|  |  |
| --- | --- |
|  |  |
| NGUYỄN TIẾN DUY | **BỘ CÔNG THƯƠNG**  **TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP HÀ NỘI**  **---------------------------------------** |
|  |
| ĐỒ ÁN TỐT NGHIỆP  NGÀNH KỸ THUẬT PHẦN MỀM |
|  |
| **XÂY DỰNG ỨNG DỤNG ĐO CHIỀU CAO CƠ THỂ NGƯỜI SỬ DỤNG THỊ GIÁC MÁY TÍNH** |
|  |
|  |
| **CBHD : ThS. Nguyễn Trung Phú** |
| KỸ THUẬT PHẦN MỀM | **Sinh viên : Nguyễn Tiến Duy** |
| **Mã số sinh viên: 2020601607** |
|  |
|  |
|  |
|  |
| Hà Nội – Năm 2024 |

A document with signature on it

Description automatically generated

MỤC LỤC

[DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ VIẾT TẮT i](#_Toc168264946)

[DANH MỤC HÌNH ẢNH ii](#_Toc168264947)

[DANH MỤC BẢNG BIỂU iii](#_Toc168264948)

[LỜI CẢM ƠN iv](#_Toc168264949)

[MỞ ĐẦU 1](#_Toc168264950)

[CHƯƠNG 1. CƠ SỞ LÝ THUYẾT 3](#_Toc168264951)

[1.1. Thị giác máy tính 3](#_Toc168264952)

[1.1.1. Định nghĩa 3](#_Toc168264953)

[1.1.2. Các lĩnh vực liên quan 3](#_Toc168264954)

[1.2. Mô hình dự đoán tư thế (Pose estimate) 4](#_Toc168264955)

[1.2.1. Giới thiệu 4](#_Toc168264956)

[1.2.2. Ứng dụng của mô hình dự đoán tư thế 5](#_Toc168264957)

[1.2.3. Các loại mô hình dự đoán tư thế 6](#_Toc168264958)

[1.3. OpenCV và MediaPipe 8](#_Toc168264959)

[1.3.1. OpenCV 8](#_Toc168264960)

[1.3.2. MediaPipe 9](#_Toc168264961)

[1.4. Học máy 13](#_Toc168264962)

[1.4.1. Định nghĩa 13](#_Toc168264963)

[1.4.2. Phân loại 13](#_Toc168264964)

[1.4.3. Mô hình hồi quy tuyến tính 14](#_Toc168264965)

[1.5. Hình học trong không gian 16](#_Toc168264966)

[1.5.1. Hình học trong không gian 2D 16](#_Toc168264967)

[1.5.2. Hình học trong không gian 3D 17](#_Toc168264968)

[1.6. Một số cách đo chiều cao hiện có 18](#_Toc168264969)

[1.6.1. Cách đo chiều cao truyền thống 18](#_Toc168264970)

[1.6.2. Cách đo chiều cao không tiếp xúc đã có 21](#_Toc168264971)

[CHƯƠNG 2. XÂY DỰNG ỨNG DỤNG 26](#_Toc168264972)

[2.1. Tổng quan quy trình 26](#_Toc168264973)

[2.2. Các bước thực hiện 26](#_Toc168264974)

[2.2.1. Thu thập và xử lý dữ liệu dữ liệu 26](#_Toc168264975)

[2.2.2. Tính toán độ dài các đoạn khung xương 29](#_Toc168264976)

[2.2.3. Dự đoán chiều cao 31](#_Toc168264977)

[CHƯƠNG 3. KẾT QUẢ 34](#_Toc168264978)

[3.1. Kết quả các phương pháp dự đoán chiều cao 34](#_Toc168264979)

[3.1.1. Kết quả phương pháp cộng trung bình từ mũi đến đỉnh đẩu 34](#_Toc168264980)

[3.1.2. Kết quả phương pháp tỉ lệ mặt 37](#_Toc168264981)

[3.1.3. Kết quả phương pháp huấn luyện mô hình 39](#_Toc168264982)

[3.1.4. So sánh và đánh giá giữa các phương pháp dự đoán 40](#_Toc168264983)

[3.2. Giao diện ứng dụng 42](#_Toc168264984)

[3.3. Các kết quả khác 43](#_Toc168264985)

[3.3.1. Thực hiện dự đoán chiều cao đối với các tư thế khác nhau 43](#_Toc168264986)

[3.3.2. Thực hiện tính toán với không gian 3D 44](#_Toc168264987)

[KẾT LUẬN 50](#_Toc168264988)

[TÀI LIỆU THAM KHẢO 52](#_Toc168264989)

DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ VIẾT TẮT

|  |  |  |
| --- | --- | --- |
| **STT** | **Viết tắt** | **Dịch nghĩa** |
| 1 | 2D | 2 Dimensional |
| 2 | 3D | 3 Dimensional |
| 3 | AI | Artificial Intelligence |
| 4 | AR | Augmented Reality |
| 5 | CNN | Convolutional Neural Network |
| 6 | ID | Identity |
| 7 | OLS | Ordinary Least Squares |
| 8 | ROI | Region of Interest |

DANH MỤC HÌNH ẢNH

[Hình 1.1 Dự đoán tư thế của một người đang di chuyển và xác định các khớp trên cơ thể 4](#_Toc168265348)

[Hình 1.2 Các mốc cơ thể người mà Mediapipe xác định 10](#_Toc168265349)

[Hình 1.3 Dữ liệu tư thế một người đứng thẳng giơ tay trong không gian 3D 11](#_Toc168265350)

[Hình 1.4 Phương pháp đo chiều cao với thước thẳng 19](#_Toc168265351)

[Hình 1.5 Thước đo chiều cao y tế dán tường 20](#_Toc168265352)

[Hình 1.6 Biểu đồ phương pháp đo chiều cao sử dụng học sâu 22](#_Toc168265353)

[Hình 1.7 Tỉ lệ vàng khuôn mặt theo chiều dọc 23](#_Toc168265354)

[Hình 1.8 Trích xuất khuôn mặt để ước tính chiều cao 23](#_Toc168265355)

[Hình 1.9 Đo chiều cao bằng nhận diện đối tượng 25](#_Toc168265356)

[Hình 2.1 Sơ đồ khối các bước thực hiện 26](#_Toc168265357)

[Hình 2.2 Mô hình các thiết lập để lấy dữ liệu và xác định chiều cao 27](#_Toc168265358)

[Hình 2.3 Một số hình ảnh trong quá trình thu thập 28](#_Toc168265359)

[Hình 2.4 Sơ đồ các đoạn khung xương 29](#_Toc168265360)

[Hình 3.1 Đồ thị mô tả phương trình tuyết tính giữa hi và hr 36](#_Toc168265361)

[Hình 3.2 Đồ thị mô tả phương trình tuyến tính giữa hi và h8 38](#_Toc168265362)

[Hình 3.3 Giao diện ứng dụng 42](#_Toc168265363)

[Hình 3.4 Hình ảnh trong không gian 3D được vẽ từ tọa độ xyz trích xuất từ MediaPipe 47](#_Toc168265364)

DANH MỤC BẢNG BIỂU

[Bảng 3.1. 39 mẫu đối tượng để tính trung bình khoảng cách từ mũi đến đầu 34](#_Toc168265365)

[Bảng 3.2 10 mẫu đối tượng sử dụng để đánh giá phương pháp đề xuất 35](#_Toc168265366)

[Bảng 3.3 Kết quả ước tính chiều cao của 10 đối tượng thử 37](#_Toc168265367)

[Bảng 3.4 10 mẫu đối tượng được sử dụng để đánh giá phương pháp đề xuất 37](#_Toc168265368)

[Bảng 3.5 Kết quả đánh giá của 10 đối tượng thử nghiệm 38](#_Toc168265369)

[Bảng 3.6 20 mẫu đối tượng được sử dụng để đánh giá phương pháp huấn luyện mô hình 39](#_Toc168265370)

[Bảng 3.7 So sánh giữa các phương pháp dự đoán 41](#_Toc168265371)

[Bảng 3.8 Kết quả của 15 tư thế khác nhau 43](#_Toc168265372)

[Bảng 3.9 So sánh các tư thế khác nhau trong không gian 2D 45](#_Toc168265373)

[Bảng 3.10 So sánh các tư thế khác nhau trong không gian 3D 45](#_Toc168265374)

[Bảng 3.11 Kết quả của 20 mẫu thử nghiệm trong không gian 3D 47](#_Toc168265375)

LỜI CẢM ƠN

Trong suốt thời gian học tập và thực hiện đồ án tốt nghiệp em luôn nhận được sự hướng dẫn nhiệt tình từ phía nhà trường và các thầy cô giáo. Vì thế, lời đầu tiên em xin chân thành cảm ơn Trường Đại học Công nghiệp Hà Nội, thầy cô giáo khoa Công nghệ thông tin đã tạo điều kiện thuận lợi và truyền dạy kiến thức cho em trong thời gian qua để em hoàn thành đề tài một cách tốt nhất. Và hơn hết em xin bày tỏ lòng biết ơn sâu sắc nhất tới thầy giáo hướng dẫn ThS. Nguyễn Trung Phú đã tận tình giúp đỡ, định hướng em rất nhiều trong suốt quá trình tìm hiểu nghiên cứu và hoàn thành đồ án tốt nghiệp.

Sản phẩm là kết quả tốt nhất trong khả năng của quá trình nghiên cứu. Rất mong nhận được sự đóng góp của tất cả các thầy, cô giáo để sản phẩm có thể hoàn thiện hơn.

Em xin chân thành cảm ơn!

MỞ ĐẦU

1. **Tên đề tài**

Xây dựng ứng dụng đo chiều cao cơ thể người sử dụng thị giác máy tính.

1. **Lý do chọn đề tài**

Việc ước tính chiều cao là một vấn đề quan trọng trong nhiều ứng dụng thị giác máy tính, ví dụ như phán đoán tư thế con người, giám sát trực quan, phân tích trong thể thao,… Trong những năm gần đây, một số phương pháp được đề xuất đo chiều cao con người bằng các kỹ thuật dựa trên thị giác. Tuy nhiên, các phương pháp hiện tại yêu cầu số lượng lớn hình ảnh lớn được gán nhãn hoặc việc tính toán chưa chính xác đối với các tư thế khác nhau của con người. Cụ thể, các phương pháp thường tập trung vào việc tính toán chiều cao của con người tính từ điểm đầu và điểm cuối, chưa quan tâm đến việc con người có thể khụy chân hoặc cúi người. Ước tính chiều cao con người bằng thị giác máy tính từ hình ảnh, video là một nhiệm vụ đầy thách thức.

Các phương pháp đo chiều cao truyền thống để ước tính chiều cao của con người sử dụng phép đo thủ công, tốn thời gian, nhân sự, cũng như lắp đặt dụng cụ, chưa kể có thể có những sai sót. Các phương pháp tiếp cận dựa trên thị giác máy tính để đo chiều cao con người có khả năng khắc phục những hạn chế này bằng cách tự động hóa quy trình và cung cấp các phép đo chính xác.

Với những thách thức và tính ứng dụng cao của đề tài, em quyết định chọn đề tài “Xây dựng ứng dụng đo chiều cao cơ thể người sử dụng thị giác máy tính” để nghiên cứu cũng như phát triển thêm các phương pháp đo chiều cao hiệu quả hơn.

1. **Mục tiêu của đề tài**

* Phát triển một ứng dụng dựa trên công nghệ thị giác máy tính để đo chiều cao cơ thể người.
* Tích hợp các thuật toán nhận dạng và đo lường để xác định chiều cao của người dùng từ hình ảnh.
* Kiểm tra và đánh giá độ chính xác và độ tin cậy của ứng dụng trong việc đo chiều cao cơ thể so với các phương pháp đo truyền thống.

1. **Phương pháp nghiên cứu**

* Phương pháp nghiên cứu lý thuyết.
* Phương pháp thu thập số liệu.
* Phương pháp toán học.
* Phương pháp thực nghiệm

1. **Đối tượng nghiên cứu**

* Các phương pháp đo chiều cao sử dụng thị giác máy tính.
* Thuật toán nhận diện khung xương cơ thể người.
* Các công thức tính toán độ dài.
* Các cá nhân có độ tuổi từ 15 trở lên, giới tính nam và nữ, với chiều cao trong khoảng 150 cm đến 180 cm.

# CƠ SỞ LÝ THUYẾT

## Thị giác máy tính

### Định nghĩa

Thị giác máy tính (Computer Vision) là một lĩnh vực trong Trí tuệ nhân tạo (AI), bao gồm các phương pháp thu nhận, xử lý ảnh kỹ thuật số, phân tích, nhận dạng các hình ảnh, nói chung là các dữ liệu đa chiều từ thế giới thực để cho ra các thông tin số hoặc biểu tượng. Việc phát triển lĩnh vực này có bối cảnh từ việc sao chép các khả năng thị giác con người bởi sự nhận diện và hiểu biết một hình ảnh mang tính điện tử.

