác Stereo

Một lĩnh vực trong thị giác máy tính mô phỏng khả năng nhìn nhận chiều sâu của con người. Thay vì chỉ nhận được thông tin từ một camera, thị giác stereo sử dụng ít nhất hai camera để tạo ra hiệu ứng thị giác tương tự như mắt người, từ đó giúp máy tính ước tính được khoảng cách đến các vật thể trong không gian 3D.

3.5.2 Tại sao cần thị giác stereo?

- **Tạo ra bản đồ độ sâu:** Giúp máy tính hiểu được cấu trúc 3D của môi trường xung quanh.
- **Nhận biết vật thể:** Phân biệt các vật thể dựa trên hình dạng và khoảng cách.
- **Điều hướng:** Định vị và điều khiển robot, xe tự lái.
- **Thực tế ảo và tăng cường:** Tạo ra các trải nghiệm tương tác thực tế hơn.

3.5.3 Các bước cơ bản trong thị giác stereo:

1. **Thu thập hình ảnh:** Sử dụng hai hoặc nhiều camera để chụp cùng một cảnh từ các góc độ khác nhau.

2. **Tìm điểm tương ứng:** Xác định các điểm tương ứng giữa hai hình ảnh, tức là các điểm đại diện cho cùng một điểm trong không gian 3D trên cả hai hình ảnh.
3. **Tính toán độ sâu:** Sử dụng thông tin về vị trí tương đối của các camera và các điểm tương ứng để tính toán khoảng cách đến các điểm trong không gian 3D.

3.6 Trừ nền và ghép ảnh

- **Trừ nền (Background Subtraction):** Là một kỹ thuật trong xử lý hình ảnh nhằm tách biệt đối tượng chuyển động khỏi nền tĩnh. Bằng cách loại bỏ phần nền, chúng ta có thể tập trung vào việc phân tích và theo dõi các đối tượng đang di chuyển trong một cảnh.
- **Image Stitching hay Ghép nối hình ảnh:** Là một kỹ thuật trong xử lý hình ảnh nhằm kết hợp nhiều hình ảnh riêng lẻ thành một hình ảnh lớn hơn, tạo ra một bức tranh toàn cảnh. Điều này đặc biệt hữu ích khi bạn muốn chụp một cảnh quá rộng để có thể nắm bắt được toàn bộ trong một khung hình duy nhất.

OpenCV cung cấp các chức năng để khớp tính năng, ước tính đồng dạng và chuyển đổi phối cảnh để căn chỉnh và pha trộn ảnh. Đánh giá có thể bao gồm kiểm tra trực quan về chất lượng đường may và đánh giá độ chính xác hình học, đảm bảo rằng kết quả gắn kết và không bị biến dạng.

3.7 Kết luận

OpenCV là một công cụ mạnh mẽ và linh hoạt trong xử lý hình ảnh và thị giác máy tính. Nó cung cấp giải pháp cho rất nhiều vấn đề trong các lĩnh vực khác nhau như các hệ thống bảo mật, y khoa hay các hệ thống sản xuất hàng hoá tự động. Với tính mã nguồn mở, cộng đồng phát triển mạnh mẽ và tài liệu học tập phong phú, OpenCV là một lựa chọn hàng đầu cho các dự án về thị giác máy tính.

Tài liệu tham khảo

Tiếng Việt

- [1] Fraidoon Omarzai. *OpenCV In Depth*. URL: <https://medium.com/@fraidoonmarzai99/opencv-in-depth-431a7b0d3e9c>.
- [2] Online. *Getting Started with OpenCV*. URL: <https://learnopencv.com/getting-started-with-opencv/> (visited on 2021).
- [3] Online. *What is OpenCV? The Complete Guide (2025)*. URL: <https://viso.ai/computer-vision/opencv/>.
- [4] Gà Lại Lập Trình. *OpenCV Python*. URL: <https://www.youtube.com/playlist?list=PLPt6-BtUI22ogXrY1V7aZS9v14ZESQLUH>.
- [5] Nhóm 6 24TNT1 Trường Đại học Khoa học tự nhiên ĐHQG HCM. *Hướng dẫn các chương trình demo*. URL: <https://github.com/trunghieu06/OpenCV-Project> (visited on 2024).

Đánh giá thành viên

Phân chia công việc và mức độ hoàn thành của từng yêu cầu		
Tên thành viên	Công việc	Mức độ hoàn thành (%)
Trần Trung Hiếu (MSSV - 24122033)	Source code, Thuyết trình + demo code, Slides	100%
Mai Hoàng Hiệp (MSSV - 24122015)	Source code, Slides, File báo cáo	100%
Nguyễn Duy Đạt (MSSV - 24122014)	Slides, Thuyết trình, File báo cáo	100%