Thị giác máy tính là một môn học khoa học liên quan đến lý thuyết đằng sau các hệ thống nhân tạo có trích xuất các thông tin từ các hình ảnh. Dữ liệu hình ảnh có thể nhiều dạng, chẳng hạn như chuỗi video, các cảnh từ đa camera, hay dữ liệu đa chiều từ máy quét y học. Thị giác máy tính còn là một môn học kỹ thuật, trong đó tìm kiếm việc áp dụng các mô hình và các lý thuyết cho việc xây dựng các hệ thống thị giác máy tính.

Các lĩnh vực con của thị giác máy tính bao gồm tái cấu trúc cảnh, dò tìm sự kiện, theo dõi video, nhận diện bố cục đối tượng, học, chỉ mục, đánh giá chuyển động và phục hồi ảnh.

### Các lĩnh vực liên quan

Các lĩnh vực liên quan của trí tuệ nhân tạo giải quyết các vấn đề như lên kế hoạch tự động hay các suy tính cho các hệ thống robot để dò đường ở một môi trường nào đó. Sự hiểu biết chi tiết của các môi trường này được yêu cầu để dò đường thông qua chúng. Thông tin về môi trường có thể được cung cấp bởi một hệ thống thị giác máy tính, hoạt động như các cảm biến và cung cấp thông tin mức độ cao về môi trường và robot.

Trí tuệ nhân tạo và thị giác máy tính chia sẻ các chủ đề như nhận dạng mẫu và các kỹ thuật học. Kết quả là thị giác máy tính đôi khi được xem là một phần của lĩnh vực trí tuệ nhân tạo hay lĩnh vực khoa học máy tính nói chung.

Thị giác máy tính theo một cách nào đó là sự đảo ngược của đồ họa máy tính. Trong khi đồ họa máy tính sản sinh hình ảnh từ mô hình 3D, thì thị giác máy tính lại thường sản sinh ra các mô hình 3D từ dữ liệu hình ảnh. Có một khuynh hướng kết hợp 2 môn học này, ví dụ như khám phá trong tăng cường thực tế.

## Mô hình dự đoán tư thế (Pose estimate)

### Giới thiệu

Dự đoán tư thế là một nhiệm vụ trong lĩnh vực thị giác máy tính, cho phép hệ thống phát hiện và hiểu được tư thế cơ thể của con người trong các video và hình ảnh. Ví dụ, nó giúp hệ thống xác định vị trí của đầu gối trong một bức ảnh. Nhiệm vụ này tập trung vào việc xác định vị trí của các khớp chính trên cơ thể, và không liên quan đến việc nhận dạng danh tính của cá nhân trong video hay hình ảnh.

A person walking with a map of lines and numbers

Description automatically generated

Hình 1.1 Dự đoán tư thế của một người đang di chuyển và xác định các khớp trên cơ thể

Mô hình ước tính tư thế có thể giúp theo dõi một đối tượng hoặc người (bao gồm nhiều người) trong không gian thế giới thực. Trong một số trường hợp, chúng có lợi thế hơn các mô hình phát hiện đối tượng, mô hình này có thể định vị các đối tượng trong một hình ảnh nhưng chỉ cung cấp bản địa hóa chi tiết thô với hộp giới hạn tạo khung cho đối tượng. Đặt ra các mô hình ước tính, so sánh, dự đoán vị trí chính xác của các điểm chính được liên kết với một đối tượng nhất định.

Đầu vào của mô hình ước tính tư thế thường là hình ảnh camera đã xử lý và đầu ra là thông tin về các điểm chính. Các điểm chính được phát hiện được đặt tên theo ID xác định, với độ lớn đã được chuẩn hóa là 0,0 – 1,0. Nó chỉ ra xác suất mà một điểm quan trọng tồn tại ở vị trí cụ thể đó.

### Ứng dụng của mô hình dự đoán tư thế

Ước tính tư thế đã trở nên phổ biến trong nhiều lĩnh vực khác nhau, bao gồm theo dõi hoạt động và chuyển động của con người, thực tế tăng cường, hoạt hình, trò chơi và robot. Công nghệ này cho phép xây dựng các mô hình dự đoán và theo dõi chuyển động của con người, hỗ trợ nhiều ứng dụng, chẳng hạn như huấn luyện viên cá nhân dựa trên AI. Trong ví dụ này, một huấn luyện viên có thể sử dụng camera để quan sát một cá nhân đang tập luyện, và mô hình ước tính tư thế sẽ đánh giá xem liệu động tác của họ có đúng kỹ thuật hay không. Ứng dụng huấn luyện viên cá nhân với tính năng ước tính tư thế giúp tăng cường sự an toàn và hiệu quả cho các bài tập tại nhà, đồng thời có thể hoạt động trên thiết bị di động mà không cần kết nối Internet, mang lại sự tiện lợi cho việc tập luyện ở các khu vực xa xôi.

Trong lĩnh vực thực tế tăng cường (AR), ước tính tư thế tạo ra trải nghiệm AR chân thực và linh hoạt bằng cách sử dụng các điểm mốc không thay đổi để định vị và theo dõi các đối tượng như giấy tờ hay nhạc cụ. Kỹ thuật ước tính tư thế cứng nhắc cho phép hệ thống xác định và theo dõi các điểm mốc này khi chúng di chuyển trong không gian thực, từ đó phủ một đối tượng AR kỹ thuật số lên đối tượng vật lý đang được theo dõi.

Trong ngành hoạt hình, ước tính tư thế giúp tự động hóa và tối ưu hóa quy trình tạo chuyển động cho nhân vật. Công nghệ này áp dụng học sâu để dự đoán tư thế và ghi lại chuyển động trong thời gian thực, loại bỏ sự cần thiết của các điểm đánh dấu hoặc bộ quần áo chuyên dụng. Dự đoán tư thế dựa trên học sâu cũng hỗ trợ tự động hóa việc ghi lại chuyển động cho các trải nghiệm trò chơi điện tử sống động, với công nghệ như Kinect của Microsoft đã mở ra tiềm năng này.

### Các loại mô hình dự đoán tư thế

Dự đoán tư thế có thể phân ra làm các loại chính.

Đầu tiên là dự đoán tư thế người 2D, dự đoán tư thế người 2D liên quan đến việc sử dụng đầu vào trực quan, chẳng hạn như hình ảnh và video, để dự đoán vị trí không gian hoặc vị trí 2D của các điểm chính trên cơ thể người. Theo truyền thống, dự đoán tư thế người 2D sử dụng các kỹ thuật trích xuất đặc điểm thủ công cho các bộ phận cơ thể riêng lẻ. Trước đây, thị giác máy tính thu được các cấu trúc tư thế toàn cầu bằng cách mô tả cơ thể con người như một hình que. May mắn thay, các phương pháp học sâu hiện đại cải thiện đáng kể hiệu suất dự đoán tư thế con người 2D đối với dự đoán tư thế của một người cũng như nhiều người.

Loại thứ hai là dự đoán tư thế người 3D, phương pháp này dự đoán vị trí các khớp của con người trong không gian 3D. Nó hoạt động trên các hình ảnh hoặc video bằng một mắt và giúp cung cấp thông tin cấu trúc 3D trên cơ thể con người. Nó có thể hỗ trợ nhiều ứng dụng khác nhau, bao gồm hoạt hình 3D, dự đoán hành động 3D, thực tế ảo và tăng cường. Dự đoán tư thế 3D có thể sử dụng nhiều quan điểm và cảm biến bổ sung, chẳng hạn như IMU và LiDAR, đồng thời hoạt động cùng với các kỹ thuật tổng hợp thông tin. Tuy nhiên, dự đoán tư thế con người 3D phải đối mặt với một thách thức lớn. Có được chú thích hình ảnh chính xác là tốn thời gian, trong khi ghi nhãn thủ công là tốn kém và không thực tế. Hiệu quả tính toán, khái quát hóa mô hình và tính mạnh mẽ của tắc cũng đặt ra những thách thức đáng kể.

Loại dự đoán tư thế thứ ba là mô hình cơ thể người 3D, đây là loại dự đoán tư thế người sử dụng vị trí của các bộ phận cơ thể người để xây dựng hình ảnh đại diện cho cơ thể người từ dữ liệu đầu vào trực quan (chẳng hạn như nó có thể tạo tư thế bộ xương cơ thể để đại diện cho cơ thể con người). Mô hình cơ thể con người đại diện cho các điểm chính và các tính năng được trích xuất từ dữ liệu đầu vào trực quan. Nó giúp suy luận và mô tả các tư thế cơ thể người và hiển thị các tư thế 3D hoặc 2D. Nó thường liên quan đến việc sử dụng mô hình động học mô tả cơ thể con người như một thực thể có các chi và khớp, bao gồm thông tin về hình dạng cơ thể và cấu trúc động học của cơ thể.

Mở rộng hơn, đó là việc dự đoán tư thế nhiều người, nó đặt ra một thách thức đáng kể vì nó yêu cầu phân tích một môi trường đa dạng. Sự phức tạp phát sinh do số lượng và vị trí của các cá thể trong một hình ảnh là không xác định. Có hai cách tiếp cận để giúp giải quyết vấn đề này:

* Cách tiếp cận từ trên xuống: Trước tiên liên quan đến việc kết hợp máy dò người, dự đoán vị trí của các bộ phận cơ thể và cuối cùng là tính toán tư thế cho từng người.
* Cách tiếp cận từ dưới lên: Liên quan đến việc phát hiện tất cả các bộ phận của mỗi người trong một hình ảnh và sau đó liên kết hoặc nhóm các bộ phận thuộc về mỗi cá nhân…

Cách tiếp cận từ trên xuống thường dễ thực hiện hơn vì việc triển khai phát hiện con người đơn giản hơn so với các thuật toán liên kết hoặc nhóm phức tạp. Tuy nhiên, việc đánh giá phương pháp nào mang lại hiệu suất tốt hơn không phải là điều đơn giản. Hiệu suất tổng thể phụ thuộc vào thành phần nào hoạt động tốt hơn: thuật toán liên kết/nhóm hoặc phát hiện con người. Cách tiếp cận từ trên xuống bắt đầu bằng cách xác định các ứng cử viên cho con người trong ảnh (phát hiện con người). Sau đó, trong các hộp giới hạn của người được phát hiện, phân tích phân đoạn sẽ dự đoán vị trí khớp của cơ thể. Tuy nhiên, cách tiếp cận này có những nhược điểm như độ nhạy cảm với độ chính xác của việc phát hiện con người và tiêu tốn thời gian, đặc biệt khi số lượng người được phát hiện ngày càng tăng. Mặt khác, cách tiếp cận từ dưới lên trước tiên sẽ phát hiện tất cả các khớp cơ thể trong một hình ảnh, sau đó tập hợp chúng thành các tư thế cho từng cá nhân. Các nhà nghiên cứu đã đề xuất nhiều phương pháp khác nhau để đạt được điều này.

Tóm lại, dự đoán tư thế bao gồm nhiều danh mục khác nhau, mỗi danh mục được điều chỉnh cho phù hợp với các ứng dụng và thách thức cụ thể. Các kỹ thuật, thuật toán và cân nhắc liên quan đến việc dự đoán chính xác tư thế cho các cá nhân hoặc nhiều người trong các tình huống 2D và 3D khác nhau giữa các danh mục này.

## OpenCV và MediaPipe

### OpenCV

OpenCV, viết tắt của Open Source Computer Vision Library, là thư viện nguồn mở hàng đầu trong lĩnh vực Thị giác Máy tính (Computer Vision) và Học Máy (Machine Learning). Thư viện này hiện nay còn được tối ưu hóa để tăng tốc các hoạt động thời gian thực.

Dự án OpenCV được khởi xướng tại Intel vào năm 1999 bởi Gary Bradsky và được giới thiệu lần đầu vào năm 2000. Sau đó, Vadim Pisarevsky gia nhập và đảm nhận vai trò quản lý nhóm phát triển. Vào năm 2005, OpenCV được sử dụng trên xe tự lái Stanley, giúp chiếc xe này vô địch giải 2005 DARPA Grand Challenge. Dự án tiếp tục được phát triển và cải tiến với sự hỗ trợ từ Willow Garage dưới sự dẫn dắt của Gary Bradsky và Vadim Pisarevsky. Hiện nay, OpenCV cung cấp hàng loạt thuật toán cho Thị giác Máy tính và Học Máy.

OpenCV có thể được sử dụng với nhiều ngôn ngữ lập trình như C++, Python, Java, và trên các hệ điều hành khác nhau như Windows, Linux, OS X, Android và iOS. Thư viện này cũng hỗ trợ sử dụng GPU để tăng tốc độ xử lý.

OpenCV chứa hơn 2000 thuật toán được tối ưu hóa cho Thị giác Máy tính và Học Máy. Dưới đây là một số ứng dụng của OpenCV trong xử lý hình ảnh:

* Chuyển đổi hình ảnh giữa các không gian màu, ví dụ như từ BGR sang HSV, từ BGR sang màu xám, v.v.
* Thực hiện ngưỡng trên hình ảnh, chẳng hạn như ngưỡng đơn giản, ngưỡng thích ứng, v.v.
* Làm mịn hình ảnh bằng cách áp dụng các bộ lọc tùy chỉnh và làm mờ hình ảnh.
* Thực hiện các phép toán hình thái trên ảnh.
* Xây dựng hình kim tự tháp.
* Trích xuất tiền cảnh từ hình ảnh bằng thuật toán GrabCut.
* Phân đoạn ảnh sử dụng thuật toán watershed.

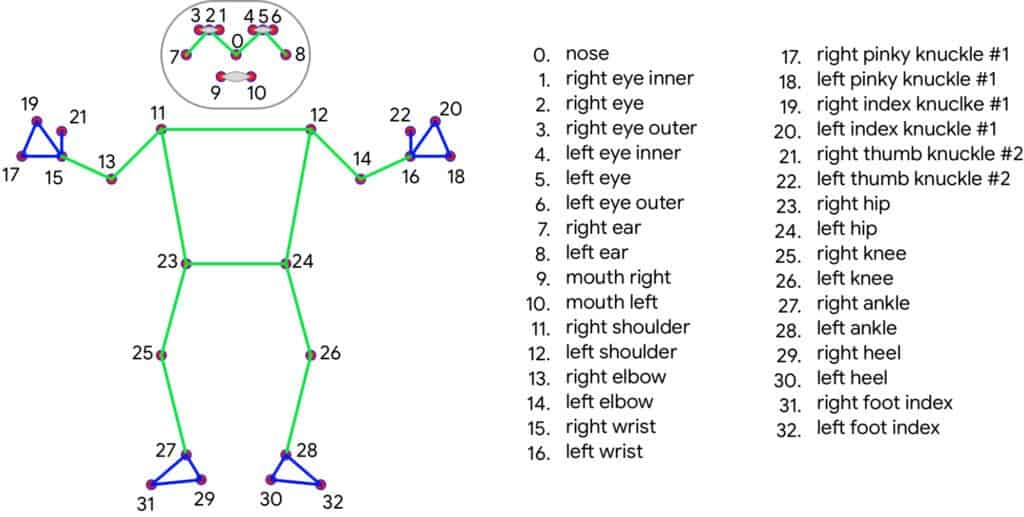
### MediaPipe

MediaPipe là một thư viện mã nguồn mở do Google AI phát triển, hỗ trợ xử lý dữ liệu trực tuyến và thời gian thực như hình ảnh, video, và âm thanh. Thư viện này cung cấp nhiều ứng dụng trong lĩnh vực nhận diện, theo dõi và phân tích hành vi của con người và các đối tượng khác trong thời gian thực.

MediaPipe cung cấp các công cụ và framework giúp phát triển dễ dàng các ứng dụng nhận diện, theo dõi chuyển động và phân tích dữ liệu hình ảnh và video. Các chức năng chính của MediaPipe bao gồm:

* Nhận diện khuôn mặt: MediaPipe có khả năng nhận diện và xác định vị trí khuôn mặt trong hình ảnh hoặc video.
* Theo dõi nhiều điểm trên cơ thể: Thư viện hỗ trợ theo dõi đồng thời nhiều điểm trên cơ thể, giúp xác định và theo dõi chuyển động của con người.
* Nhận diện cử chỉ và hành động: MediaPipe có thể nhận diện và phân tích các cử chỉ và hành động từ video, bao gồm cả cử chỉ tay và các hành động khác.
* Nhận diện đối tượng: MediaPipe có khả năng nhận diện và định vị các đối tượng trong hình ảnh và video, tuy nhiên, hỗ trợ nhận diện đối tượng của nó có hạn chế so với các thư viện chuyên biệt.
* Nhận diện và phân tích âm thanh: Thư viện hỗ trợ các công cụ nhận diện và phân tích âm thanh, cho phép xử lý tín hiệu âm thanh trong thời gian thực.

MediaPipe Pose là một giải pháp theo dõi tư thế cơ thể với độ chính xác cao, hiển thị 33 điểm mốc 3D và mặt nạ phân đoạn nền toàn cơ thể từ các khung hình RGB. Nó sử dụng cấu trúc BlazePose, một siêu tập hợp của các cấu trúc COCO, BlazeFace và BlazePalm.



Hình 1.2 Các mốc cơ thể người mà Mediapipe xác định

Mediapipe sử dụng TensorFlow Lite trong phần phụ trợ. Ban đầu, hệ thống dùng máy dò để xác định vùng quan tâm (ROI) chứa người trong khung hình. Sau đó, vùng ROI này được cắt ra và sử dụng làm đầu vào để dự đoán các điểm mốc trong ROI. Công cụ ước tính tư thế của Mediapipe phát hiện tổng cộng 33 điểm mốc. Nó xác định tọa độ x, y, và z cho mỗi điểm mốc trên cơ thể người trong hệ tọa độ 3 chiều, và những tọa độ này đã được chuẩn hóa. Trục x đại diện cho chiều ngang (chiều rộng), có chiều dương từ trái sang phải; trục y đại diện cho chiều dọc (chiều cao), có chiều dương từ trên xuống dưới; trục z cung cấp thông tin về độ sâu của điểm mốc. Để chuyển đổi giá trị các điểm mốc về kích thước thực tế, cần nhân giá trị x và y với chiều rộng và chiều cao của hình ảnh đầu vào tương ứng.

Tuy nhiên, Mediapipe cũng gặp phải một số hạn chế khi phát hiện tư thế con người trong không gian 3D. Yiqiao Lin, Xueyan Jiao và Lei Zhao đã đề xuất một giải pháp cải tiến để nâng cao khả năng phát hiện và nhận dạng tư thế con người. Để chuẩn hóa trục z, họ đã thống kê độ nghiêng của cơ thể người. Trong Mediapipe, việc dự đoán tư thế con người trong không gian 3D được ánh xạ từ hình ảnh 2D, dẫn đến việc hình ảnh 3D của người đứng thẳng bị nghiêng. Điều này ảnh hưởng lớn đến việc phát hiện độ sâu và giá trị tọa độ z mà Mediapipe trả về. Dựa trên sự thống kê góc nghiêng của cánh tay giữa Mediapipe và thực tế, Yiqiao Lin, Xueyan Jiao, và Lei Zhao đã tìm ra mối tương quan và hiệu chỉnh lại độ nghiêng sai lệch trong Mediapipe để phù hợp với thực tế.

A graph of a graph of a graph

Description automatically generated with medium confidence

Hình 1.3 Dữ liệu tư thế một người đứng thẳng giơ tay trong không gian 3D

Để tính hiệu chỉnh Z, cụ thể họ lấy dữ liệu trên một người đứng thẳng, lựa chọn trung tâm hông làm mốc và đi tìm giá trị Z của điểm trung tâm vai theo công thức sau:

Trong đó, *Zthực tế* là *Z* chính xác của điểm trung tâm vai, *ΔY* là chênh lệch giá trị *Y* giữa tâm vai và tâm hông, và *Tanthực tế* là độ nghiêng của cơ thể con người trong thực tế. Từ đó, họ xác định được hệ số tỉ lệ giữa Z thực và Z gốc do Mediapipe trả về là α và giá trị Z của các điểm khớp còn lại theo phương trình sau:

Trong nghiên cứu này, các nhà khoa học đã thực hiện các thí nghiệm so sánh để kiểm tra hiệu quả của việc cải tiến Mediapipe trong việc nhận diện tư thế 3D của con người. Các thí nghiệm được tiến hành trên nhiều đối tượng với độ tuổi từ 12 đến 50, chiều cao từ 1,5 m đến 1,95 m và cân nặng từ 40 kg đến 90 kg, bao gồm cả nam, nữ và trẻ em.

Phương pháp điều chỉnh độ nghiêng của cơ thể người đã giúp tăng độ chính xác từ 32,43% lên 81,57%, và phương pháp mô phỏng cơ thể đã cải thiện độ chính xác từ 81,57% lên 91,54%. Cuối cùng, sử dụng bộ lọc một euro và bộ lọc trung bình đã đạt kết quả tốt nhất với độ chính xác trung bình lần lượt là 95,87% và 93,916%. Độ chính xác tổng thể cũng được nâng từ 34,26% lên 96,12%. Các cải tiến này đã giúp dữ liệu 3D của cơ thể người gần hơn với tư thế thực tế và loại bỏ hiệu ứng nhiễu rõ rệt trên biểu đồ tín hiệu số.

Kết quả thí nghiệm đã chứng minh tính khả thi và hiệu quả của việc cải tiến Mediapipe trong nhận diện tư thế 3D của con người. Những cải tiến này đã nâng cao đáng kể độ chính xác và độ tin cậy trong việc nhận diện tư thế, giảm thiểu sai số và mở rộng khả năng ứng dụng của Mediapipe trong nhiều lĩnh vực như xây dựng hình ảnh 3D, điều khiển chuyển động và theo dõi sức khỏe.

## Học máy

### Định nghĩa

Ngày nay, Machine learning được coi là rất quan trọng. Nhiều công ty hàng đầu hiện nay, chẳng hạn như Facebook, Google và biến Machine learning thành một phần trung tâm trong hoạt động của họ.

Machine learning (học máy) là một lĩnh vực của trí tuệ nhân tạo (AI) tập trung vào việc phát triển các hệ thống và thuật toán có khả năng học hỏi và cải thiện hiệu suất của chúng qua kinh nghiệm từ dữ liệu. Thay vì viết cụ thể các quy tắc và lôgic, học máy cho phép máy tính tự động tìm ra các mô hình và cấu trúc từ dữ liệu đã cho, mà không cần được lập trình trực tiếp cho từng tác vụ cụ thể.

Trong quá trình học, hệ thống máy tính sẽ sử dụng dữ liệu đào tạo để hiểu các mẫu và xu hướng trong dữ liệu. Sau đó, khi gặp phải dữ liệu mới, hệ thống này có thể dự đoán, phân loại hoặc đưa ra quyết định một cách tự động dựa trên những gì đã học từ dữ liệu đã được cung cấp trước đó.

Machine learning có nhiều ứng dụng rộng rãi trong cuộc sống và công nghiệp, bao gồm nhận diện giọng nói, xử lý ngôn ngữ tự nhiên, nhận diện hình ảnh, dự đoán tín dụng, phân tích dữ liệu tài chính, tư vấn sản phẩm, và nhiều lĩnh vực khác. Nó đã đóng vai trò quan trọng trong việc tạo ra các công nghệ thông minh nhân tạo hiện đại và tiến bộ đáng kể trong các ứng dụng thực tế.

### Phân loại

Machine learning chỉ hữu ích nếu có nhiều dữ liệu và có bài toán đặt ra để hỏi để dữ liệu có thể trả lời. Các loại thuật toán Machine learning khác nhau sử dụng các cách tiếp cận khác nhau cho các tình huống khác nhau để xử lý các bài toán khác nhau như hiện nay chúng ta có các loại thuật toán Machine learning chính: học có giám sát, học không giám sát và học tăng cường.

* **Học có giám sát:**

Trong loại này, mô hình được huấn luyện trên dữ liệu đã được gán nhãn, với các đầu ra mong muốn có sẵn. Mô hình học từ mối quan hệ giữa đầu vào và đầu ra để có thể dự đoán kết quả cho dữ liệu mới. Các ví dụ bao gồm nhiệm vụ phân loại và dự đoán.

* **Học không giám sát:**

Trong loại này, mô hình không có dữ liệu đầu ra đã được gán nhãn để học từ. Thay vào đó, mô hình cố gắng hiểu cấu trúc và mối quan hệ trong dữ liệu đầu vào. Điều này thường bao gồm việc tìm ra các cụm, nhóm hoặc mẫu trong dữ liệu. Các ví dụ bao gồm việc gom cụm dữ liệu và giảm chiều dữ liệu.

* **Học tăng cường:**

Trong loại này, mô hình học thông qua tương tác với môi trường, nhận được phần thưởng hoặc phạt dựa trên các hành động mà nó thực hiện. Mục tiêu của mô hình là tối ưu hóa hành vi để đạt được các mục tiêu dài hạn. Các ví dụ bao gồm việc huấn luyện robot để thực hiện các nhiệm vụ cụ thể.

Mỗi loại thuật toán có các ứng dụng riêng biệt và phù hợp cho các vấn đề cụ thể. Sự lựa chọn của thuật toán thích hợp phụ thuộc vào loại dữ liệu bạn có và các mục tiêu bạn muốn đạt được với học máy.

### Mô hình hồi quy tuyến tính

Hồi quy tuyến tính là một trong những vấn đề quan trọng trong học máy. Trong hồi quy tuyến tính, mục tiêu là xây dựng một mô hình để dự đoán giá trị đầu ra dựa trên các biến đầu vào. Mô hình này giả định rằng mối quan hệ giữa các biến đầu vào và đầu ra là tuyến tính. Cụ thể, trong hồi quy tuyến tính, chúng ta cố gắng tạo ra đường thẳng (hoặc siêu mặt phẳng trong không gian có nhiều chiều hơn) sao cho nó trùng với dữ liệu đầu vào càng chính xác càng tốt. Đường thẳng này được gọi là "đường hồi quy" và được định nghĩa bởi các hệ số và độ dốc. Mục tiêu của hồi quy tuyến tính là tìm các giá trị của các hệ số sao cho đường hồi quy này có thể dự đoán giá trị đầu ra một cách chính xác nhất dựa trên dữ liệu đầu vào đã cho. Để đạt được điều này, chúng thường sử dụng phương pháp bình phương tối thiểu để giảm thiểu sai số giữa các dự đoán và giá trị đầu ra thực tế. Các vấn đề hồi quy tuyến tính thường được sử dụng trong các tình huống như dự đoán giá cổ phiếu, dự đoán giá nhà dựa trên các yếu tố như diện tích và vị trí, dự đoán lượng bán hàng dựa trên quảng cáo, và nhiều ứng dụng khác.

Mô hình hồi quy tuyến tính là một phương pháp trong lĩnh vực học máy được sử dụng để dự đoán một biến phụ thuộc dựa trên một hoặc nhiều biến độc lập. Ý tưởng cơ bản của hồi quy tuyến tính là xây dựng một đường thẳng (đối với hồi quy tuyến tính đơn giản) hoặc một siêu mặt phẳng (đối với hồi quy tuyến tính đa biến) để xấp xỉ mối quan hệ tuyến tính giữa các biến đầu vào và biến đầu ra. Trong mô hình hồi quy tuyến tính đa biến, giả định là mối quan hệ giữa biến phụ thuộc (được ký hiệu là Y) và biến độc lập (được ký hiệu là X1, X2, ..., Xn) được biểu diễn bằng một phương trình tuyến tính như sau.

*Y = β0 + β1 \* X1 + β2 \* X2 + ... + βn \* Xn + ε*

Trong đó:

Y là biến phụ thuộc cần được dự đoán.

X1, X2, ..., Xn là các biến độc lập.

β0, β1, β2, ..., βn là các hệ số của mô hình, cho biết mức độ ảnh hưởng của mỗi biến độc lập lên biến phụ thuộc.

ε là sai số, biểu thị sự khác biệt giữa giá trị thực tế của Y và giá trị được mô hình dự đoán.

Quá trình huấn luyện một mô hình hồi quy tuyến tính là quá trình tìm kiếm các hệ số β0, β1, β2, ..., βn sao cho sai số ε được tối thiểu hóa. Phương pháp phổ biến để đạt được điều này là sử dụng phương pháp Ordinary Least Squares (OLS), mà tối thiểu hóa tổng của bình phương các sai số.

Mô hình hồi quy tuyến tính có thể được mở rộng để xử lý nhiều biến độc lập, và nếu mối quan hệ không tuyến tính, ta có thể sử dụng biến đổi hoặc kỹ thuật kỹ thuật đặc trưng để biểu diễn mối quan hệ phức tạp hơn giữa các biến. Nó cũng có thể được sử dụng cho cả nhiệm vụ với biến phụ thuộc đơn giản và đa biến (hồi quy đa biến).

## Hình học trong không gian

### Hình học trong không gian 2D

Trong hình học không gian 2D, các đối tượng thường được biễu diễn thông qua hệ trục tọa Oxy, được áp dụng để phân tích các thuộc tính của hình dạng, tính khoảng cách, đo góc, xác định diện tích và chu vi cũng như giải các bài toán hình học khác nhau. Sau đây là một số khái niệm trong không gian 2D, yếu tố cơ bản nhất trong hình học không gian 2D là một điểm không có kích thước hoặc hình dạng và được biểu thị bằng một dấu chấm, đường thẳng là tập hợp các điểm kéo dài vô tận theo hai hướng ngược nhau, đoạn thẳng là một phần của đoạn thẳng có hai điểm cuối và không giống như đường thẳng, nó có độ dài xác định. Giả sử em có ba điểm D (, E (, F (), và một đường thẳng với E, F . Em có thể tính toán hình học bằng các công thúc dưới đây:

Tọa độ của một vector, :

Khoảng cách giữa hai điểm trong hệ tọa độ (độ dài đoạn thẳng EF):

Tọa độ trung điểm của đoạn EF gọi là điểm P có tọa độ như sau:

P

Phương trình đường thẳng với vecto pháp tuyến

|  |
| --- |
|  |
| hay |

Khoảng cách từ một điểm đến một đường thẳng, khoảng cách từ điểm D đến đường thẳng có công thức là:

### Hình học trong không gian 3D

Trong hình học không gian 3D, các đối tượng được biểu diễn bằng các điểm, đường thẳng và mặt phẳng. Hệ tọa độ Oxyz thường được sử dụng trong hình học không gian 3D, trong đó các điểm, đối tượng được xác định bởi ba tọa độ x, y và z. Những đối tượng này có thể được phân tích và thao tác bằng các nguyên tắc hình học khác nhau. Nó liên quan đến việc phân tích và đo lường khoảng cách, góc, diện tích, thể tích, các tính chất và mối quan hệ khác của các vật thể rắn nhưng ở đây chỉ nhắc đến một số trong đó. Một số khái niệm trong hình học không gian 3D như điểm, đường thẳng, đoạn thẳng,… gần giống như trong hình học không gian 2D. Trong hệ tọa độ Oxyz, giả sử cho ba điểm A (, B (, C ( và đường thẳng với A, B . Từ đó, suy ra được các công thức sau:

Tọa độ của một vector, :

|  |  |
| --- | --- |
|  |  |

Khoảng cách giữa hai điểm trong hệ tọa độ (độ dài đoạn thẳng AB):

|  |  |
| --- | --- |
|  |  |

Tọa độ trung điểm của đoạn AB gọi là điểm M được tính như sau:

Khoảng cách từ một điểm đến một đường thẳng, khoảng cách từ điểm C đến đường thẳng có công thức là, với là vecto chỉ phương của đường thẳng :

|  |  |
| --- | --- |
|  |  |

Cho hai vecto và , công thức tính tích có hướng của hai vecto như sau:

|  |  |
| --- | --- |
|  |  |

## Một số cách đo chiều cao hiện có

### Cách đo chiều cao truyền thống

Để thực hiện phương pháp đo chiều cao một cách chính xác nhất, điều đầu tiên chúng ta cần chuẩn bị là một thước đo có độ dài lớn hơn chiều cao của người đo. Có nhiều loại thước đo khác nhau như thước dây, thước đứng, thước đo chiều cao y tế,... Đối với bất kỳ loại thước đo nào, việc tìm một mặt phẳng như tường hoặc sàn để đặt thước đo và làm điểm chuẩn là rất quan trọng. Đây là bước chuẩn bị cơ bản để đảm bảo kết quả đo chính xác. Sau đó, cần điều chỉnh tư thế và vị trí đứng phù hợp với từng loại thước đo để đạt kết quả tốt nhất.

Đầu tiên là phương pháp sử dụng thước đứng. Loại thước này được đánh giá cao về độ chính xác và thường được sử dụng trong các phòng khám, trung tâm y tế, phòng tập gym,… Thước đứng có thiết kế với các vạch chia chiều cao theo phương thẳng đứng và thanh trượt ngang trên đầu; nhiều loại còn tích hợp với cân để đo trọng lượng cơ thể người. Thanh trượt ở phía trên giúp điều chỉnh dễ dàng và cải thiện độ chính xác khi đo.

Các bước đo chiều cao bằng thước đứng như sau:

* Đặt thước ở vị trí bằng phẳng và đưa thanh trượt về điểm bắt đầu.
* Đứng thẳng lưng, áp sát người vào thước, giữ cổ, lưng, và gót chân thẳng hàng, mắt hướng về phía trước, và không mang giày dép.
* Điều chỉnh thanh trượt sao cho chạm nhẹ đỉnh đầu và giữ nguyên.
* Cố định thanh trượt ở vị trí đã đo được và bước ra khỏi thước.

Khi sử dụng thước đứng thông thường, chiều cao của người đo được xác định bằng khoảng cách từ mặt đất đến vị trí của thanh trượt.

A person in a white coat standing next to a child

Description automatically generated

Hình 1.4 Phương pháp đo chiều cao với thước thẳng

Tiếp theo là phương pháp đo chiều cao bằng thước dây hoặc thước cuộn, cần chọn vị trí sàn bằng phẳng, tốt nhất là cạnh cửa hoặc trong góc phòng. Đây là phương pháp phổ biến để đo chiều cao tại nhà hoặc tự đo do tính đơn giản, thuận tiện và chi phí thấp. Tuy nhiên, khi tự đo chiều cao, kết quả có thể không chính xác bằng khi có người hỗ trợ. Các bước thực hiện như sau:

* Đứng thẳng người, áp sát vào tường, điều chỉnh tư thế sao cho mắt nhìn thẳng và cằm hướng xuống.
* Sử dụng một chiếc hộp nhỏ hoặc thước kẻ đặt nhẹ lên đỉnh đầu và vuông góc với tường.
* Dùng bút chì đánh dấu điểm tiếp xúc trên tường.
* Dùng thước dây đo khoảng cách từ sàn đến điểm đã đánh dấu, đó chính là chiều cao của bạn.

Ngoài ra, loại thước đo chiều cao y tế có thể dán tường cũng được sử dụng phổ biến. Thước này được thiết kế gọn nhẹ và thường được sử dụng ở các phòng khám sức khỏe, trung tâm y tế, bệnh viện, trường học hoặc ngay tại gia đình. Với chiều dài khoảng 1,7 – 2m, thước này phù hợp để đo chiều cao cho hầu hết mọi người. Khi cố định thước trên tường, việc đọc kết quả trở nên dễ dàng và chính xác hơn. Các bước thực hiện tương tự như khi sử dụng thước đứng và thước dây.



Hình 1.5 Thước đo chiều cao y tế dán tường

Ngoài tư thế đứng, khi đo chiều cao cho trẻ em hoặc những người không thể đứng vững, phương pháp đo nằm ngang là một lựa chọn phù hợp. Trong trường hợp này, người đo cần nằm ngửa trên một mặt phẳng ngang như sàn hoặc bàn. Một người sẽ giữ đầu người đo chạm nhẹ vào mặt tường hoặc đầu giường, trong khi người khác nắm giữ hai cổ chân và kéo thẳng. Thước đo được đặt dọc theo người đo và kết quả sẽ được đánh dấu từ đỉnh đầu đến gót chân. Khi đo chiều cao trên bề mặt nằm ngang, kết quả có thể khác so với khi đo đứng thẳng. Đặc biệt, đối với những người không thể duỗi thẳng cơ thể, kết quả đo có thể sai lệch đáng kể so với chiều cao thực tế.

Ngoài ra, với đối tượng nằm, ta có thể sử dụng phương pháp đo từng đoạn xương. Để thực hiện phương pháp này, cần chuẩn bị thước dây hoặc thước kẻ và đo lần lượt từng đoạn xương từ trên xuống dưới. Tổng độ dài của các đoạn đo sẽ cho biết chiều cao cơ thể người. Tuy nhiên, phương pháp này đòi hỏi nhiều thời gian và công sức, và sai số trong từng lần đo có thể dẫn đến sai lệch lớn trong kết quả tổng chiều cao.

### Cách đo chiều cao không tiếp xúc đã có

* **Sử dụng công nghệ học sâu (Deep Learning)**

Hiện nay, các thuật toán học sâu (Deep Learning) được nhiều nhà khoa học và doanh nghiệp sử dụng để giải quyết hàng trăm vấn đề khác nhau. Deep Learning là một lĩnh vực con của Machine Learning (học máy) trong công nghệ trí tuệ nhân tạo (AI), sử dụng mạng nơ-ron nhân tạo để mô phỏng bộ não con người. Các máy tính có thể học hỏi từ lượng dữ liệu lớn để liên tục cải tiến khả năng của chúng.

Một kỹ thuật Deep Learning mà em muốn nhắc đến là mạng nơ-ron tích chập (CNN). Đây là một kiến trúc mạng nơ-ron nhân tạo tiên tiến, được thiết kế để xử lý các bài toán liên quan đến hình ảnh phức tạp. Thuật ngữ "tích chập" xuất phát từ xử lý tín hiệu số, liên quan đến việc biến đổi thông tin đầu vào bằng cách sử dụng bộ lọc để tạo ra tín hiệu mới, chứa các đặc trưng chính và quan trọng nhất. CNN được ứng dụng rộng rãi trong các lĩnh vực như nhận diện, phân tích và phân khúc hình ảnh, cũng như phân tích video.

A diagram of body height

Description automatically generated

Hình 1.6 Biểu đồ phương pháp đo chiều cao sử dụng học sâu

Kỹ thuật này đã được Fukun Yin và Shizhe Zhou sử dụng để ước tính chính xác chiều cao của một người từ một hình ảnh độ sâu duy nhất, bất kể tư thế của người đó. Quá trình này, bắt đầu từ việc sử dụng một điểm đại diện trung gian dựa trên các bộ phận cơ thể và mạng nơ-ron sâu phức tạp qua bốn giai đoạn. Quá trình này mô tả cơ thể con người bằng cách phân đoạn thành bốn phần và dự đoán độ dài của từng phần thông qua ba CNN. Sau đó, họ xây dựng một CNN độc lập khác để kết hợp biểu diễn trung gian, độ dài các bộ phận và thông tin độ sâu, cuối cùng dự đoán chiều cao cơ thể.

Đây là phương pháp đầu tiên ước tính chiều cao chỉ từ một hình ảnh độ sâu duy nhất, và độ chính xác trung bình có thể đạt tới 99% đối với những người ở các vị trí và tư thế khác nhau.

* **Sử dụng độ dài của một bộ phận để ước tính chiều cao cơ thể**

Mặc dù việc đo chiều cao khi đứng là tiêu chuẩn vàng, nhưng phương pháp này không khả thi đối với những người bị khuyết tật về thể chất hoặc không thể đứng thẳng. Một số nghiên cứu đã chỉ ra rằng có thể ước tính chiều cao cơ thể dựa trên các bộ phận khác của cơ thể, đây là phương pháp thay thế không yêu cầu người đo phải đứng thẳng. Hiện nay, các bộ phận cơ thể thường được sử dụng trong nghiên cứu ước tính chiều cao bao gồm đầu và đầu gối.

Trong một nghiên cứu của Ye-Peng Guan, tỷ lệ khuôn mặt của một người đã được sử dụng để ước tính chiều cao cơ thể. Tỷ lệ theo chiều dọc của khuôn mặt đã được phát hiện có mối tương quan mạnh mẽ với chiều cao cơ thể và duy trì tương đối ổn định trong suốt quá trình tăng trưởng và phát triển của con người. Ye-Peng Guan đã ước tính chiều cao dựa trên tỷ lệ chiều cao theo chiều dọc của khuôn mặt theo tỷ lệ vàng. Các đặc điểm khuôn mặt được trích xuất thông qua xử lý hình ảnh tự động bằng cách sử dụng các toán tử phân tích hình ảnh. Những đặc điểm này sau đó được sử dụng để ước tính chiều cao cơ thể từ các hình ảnh chụp bằng máy ảnh tiêu chuẩn.

A comparison of a face and a face

Description automatically generated with medium confidence

Hình 1.7 Tỉ lệ vàng khuôn mặt theo chiều dọc

Để đánh giá độ chính xác, Ye-Peng Guan đã tiến hành thử nghiệm với nhiều người bằng hình ảnh khuôn mặt. Kết quả của phương pháp này cho thấy một mức độ độ chính xác tương đối cao, với sai lệch tối đa so với chiều cao thực tế là 17,80 mm, và có thể xem xét là một công cụ để ước tính chiều cao.

*A person in a green mask

Description automatically generated*

Hình 1.8 Trích xuất khuôn mặt để ước tính chiều cao

Ngoài phương pháp trên, một số nghiên cứu đã khám phá việc đo chiều cao đầu gối như một cách thay thế để ước tính chiều cao tổng thể, vì phương pháp này có thể thực hiện nhanh chóng và chỉ cần một người thực hiện. Chẳng hạn, nghiên cứu của Nopphanath Chumpathat đối với phụ nữ Thái Lan đã phát hiện rằng chiều cao đầu gối và tuổi là những yếu tố dự đoán quan trọng nhất để ước tính chiều cao tổng thể. Nghiên cứu chỉ ra rằng việc sử dụng chiều cao của bất kỳ đầu gối nào cũng có thể cung cấp và đánh giá tỷ lệ chính xác cho phương trình ước tính chiều cao. Độ chính xác của phương trình đã được cải thiện so với các nghiên cứu trước đó, tuy nhiên, vẫn còn nhiều yếu tố như độ tuổi và lối sống của con người có thể gây sai số và ảnh hưởng đến độ chính xác của phương trình.

* **Phương pháp đo chiều cao của con người bằng nhận diện đối tượng và hồi quy tuyến tính**

Đây là một phương pháp an toàn để đo chiều cao con người mà không cần tiếp xúc. Aji Bijaksana Abadi và Syahfrizal Tahcfulloh đã sử dụng các thuật toán xử lý ảnh để nhận diện đối tượng con người, cố định các thông số như khoảng cách từ người đến camera, chiều cao camera, độ sáng (dùng cho thuật toán xử lý ảnh). Từ đó lấy được chiều cao của đối tượng được tính bằng pixel trên ảnh, sau đó thu thập đủ lượng dữ liệu để thực hiện với các công thức hồi quy tuyến tính để đưa ra kết quả cuối cùng cho các bộ dữ liệu thực tế.

Các làm này đạt độ chính xác 99.41% đối với phương pháp hồi quy bình phương nhỏ nhất, nhưng có một điểm yếu đối tượng bắt buộc phải đứng thẳng để có thể đưa ra kết quả chính xác.

A collage of two people wearing a mask

Description automatically generated

Hình 1.9 Đo chiều cao bằng nhận diện đối tượng

# XÂY DỰNG ỨNG DỤNG

Dựa vào các lý thuyết, kiến thức nền tảng đã nêu ở các chương 1, chương 2 này sẽ phân tích, xây dựng và thiết kế phương pháp ước tính chiều cao. Chương này cũng sẽ đề cập chi tiết quá trình xây dựng, thiết kế, đưa ra lựa chọn phương pháp tối ưu cũng như quá trình chuẩn bị dữ liệu.

## Tổng quan quy trình

Phương pháp em thực hiện là đo chiều cao không tiếp xúc ở trong không gian Oxy, sử dụng kết hợp hai thư viện OpenCV và Mediapipe để trích xuất tọa độ x và tọa độ y của điểm landmark được vẽ trên cơ thể người. Tọa độ này được trích xuất từ ảnh được chụp lại từ camera. Với dữ liệu tọa độ landmark đã trích xuất từ ảnh, tiến hành tính toán và ứng dụng các công thức toán học để xây dựng mô hình dự đoán chiều cao. Các phương pháp dự đoán được áp dụng để đưa ra kết quả cuối cùng về chiều cao của đối tượng cần đo. Quá trình này được thực hiện với sự cân nhắc kỹ lưỡng và chính xác để đảm bảo tính chính xác của kết quả đo.

Việc trích xuất các mốc cơ thể có thể giúp không chỉ tính toán được cơ thể ở tư thế đứng thẳng mà có thể dự đoán được chiều cao ở nhiều tư thế khác nhau.



Hình 2.1 Sơ đồ khối các bước thực hiện

## Các bước thực hiện

### Thu thập và xử lý dữ liệu dữ liệu

Tiến hành thu thập trên 100 mẫu dữ liệu là hình ảnh cơ thể người được thiết lập theo đúng các thông số đề ra, đi kèm với chiều cao thực tế được đo chính xác khi đứng thẳng và không đi giày, dép của đối tượng tương ứng.

Đầu ra của bước này là các tọa độ các mốc cơ thể của các đối tượng trong các ảnh thu thập được.

#### Thu thập hình ảnh

A drawing of a person standing in a line

Description automatically generated

Hình 2.2 Mô hình các thiết lập để lấy dữ liệu và xác định chiều cao

Trong công thức trên, d là khoảng cách từ camera đến vị trí người đứng,

hc là độ cao của camera so với mặt đất, φ là độ rộng góc nhìn của camera, dh là độ rộng nhìn thấy được của camera.

Trong khuôn khổ nghiên cứu và thực hiện sản phẩm này, khi thu tập dữ liệu, em cố định chiều cao của camera và khoảng cách từ camera đến vị trí cơ thể người .

**A collage of images of a person standing in a hallway

Description automatically generated**

Hình 2.3 Một số hình ảnh trong quá trình thu thập

#### Xử lý dữ liệu

Ở bước này, em sử dụng các mẫu ảnh thu thập được trước đó, thông qua OpenCV và MediaPipe trích xuất các tọa độ x, tọa độ y của các mốc cơ thể.

Em sử dụng 9 trong 33 mốc cơ thể bao gồm: mốc 0 – mũi, mốc 11 và 12 – vai, mốc 23 và 24 – hông, mốc 26 – đầu gối phải, mốc 28 – mắt cá chân phải, mốc 30 - gót chân phải, mốc 32 – ngón chân phải.

A drawing of a person with numbers and lines

Description automatically generated

Hình 2.4 Sơ đồ các đoạn khung xương

### Tính toán độ dài các đoạn khung xương

Từ các công thức các công thức hình học trong không gian 2D ở chương 1, ta có thể tính được khoảng cách giữa các mốc và độ dài các đoạn khung xương cơ thể, cụ thể:

Để tính khoảng cách từ vai đến hông (đoạn h1), ta tính khoảng cách giữa trung điểm hai bên vai và trung điểm của hai bên hông. Tìm được tọa độ trung điểm giữa vai trái và vai phải là ; ). Tương tự, ta có tọa độ trung điểm của hai bên hông là ; ). Từ đó, áp dụng công thức ta có được khoảng cách từ vai đến hông (h1) như sau:

|  |
| --- |
|  |
| *Hay* |

Đối với đoạn khung xương h2, khoảng cách từ hông đến đầu gối. Trong phương pháp này, em lựa chọn chân bên phải để xác định chiều cao một cách chính xác. Khoảng cách từ hông đến đầu gối được xác định là khoảng cách giữa hai mốc 24 và 26. Như vậy, áp dụng công thức, ta có khoảng cách h2 như phương trình sau:

Tiếp theo, khoảng cách h3 là khoảng cách từ đầu gối tới mắt cá chân. Tương tự như cách tính khoảng cách từ hông đến đầu gối, công thức tính khoảng cách từ đầu gối (mốc 26) đến mắt cá chân (mốc 28) là h3 như phương trình sau:

Tiếp đó, khoảng cách h4 là khoảng cách từ mắt cá chân đến lòng bàn chân. Ta tìm khoảng cách h4 là khoảng cách từ mắt cá chân (mốc 28) đến lòng bàn chân. Ta thấy rằng đoạn h4 là khoảng cách từ điểm 28 đến đường thẳng đi qua hai điểm 30 (gót chân phải) và 32 (ngón chân phải). Đầu tiên, ta tìm phương trình đường thẳng đi qua hai điểm mốc 30 và 32. Kết quả tìm được như phương trình sau:

Ta có phương trình tính khoảng cách từ mắt cá chân (mốc 28) đến bàn chân phải:

Cuối cùng, khoảng cách h5 là khoảng cách từ mũi (mốc 0) tới trung điểm của hai vai có tọa độ trong không gian 2D là ; ). Em tìm ra được khoảng cách từ mũi đến trung điểm hai vai có dạng như phương trình sau:

Kết quả cuối cùng ta thu được các chiều dài h­­­1, h2, h3, h4, h5 tương ứng.

### Dự đoán chiều cao

#### Phương pháp cộng trung bình từ mũi đến đỉnh đầu

Phương pháp này, ta ước tính chiều cao còn người bằng cách cộng tất cả các đoạn khung xương đã được đề cập. Công thức có dạng là:

Trong đó, h là chiều cao con người và , với h1 là khoảng cách từ vai đến hông, h2 là khoảng cách từ hông đến đầu gối, h3 là khoảng cách từ đầu gối tới mắt cá chân, h4 là khoảng cách từ mắt cá chân đến lòng bàn chân, h5 là khoảng cách từ mũi tới trung điểm của của vai, h6 là khoảng cách từ mũi đến đỉnh đầu lấy trung bình nhiều 39 người (đơn vị - cm). Hệ số k và k0 là hệ số tỉ lệ sau khi sử dụng hàm hồi quy tuyến tính khảo sát từ 39 người có phạm vi chiều cao từ 153 cm đến 179 cm. Ở phương pháp này ta sử dụng 39 đối tượng (với tư thế đứng thẳng) để lấy khoảng cách từ mũi đến đỉnh đầu ở thực bằng hệ thống thước đo, rồi tính trung bình của khoảng cách này với 39 đối tượng, ta có được h6. Sau đó sẽ thử nghiệm với 10 người với tư thế đứng thẳng phạm vi chiều cao 153 cm đến 177.5 cm, không ở trong 39 người trên.

#### Phương pháp tỉ lệ mặt

Phương pháp này, ta ước tính chiều cao còn người bằng cách cộng tất cả các đoạn khung xương đã được đề cập ở trên. Công thức có dạng là:

Trong đó, h’ là chiều cao của con người và , với h1 là khoảng cách từ vai đến hông, h2 là khoảng cách từ hông đến đầu gối, h3 là khoảng cách từ đầu gối tới mắt cá chân, h4 là khoảng cách từ mắt cá chân đến lòng bàn chân, h5 là khoảng cách từ mũi tới trung điểm của của vai, h8 là khoảng cách từ mũi đến đỉnh đầu được tính theo công thức tỉ lệ mặt; k1 và k01 là hệ số tỉ lệ sau khi sử dụng hàm hồi quy tuyến tính với đầu vào tổng độ dài các khung xương từ h1 đến h5 và h8 của 39 mẫu đối tượng có phạm vị chiều cao từ 153 cm đến 179 cm, đầu ra là chiều cao thực của các đối tượng đó. Để thử nghiệm độ chính xác của hàm, ta sẽ kiểm tra trên 10 đối tượng khác có chiều cao từ 153 cm đến 177.5 cm, với 39 mẫu huấn luyện đề cập ở trên và đều ở tư thế đứng thẳng.

Trong phương pháp này, khoảng cách h8 từ mũi đến đỉnh đầu được tính dựa trên tỷ lệ theo chiều dọc của khuôn mặt trong nghiên cứu của Ye-Peng Guan và cộng sự của mình. Khoảng cách h8 được tính toán dựa trên một phần của khuôn mặt, và ở đây em lựa chọn khoảng cách từ đỉnh mũi đến miệng làm khoảng cách tham chiếu để tìm ra khoảng cách từ mũi đến đỉnh đầu. Tương tự như công thức tính h5, ta tìm ra được khoảng cách từ mũi (điểm 0) đến miệng (điểm 9 và 10) trong không gian 2D như phương trình:

Em tìm ra được tỉ lệ giữa khoảng cách từ mũi đến miệng và khoảng cách từ mũi đến đỉnh đầu là 1/6, từ đó ta có được công thức:

#### Phương pháp huấn luyện mô hình

Phương pháp đề suất ở đây là sử dụng độ dài các đoạn khung xương từ mũi tới chân (bao gồm từ h1 đến h5) đã được tính toán ở trên đưa vào hàm hồi quy tuyến tính bội để tìm ra phương trình tương quan giữa chiều dài các đoạn khung xương với chiều cao thực của con người. Phương trình có dạng:

Trong đó, h” là chiều cao con người, h1 là khoảng cách từ vai đến hông, h2 là khoảng cách từ hông đến đầu gối, h3 là khoảng cách từ đầu gối tới mắt cá chân, h4 là khoảng cách từ mắt cá chân đến lòng bàn chân, h5 là khoảng cách từ mũi tới trung điểm của của vai; β0, β1, β2, β3, β4, β5 là hệ số tỉ lệ sau khi sử dụng hàm hồi quy tuyến tính bội trên 71 mẫu đối tượng khác nhau với tư thế đứng thẳng.

Ta sử dụng phương trình tuyến tính bội là kết quả huấn luyện của 71 mẫu đối tượng khác nhau với chiều cao thực. Phạm vi chiều cao đo từ 152.5 cm đến 180 cm ở độ tuổi từ 18 – 23 tuổi. Sau đó, để đánh giá độ chính xác của thuật toán đối với tư thế đứng thẳng như, ta sẽ kiểm tra trên 20 đối tượng khác với 71 đối tượng trên. Phạm vị chiều cao thử nghiệm từ 153 cm đến 177 cm với đối tượng từ 18 – 23 tuổi.

# KẾT QUẢ

## Kết quả các phương pháp dự đoán chiều cao

### Kết quả phương pháp cộng trung bình từ mũi đến đỉnh đẩu

Bảng 3.1. 39 mẫu đối tượng để tính trung bình khoảng cách từ mũi đến đầu

| **STT** | **Chiều cao (cm)** | **h6 (cm)** |
| --- | --- | --- |
| 1 | 152.5 | 15 |
| 2 | 153 | 15 |
| 3 | 153 | 15 |
| 4 | 153 | 15 |
| 5 | 153 | 15 |
| 6 | 157.5 | 16.5 |
| 7 | 157.5 | 16.5 |
| 8 | 157.5 | 16.5 |
| 9 | 160 | 18 |
| 10 | 160 | 17 |
| 11 | 162 | 16 |
| 12 | 162 | 18.5 |
| 13 | 163 | 18 |
| 14 | 163.5 | 18.5 |
| 15 | 165 | 15 |
| 16 | 165.5 | 14 |
| 17 | 166 | 17 |
| 18 | 167 | 15.5 |
| 19 | 167 | 15 |
| 20 | 167 | 15.5 |
| 21 | 168 | 19 |
| 22 | 168 | 17 |
| 23 | 168.5 | 18.5 |
| 24 | 169.5 | 16.5 |
| 25 | 170 | 15 |
| 26 | 170 | 17 |
| 27 | 171 | 18 |
| 28 | 171 | 18 |
| 29 | 171.5 | 17.5 |
| 30 | 171.5 | 18 |
| 31 | 173 | 17.5 |
| 32 | 174.5 | 17.5 |
| 33 | 175 | 17 |
| 34 | 175 | 15 |
| 35 | 176 | 15 |
| 36 | 176 | 18 |
| 37 | 176 | 14.5 |
| 38 | 179 | 16 |
| 39 | 180 | 15 |
| **Trung bình (cm)** | | **16.4744** |

Bảng 3.2 10 mẫu đối tượng sử dụng để đánh giá phương pháp đề xuất

| **STT** | **Chiều cao (cm)** | **hi** | **h6 (cm)** |
| --- | --- | --- | --- |
| 1 | 153 | 0.6706 | 15 |
| 2 | 157.5 | 0.6839 | 16.5 |
| 3 | 163 | 0.7220 | 17 |
| 4 | 165.5 | 0.7705 | 16 |
| 5 | 165 | 0.7105 | 19 |
| 6 | 167 | 0.7367 | 15.5 |
| 7 | 168 | 0.7926 | 18 |
| 8 | 170 | 0.7601 | 18 |
| 9 | 174.5 | 0.7879 | 19.5 |
| 10 | 177.5 | 0.8023 | 16 |

Kết quả đo lường của phương pháp này có độ sai số dao động từ 0 cm đến 4.8 cm so với chiều cao thực của các đối tượng đo, chiều cao thực thử nghiệm trong dải từ 153 cm đến 177.5 cm. Phương pháp đề xuất có trung bình sai số ước tính chiều cao xấp xỉ 2.8 cm (trung bình sai số theo phần trăm là 1.67%).

A graph with numbers and points

Description automatically generated

Hình 3.1 Đồ thị mô tả phương trình tuyết tính giữa hi và hr

Bảng 3.3 Kết quả ước tính chiều cao của 10 đối tượng thử

| **STT** | **Chiều cao**  **(cm)** | **Chiều cao ước tính**  **(cm)** | **Độ sai số**  **(cm)** | **Độ sai số**  **(%)** |
| --- | --- | --- | --- | --- |
| 1 | 153 | 156.7758 | 3.7758 | 2.47 |
| 2 | 157.5 | 158.3817 | 0.8817 | 0.56 |
| 3 | 163 | 162.9973 | 0.0027 | 0.0017 |
| 4 | 165 | 168.8736 | 3.3940 | 2.06 |
| 5 | 165.5 | 161.6060 | 3.3736 | 2.04 |
| 6 | 167 | 164.7847 | 2.2153 | 1.33 |
| 7 | 168 | 171.5536 | 3.5536 | 2.12 |
| 8 | 170 | 167.6143 | 2.3857 | 1.40 |
| 9 | 174.5 | 170.9827 | 3.5173 | 2.02 |
| 10 | 177.5 | 172.7296 | 4.7704 | 2.69 |
| **Trung bình** | | | **2.7870** | **1.67** |

### Kết quả phương pháp tỉ lệ mặt

Bảng 3.4 10 mẫu đối tượng được sử dụng để đánh giá phương pháp đề xuất

| **STT** | **Chiều cao (cm)** |  |  |
| --- | --- | --- | --- |
| 1 | 153 | 0.6706 | 0.1250 |
| 2 | 157.5 | 0.6839 | 0.1175 |
| 3 | 163 | 0.7220 | 0.1354 |
| 4 | 165 | 0.7105 | 0.1398 |
| 5 | 165.5 | 0.7705 | 0.1542 |
| 6 | 167 | 0.7367 | 0.1297 |
| 7 | 168 | 0.7926 | 0.1470 |
| 8 | 170 | 0.7601 | 0.1408 |
| 9 | 174.5 | 0.7879 | 0.1447 |
| 10 | 177.5 | 0.8023 | 0.1518 |

Phương pháp này cho kết quả sai số trung bình là 2.8 cm (với tỷ lệ sai số trung bình là 1.67%). Sai số ước tính tối đa lên đến 5.425 cm tương đương với 3.29% ở chiều cao 165.5 cm.

A graph with numbers and dots

Description automatically generated

Hình 3.2 Đồ thị mô tả phương trình tuyến tính giữa hi và h8

Bảng 3.5 Kết quả đánh giá của 10 đối tượng thử nghiệm

| **STT** | **Chiều cao**  **(cm)** | **Dự đoán**  **(cm)** | **Độ lệch (cm)** | **Phần trăm sai số (%)** |
| --- | --- | --- | --- | --- |
| 1 | 153 | 157.5560 | 4.5560 | 2.98 |
| 2 | 157.5 | 158.1560 | 0.6560 | 0.42 |
| 3 | 163 | 163.9677 | 0.9677 | 0.59 |
| 4 | 165 | 163.2339 | 1.7661 | 1.07 |
| 5 | 165.5 | 170.9525 | 5.4525 | 3.29 |
| 6 | 167 | 164.9109 | 2.0891 | 1.25 |
| 7 | 168 | 172.5025 | 4.5025 | 2.68 |
| 8 | 170 | 168.4781 | 1.5219 | 0.90 |
| 9 | 174.5 | 171.7738 | 2.7262 | 1.56 |
| 10 | 177.5 | 174.0018 | 3.4982 | 1.97 |
| **Trung bình** | | | **2.7736** | **1.67** |

Dựa trên thuật toán của và các tính toán được thực hiện với 49 mẫu, trong đó có 39 mẫu được sử dụng làm đầu vào cho mô hình hồi quy đa thức và 10 mẫu được dự trữ để đánh giá hiệu suất của phương pháp đề xuất, ta thu được kết quả như *Bảng 3.5*. Từ kết quả, có thể thấy rằng kết quả đo của phương pháp này có sai số từ 0 cm đến 5.5 cm so với chiều cao thực tế của các đối tượng được đo. Các chiều cao được thử nghiệm nằm trong khoảng từ 153 cm đến 177.5 cm. Phương pháp đề xuất cho kết quả sai số trung bình là 2.8 cm (với tỷ lệ sai số trung bình là 1.67%). Sai số ước tính tối đa lên đến 5.425 cm tương đương với 3.29% ở chiều cao 165.5 cm.

### Kết quả phương pháp huấn luyện mô hình

Bảng 3.6 20 mẫu đối tượng được sử dụng để đánh giá phương pháp huấn luyện mô hình

| **STT** | **Chiều cao (cm)** | **h1** | **h2** | **h3** | **h4** | **h5** |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | 153 | 0.2260 | 0.1617 | 0.1571 | 0.0197 | 0.1061 |
| 2 | 153 | 0.2324 | 0.1554 | 0.1521 | 0.0217 | 0.1041 |
| 3 | 157.5 | 0.2323 | 0.1712 | 0.1540 | 0.0249 | 0.0999 |
| 4 | 157.5 | 0.2283 | 0.1718 | 0.1571 | 0.0241 | 0.1027 |
| 5 | 157.5 | 0.2289 | 0.1777 | 0.1509 | 0.0216 | 0.0992 |
| 6 | 163 | 0.2539 | 0.1788 | 0.1528 | 0.0203 | 0.1162 |
| 7 | 165 | 0.2645 | 0.1838 | 0.1654 | 0.0171 | 0.1191 |
| 8 | 165.5 | 0.2830 | 0.1794 | 0.1632 | 0.0171 | 0.1278 |
| 9 | 165.5 | 0.2470 | 0.1869 | 0.1574 | 0.0175 | 0.1155 |
| 10 | 166 | 0.2497 | 0.1800 | 0.1767 | 0.0192 | 0.1165 |
| 11 | 167 | 0.2469 | 0.1934 | 0.1633 | 0.0170 | 0.1162 |
| 12 | 167 | 0.2468 | 0.1905 | 0.1655 | 0.0201 | 0.1160 |
| 13 | 167 | 0.2490 | 0.1919 | 0.1641 | 0.0187 | 0.1216 |
| 14 | 168 | 0.2676 | 0.1958 | 0.1780 | 0.0253 | 0.1258 |
| 15 | 170 | 0.2620 | 0.1939 | 0.1773 | 0.0161 | 0.1108 |
| 16 | 170 | 0.2527 | 0.1942 | 0.1727 | 0.0143 | 0.1294 |
| 17 | 174 | 0.2785 | 0.2034 | 0.1864 | 0.0261 | 0.1438 |
| 18 | 174.5 | 0.2678 | 0.2080 | 0.1724 | 0.0075 | 0.1291 |
| 19 | 175 | 0.2732 | 0.2087 | 0.1738 | 0.0200 | 0.1301 |
| 20 | 177 | 0.2927 | 0.2086 | 0.1942 | 0.0255 | 0.1393 |

Kết quả đo của phương pháp này có sai số từ 0 cm đến 2.2 cm so với chiều cao thực sự của các đối tượng thử nghiệm, với chiều cao từ 153 cm đến 177 cm. Phương pháp đề xuất cho kết quả sai số trung bình là khoảng 1.3 cm (với tỷ lệ sai số trung bình là 0.8%).

### So sánh và đánh giá giữa các phương pháp dự đoán

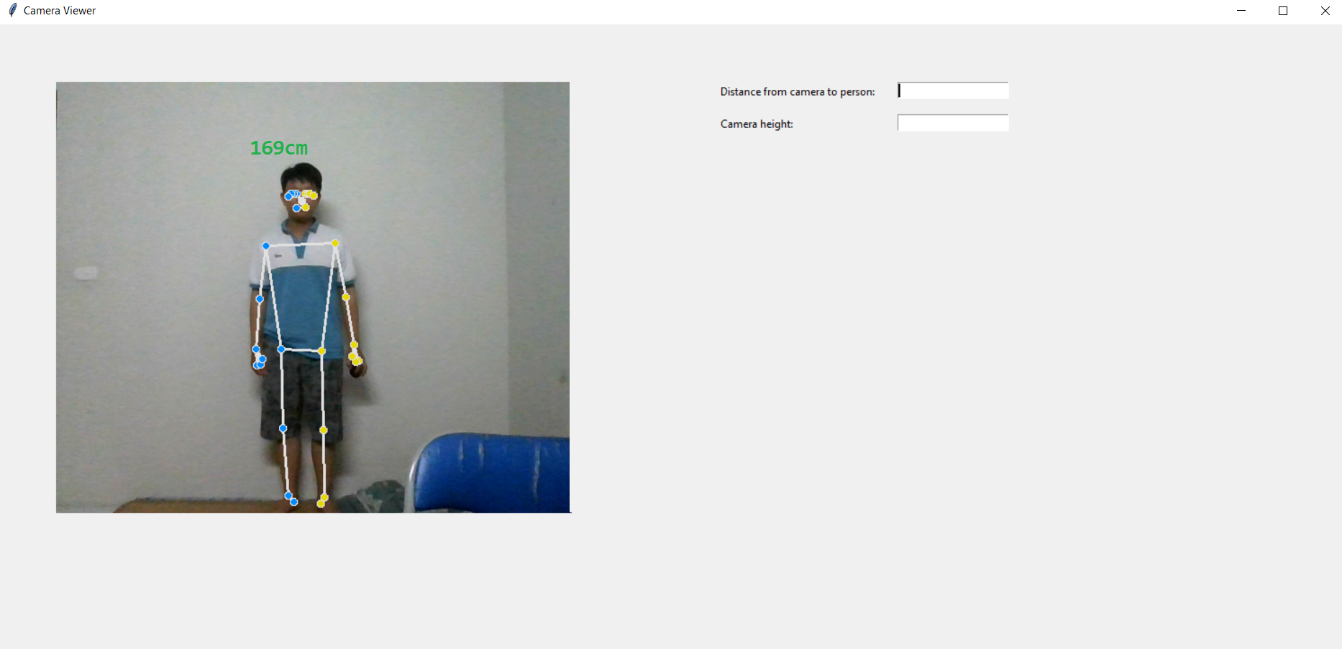
Bảng 3.7 So sánh giữa các phương pháp dự đoán

| **STT** | **Phương pháp** | **Trung bình sai số (cm)** | **Trung bình sai số (%)** | **Chú thích** |
| --- | --- | --- | --- | --- |
| 1 | Cộng trung bình khoảng cách từ mũi đến đỉnh đầu | 2.7870 | 1.67 | Sử dụng 39 mẫu huấn luyện, 10 mẫu kiểm tra; Đối tượng: nam và nữ từ 18 – 23 tuổi; Phạm vi chiều cao đo: 153 cm – 177.5 cm |
| 2 | Tỷ lệ mặt | 2.7736 | 1.67 |
| 3 | Training model | 1.3151 | 0.80 | Sử dụng 71 mẫu huấn luyện, 20 mẫu kiểm tra; Đối tượng: nam và nữ từ 18 – 23 tuổi; Phạm vi chiều cao đo: 152.5 cm – 180 cm |

Từ *Bảng 3.7*, chúng ta thấy rằng phương pháp lấy trung bình chiều dài từ mũi đến đỉnh đầu và phương pháp sử dụng tỷ lệ khuôn mặt đều cho sai số trung bình tương tự nhau là khoảng 2,8 cm (khoảng 1,7%). Tuy nhiên, phương pháp 1 có sai số cao nhất khoảng 4,8 cm, trong khi phương pháp 2 có sai số cao nhất khoảng 5,4 cm. Mặt khác, phương pháp 3 có sai số trung bình nhỏ hơn so với hai phương pháp còn lại, xấp xỉ 1,3 cm (0,80%) và sai số tối đa so với chiều cao thực tế là 2,2 cm. Vì vậy, có thể kết luận rằng phương pháp 3 tốt hơn hai phương pháp còn lại.

Dựa trên cơ sở lý luận này và kết quả thu được đối với các đối tượng đứng thẳng, em chọn phương pháp sử dụng mô hình đã được huấn luyện để sử dụng làm sản phẩm cuối cùng.

## Giao diện ứng dụng



Hình 3.3 Giao diện ứng dụng

## Các kết quả khác

### Thực hiện dự đoán chiều cao đối với các tư thế khác nhau

Sau khi thực hiện thành công các thí nghiệm với các đối tượng đứng thẳng, thu được sai số tối đa chỉ là 2.2 cm so với chiều cao thực tế, em tiến hành thử nghiệm các tư thế cơ thể khác nhau của một nữ đối tượng có chiều cao 167 cm, Kết quả của thí nghiệm như *Bảng 3.8*.

Bảng 3.8 Kết quả của 15 tư thế khác nhau

| **STT** | **Chiều cao (cm)** | **Tư thế** | **Dự đoán (cm)** | **Độ lệch** | **Phần trăm sai số (%)** |
| --- | --- | --- | --- | --- | --- |
| 1 | 167 | Tư thế 1 | 164.7150 | 2.2850 | 1.37 |
| 2 | 167 | Tư thế 2 | 173.6915 | 6.6915 | 4.01 |
| 3 | 167 | Tư thế 3 | 169.6496 | 2.6496 | 1.59 |
| 4 | 167 | Tư thế 4 | 169.4928 | 2.4928 | 1.49 |
| 5 | 167 | Tư thế 5 | 170.2463 | 3.2463 | 1.94 |
| 6 | 167 | Tư thế 6 | 165.7844 | 1.2156 | 0.73 |
| 7 | 167 | Tư thế 7 | 168.7023 | 1.7023 | 1.02 |
| 8 | 167 | Tư thế 8 | 166.8917 | 0.1083 | 0.06 |
| 9 | 167 | Tư thế 9 | 165.4133 | 1.5867 | 0.95 |
| 10 | 167 | Tư thế 10 | 175.9146 | 8.9146 | 5.34 |
| 11 | 167 | Tư thế 11 | 171.1740 | 4.1740 | 2.50 |
| 12 | 167 | Tư thế 12 | 171.0571 | 4.0571 | 2.43 |
| 13 | 167 | Tư thế 13 | 169.7709 | 2.7709 | 1.66 |
| 14 | 167 | Tư thế 14 | 172.6179 | 5.6179 | 3.36 |
| 15 | 167 | Tư thế 15 | 168.0741 | 1.0741 | 0.64 |
| **Trung bình** | | | **169.5464** | **3.2391** | **1.94** |

Có thể thấy rằng sai số đo trung bình là 3.2391 cm, tương đương với 1.94%. Tuy nhiên, sai số biến đổi đáng kể, dao động từ 0.1038 cm đến 8.9146 cm. Điều này cho thấy sự biến đổi đáng kể giữa các tư thế cơ thể khác nhau, với một số tư thế dẫn đến sai số cao hơn so với các tư thế khác. Một số tư thế như "Tư thế 8", "Tư thế 13" và "Tư thế 15" có sai số rất nhỏ, dưới 1.5 cm và sai số phần trăm dưới 1%. Ngược lại, "Tư thế 2", "Tư thế 10", "Tư thế 11", "Tư thế 12" và "Tư thế 14" có sai số cao hơn đáng kể, vượt quá 4.0 cm và sai số phần trăm trên 2.4%. Điều này cho thấy rằng các tư thế cơ thể khác nhau trong không gian 2D dẫn đến sai số lớn hơn so với tư thế thẳng đứng. Do đó, em tiếp tục thử nghiệm trong không gian 3D bằng cách sử dụng hệ thống tọa độ x, y, z để nâng cao độ chính xác của hệ thống. Đánh giá độ chính xác đối với không gian 3D.

### Thực hiện tính toán với không gian 3D

Sau khi thực hiện các thí nghiệm trong không gian 2D, em đã thấy rằng tư thế đứng thẳng đưa ra một mức độ sai số đo chấp nhận được, nhưng sai số tăng đáng kể cho các tư thế khác. Do đó, em đã cố gắng cải thiện độ chính xác bằng cách tích hợp tọa độ z bằng cách sử dụng thuật toán tương tự không gian 2D.

Để đánh giá độ chính xác của z trong không gian 3D, em thực hiện thí nghiệm với 1 camera ở các tư thế khác nhau: Tư thế đứng thẳng, tư thế đứng thẳng nhưng lệch phải, tư thế đứng thẳng dang chân, tư thế nghiêng người. Quá trình đo chiều cao bao gồm việc chụp các tư thế khác nhau cùng một đối tượng cao 167cm, không thay đổi vị trí camera trong suốt quá trình đo. Sau đó thực hiện, trích xuất tọa độ các điểm mốc để tính toán độ dài các khung xương trong từng ảnh.

Bảng 3.9 So sánh các tư thế khác nhau trong không gian 2D

| **Tư thế** |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Đứng thẳng | 0.106282 | 0.225258 | 0.207854 | 0.178026 | 0.021090 |
| Đứng thẳng lệch phải | 0.105377 | 0.226758 | 0.202578 | 0.183382 | 0.020024 |
| Đứng thẳng dang chân | 0.103929 | 0.226978 | 0.204179 | 0.179448 | 0.019959 |
| Ngiêng người | 0.110054 | 0.237992 | 0.214180 | 0.173658 | 0.020595 |
| Độ lệch | 0.006125 | 0.011233 | 0.011602 | 0.009724 | 0.001130 |
| **Phần trăm** | **5.576%** | **4.72%** | **5.42%** | **5.30%** | **5.36%** |

Bảng 3.10 So sánh các tư thế khác nhau trong không gian 3D

| Tư thế |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Đứng thẳng | 0.375538 | 0.232467 | 0.223815 | 0.307214 | 0.023854 |
| Đứng thẳng lệch phải | 0.604325 | 0.345552 | 0.217216 | 0.324277 | 0.019314 |
| Đứng thẳng dang chân | 0.502588 | 0.258887 | 0.216149 | 0.320804 | 0.017748 |
| Nghiêng người | 0.341456 | 0.252658 | 0.218266 | 0.267857 | 0.031521 |
| Độ lệch | 0.262868 | 0.113085 | 0.007667 | 0.056419 | 0.013773 |
| **Phần trăm** | **43.49%** | **32.72%** | **3.43%** | **17.40%** | **43.69%** |

Em có thể thấy độ lệch giữa đoạn xương cao nhất ở các tư thế so với đoạn xương thấp nhất trong không gian 2 chiều có độ lệch dao động từ 4,72% đến 5,56%.

Mặt khác, kết quả các tư thế khác nhau trong không gian 3 chiều, xét từng đoạn xương với nhau, cho thấy độ lệch giữa tư thế này so với tư thế kia có độ sai lệch lớn nhất giữa các từ thế lên đến 43,69%. Đặc biệt là đoạn – khoảng cách từ trung điểm vai đến mũi và – khoảng cách từ mắt cá chân đến bàn chân có độ lệch là lớn nhất. Điều này có thế kết luận rằng độ ảnh hưởng của z trong không gian 3 chiều là khá lớn.

Dựa vào nghiên cứu của em tìm hiểu và sau khi khảo sát, sử dụng Mediapipe, mô hình được đào tạo của Mediapipe có khuyết điểm trong không gian 3D. Từ tài liệu chính thức của Mediapipe cung cấp, đối với tọa độ x và y, tọa độ các mốc được chuẩn hóa trong khoảng từ 0,0 đến 1,0 lần lượt theo chiều rộng (tọa độ x) và chiều dài (tọa độ y) của ảnh. Trong khi đó, tọa độ z là độ sâu của mốc, với gốc tọa độ là tại độ sâu của trung điểm ở hông, giá trị càng nhỏ thì mốc càng gần camera. Em thấy rằng, đối với tọa độ x và y, Mediapipe trích xuất giá trị với độ sai số có thể chấp nhận được.

Mặt khác, khi sử dụng một camera, tọa độ z được Mediapipe trích xuất sẽ không được chính xác, Mediapipe phát hiện tư thế người trong không gian 3D sẽ bị nghiêng và lệch so với phương thẳng. Khi sử dụng hình ảnh đầu vào của một người ở tư thế thẳng, biểu đồ 3D được tạo từ tọa độ x, y, z do MediaPipe trích xuất cho thấy tư thế của con người bị lệch và nghiêng. Tọa độ z được trích xuất từ MediaPipe phụ thuộc vào độ sâu của từng điểm trong ảnh và độ sâu của hình ảnh phụ thuộc vào các yếu tố như điều kiện ánh sáng, tiêu cự của máy ảnh, môi trường, v.v. Điều này dẫn đến kết quả không mong muốn và không chính xác với một mức độ sai sót cao.

A graph with lines drawn on it

Description automatically generated

Hình 3.4 Hình ảnh trong không gian 3D được vẽ từ tọa độ xyz trích xuất từ MediaPipe

Kết quả cuối cùng thu được

Bảng 3.11 Kết quả của 20 mẫu thử nghiệm trong không gian 3D

| **STT** | **Chiều cao(cm)** | **Dự đoán(cm)** | **Độ lệch (cm)** | **Phần trăm sai số (%)** |
| --- | --- | --- | --- | --- |
| 1 | 153 | 157.1069 | 4.1069 | 2.68 |
| 2 | 153 | 161.7207 | 8.7207 | 5.69 |
| 3 | 157.5 | 158.7305 | 1.2305 | 0.78 |
| 4 | 157.5 | 154.817 | 2.683 | 1.70 |
| 5 | 157.5 | 155.617 | 1.883 | 1.19 |
| 6 | 163 | 159.868 | 3.132 | 1.92 |
| 7 | 165 | 159.867 | 5.133 | 3.11 |
| 8 | 165.5 | 169.391 | 3.891 | 2.35 |
| 9 | 165.5 | 162.9652 | 2.5348 | 1.53 |
| 10 | 166 | 168.3054 | 2.3054 | 1.39 |
| 11 | 167 | 168.8903 | 1.8903 | 1.13 |
| 12 | 167 | 169.518 | 2.518 | 1.51 |
| 13 | 167 | 169.0737 | 2.0737 | 1.24 |
| 14 | 168 | 171.159 | 3.159 | 1.88 |
| 15 | 170 | 173.6501 | 3.6501 | 2.15 |
| 16 | 170 | 169.1491 | 0.8509 | 0.50 |
| 17 | 174 | 170.8935 | 3.1065 | 1.79 |
| 18 | 174.5 | 167.8863 | 6.6137 | 3.79 |
| 19 | 175 | 174.0563 | 0.9437 | 0.54 |
| 20 | 177 | 173.6722 | 3.3278 | 1.88 |
| **Trung bình** | | | **3.2329** | **2.01** |

Có thể thấy sai số đo chiều cao của hệ thống so với chiều cao thực tế của đối tượng nằm trong khoảng từ 0.851 cm đến 8.721 cm (chiều cao từ 153 cm đến 177 cm), với sai số phần trăm từ 0.50% đến 5.69%. Đáng chú ý, ngay cả ở cùng một chiều cao (ví dụ, 153 cm), cũng có sự không đồng nhất trong sai số đo. Ví dụ, đối với một đối tượng cao 153 cm, sai số là 4.107 cm (2.68%), trong khi đối với một đối tượng khác có cùng chiều cao, sai số là 8.721 cm (5.69%). Sai số cao nhất là 8.721cm được quan sát cho một đối tượng có chiều cao 153 cm, trong khi sai số thấp nhất là 0.851 cm được thấy cho một đối tượng có chiều cao 170 cm. Điều này làm nổi bật tác động đáng kể của tọa độ z đối với thuật toán mà em đã phát triển.

KẾT LUẬN

Trong quá trình phát triển Ứng dụng đo chiều cao cơ thể người sử dụng thị giác máy tính, em đã sử dụng framework MediaPipe để trích xuất tọa độ khung xương trong không gian 2D và 3D từ các ảnh được ghi lại bằng một máy ảnh của các đối tượng có chiều cao đã biết. Hơn nữa, vị trí của máy ảnh trong thực nghiệm là cố định. Sau đó, em áp dụng các công thức toán học để tính toán kết quả đo chiều cao cuối cùng của các đối tượng.

Trong không gian 2D, em sử dụng ba phương pháp để xác định chiều cao của đối tượng được đo, bao gồm: phương pháp trung bình từ mũi đến đỉnh đầu, phương pháp tỷ lệ khuôn mặt và phương pháp huấn luyện mô hình để ước tính độ chính xác của hệ thống so với chiều cao thực tế. Sau khi thực hiện các thí nghiệm, sử dụng phương pháp trung bình từ mũi đến đỉnh đầu, em thu được kết quả lỗi trung bình là 2.787cm, tương đương với 1.67%, với lỗi tối đa lên đến 4.77cm (2.69%). Tương tự, với phương pháp tỷ lệ khuôn mặt, em đạt được kết quả lỗi trung bình là 2.774cm, tương đương với 1.67%, với lỗi tối đa lên đến 5.45cm (3.29%). Phương pháp huấn luyện mô hình cho kết quả lỗi trung bình là 1.315cm, tương đương với 0.80%, với lỗi tối đa lên đến 2.21cm (1.33%).

Sau khi triển khai các phương pháp được đề cập ở trên trong không gian 2D, em đã thấy rằng phương pháp huấn luyện mô hình đem lại sai số nhỏ nhất. So với nghiên cứu sử dụng độ dài đầu gối để ước tính chiều cao cơ thể con người trong bài báo của Nopphanath Chumpatha và cho thấy một sai số trung bình lên đến 2.8cm, kết quả của em cho thấy một mức độ chính xác cao hơn với 1.315cm. Tuy nhiên, độ chính xác của phương pháp của em thấp hơn khi so sánh với các kỹ thuật khác như phương pháp Sử dụng Học Sâu (Deep Learning) đạt được độ chính xác lên đến 99% trong nhiều vị trí và tư thế khác nhau. Ngược lại, phương pháp của em sử dụng công nghệ đơn giản hơn và thực hiện đơn giản hơn. Có thể khẳng định rằng em đã đạt được tiến bộ thành công ban đầu trong việc xác định chiều cao của cơ thể con người bằng cách tổng hợp các khung xương trong không gian 2D với các đối tượng đứng thẳng.

Sau đó, em đã tiến hành các thí nghiệm với 15 tư thế khác nhau của cùng một đối tượng bằng phương pháp huấn luyện mô hình trong không gian 2D. Kết quả cho thấy sai số trung bình là 3.2391 cm, tương đương với 1.94%. Tuy nhiên, một số tư thế cụ thể có sai số lên đến 6.69cm (4.01%) và 8.91cm (5.34%).

Do đó, em tiếp tục thử nghiệm hệ thống trong không gian 3D để đạt được độ chính xác cao nhất để sử dụng với các bệnh nhân cấp cứu ở nhiều vị trí khác nhau. Trong không gian 3D, ban đầu em thực hiện các thí nghiệm sử dụng tọa độ x, y, z với các mẫu đối tượng đã được nghiên cứu trước đó trong không gian 2D. Sau khi huấn luyện mô hình, em thu được kết quả sai số tối đa là 8.7207cm, tương đương với 5.69%.

Nhận thấy rằng tọa độ z ảnh hưởng đáng kể đến kết quả đo chiều cao, em đã tiến hành một nghiên cứu bằng cách tính độ dài các khung xương dọc theo các tọa độ x, y, z được trích xuất từ MediaPipe cho một đối tượng có chiều cao đã biết trong 4 tư thế khác nhau. Kết quả cho thấy sai số ở tọa độ z có thể lên đến 43.69%. Do đó, em đã tiến hành nghiên cứu về các phương pháp để chuẩn hóa tọa độ z. Tuy nhiên, do hạn chế về thời gian, sau một giai đoạn thử nghiệm ngắn, em không thể đạt được kết quả tối ưu.

Tóm lại, dựa trên kết quả nghiên cứu trong cả không gian 2D và 3D, việc xác định chiều cao không tiếp xúc thông qua phương pháp tổng hợp các khung xương hoàn toàn khả thi.

Trong tương lai, em sẽ tiếp tục nghiên cứu để tiêu chuẩn hóa tọa độ z để đạt được kết quả tốt nhất cho việc xây dựng một hệ thống ước tính chiều cao không tiếp xúc với độ chính xác cao nhất cho các tư thế khác nhau, phát triển thêm các mô hình nhận diện tư thế để dự đoán chính xác cho từng loại tư thế.

TÀI LIỆU THAM KHẢO

|  |  |
| --- | --- |
| [1] | H. O. Velesaca, S. Araujo, P. L. Suarez, A. S ´ anchez, and A. D. Sappa, “Off-the-shelf based system for urban environment video analytics,” in Int. Conf. on Systems, *Signals and Image Processing*. IEEE, 2020, pp. 459–464. |
| [2] | J. L. Charco, A. D. Sappa, B. Vintimilla, and H. O. Velesaca, “Camera pose estimation in multi-view environments: From virtual scenarios to the real world,” *Image and Vision Computing*, vol. 110, p. 104182, 2021. |
| [3] | Jaehoon Jung, Inhye Yoon, Sangkeun Lee, Joonki Paik, "Object Detection and Tracking-Based Camera Calibration for Normalized Human Height Estimation," *Journal of Sensors,* 2016. |
| [4] | Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, "You Only Look Once: Unified, Real-Time Object Detection", 2016. |
| [5] | Kaehler, Adrian; Bradski, Gary, Learning OpenCV 3: Computer Vision in C++ with the OpenCV Library, O'Reilly Media, Inc., 2016. |
| [6] | Miaopeng Li, Zimeng Zhou, Jie Li, Xinguo Liu, "Bottom-up Pose Estimation of Multiple Person with Bounding Box Constraint," 26 Jun 2018. |
| [7] | Nopphanath Chumpathat, Ram Rangsin, Supranee Changbumrung, Ngamphol Soonthornworasiri, Vanida Durongritichai, Karunee Kwanbunjan, "Use of knee height for the estimation of body height in Thai adult women," *Original Article,* 2015. |
| [8] | Yiqiao Lin, Xueyan Jiao, Lei Zhao, "Detection of 3D Human Posture Based on Improved Mediapipe," *Journal of Computer and Communications,* 2023. |
| [9] | Zhan Shi, Ziming Xu, Tongxi Wang, "A method for detecting pedestrian height and distance based on monocular vision technology," *Measurement,* 2022. |
|  |  |