TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**



**BÁO CÁO MÔN NHẬP MÔN XỬ LÝ ẢNH SỐ**

**Bài toán: Nhận dạng bảng điểm,**

**Nhận biết cử chỉ bàn tay**

*Người hướng dẫn*: **TS. PHẠM VĂN HUY**

*Người thực hiện*: **NGUYỄN TRẦN QUANG HUY – 52000668**

**LƯƠNG GIA BẢO – 52000630**

Lớp **: 20050201**

Khoá  **: 24**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2023**

TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**

A blue and red logo

Description automatically generated

**BÁO CÁO MÔN NHẬP MÔN XỬ LÝ ẢNH SỐ**

**Bài toán: Nhận dạng bảng điểm,**

**Nhận biết cử chỉ bàn tay**

*Người hướng dẫn*: **TS. PHẠM VĂN HUY**

*Người thực hiện*: **NGUYỄN TRẦN QUANG HUY – 52000668**

**LƯƠNG GIA BẢO – 52000630**

Lớp **: 20050201**

Khoá  **: 24**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2023**

LỜI CẢM ƠN

Để hoàn thành được báo cáo này, chúng em xin gửi lời cảm ơn chân thành đến thầy **Phạm Văn Huy** đã từng bước giảng dạy, hỗ trợ và bồi đắp kiến thức cho chúng em và các bạn sinh viên khác.

Đồng thời chúng em xin gửi lời cảm ơn đến khoa **Công Nghệ Thông Tin** của trường **Đại Học Tôn Đức Thắng** đã có những giáo trình chất lượng, tài liệu giảng dạy phục vụ cho sinh viên và giảng viên. Trường đã tạo những điều kiện tốt để quá trình học của sinh viên luôn được diễn ra suôn sẻ, hỗ trợ sinh viên được tiếp cận với nhiều giảng viên ở đa dạng bộ môn khác nhau và giúp cho quá trình học của sinh viên trở nên dễ dàng và thuận lợi hơn.

Trong quá trình làm và hoàn thiện báo cáo chắc chắn không thể tránh được những sai sót nên mong thầy nhận xét và đóng góp để chúng em có thể chỉnh sửa cũng như là cơ hội để chúng em có thể trau dồi lại những kiến thức bị hỏng, mất gốc. Chúng em mong nhận được sự góp ý tận tình của thầy để chúng em cải thiện tốt hơn và nâng cao kiến thức của mình hơn.

Chúng em xin chân thành cảm ơn..

TÓM TẮT

Sau một khoảng thời gian tìm hiều, học hỏi và tiếp cận với các khái niệm chuyên ngành mới mẻ cùng với đó là quá trình học tập liên tục đòi hỏi ở sự khai thác, sáng tạo của con người, đặc biệt là thế hệ trẻ trong xã hội.

Bản thân em một lần nữa xin gửi lời cảm ơn chân thành đến thầy **Phạm Văn Huy** và Thầy, Cô trong khoa đã hướng dẫn, giúp đỡ và tạo điều kiện thuận lợi để em hoàn thành tốt báo cáo này.

Các hình ảnh, tài liệu, đường link sử dụng trong bài báo cáo này được tổng hợp lại ở github của em ([@github.com:nguyenhuy158](https://github.com/nguyenhuy158/SQL-Tuning)) hoặc truy cập trực tiếp vào đường dẫn sau <https://github.com/nguyenhuy158/final_dip>. Và đường dẫn demo ở kênh youtube <https://www.youtube.com/@ntqhuy2k2>.

MỤC LỤC

[LỜI CẢM ƠN i](#_Toc142661512)

[TÓM TẮT ii](#_Toc142661513)

[MỤC LỤC iii](#_Toc142661514)

[DANH MỤC HÌNH VẼ v](#_Toc142661515)

[PHẦN 1: Lý thuyết và hiện thực phần chung: Nhận dạng bảng điểm 1](#_Toc142661516)

[1.1 Giới thiệu về bài toán 1](#_Toc142661517)

[1.2 Các thư viện được sử dụng 1](#_Toc142661518)

[1.3 Hiện thực hóa và kết quả 1](#_Toc142661519)

[PHẦN 2: Lý thuyết và hiện thực phần riêng: Nhận biết cử chỉ bàn tay 2](#_Toc142661520)

[2.1 Các ứng dụng được thực hiện 2](#_Toc142661521)

[2.1.a Trò chơi kéo - búa - bao 2](#_Toc142661522)

[2.1.b Trò chơi khủng long 2](#_Toc142661523)

[2.1.c AirGesture – nhận diện hình vẽ 3](#_Toc142661524)

[2.1.d Thoát trò chơi, quay trở lại menu 4](#_Toc142661525)

[2.2 Thư viện (framework) được sử dụng 5](#_Toc142661526)

[2.2.a MediaPipe: Google's open-source framework 5](#_Toc142661527)

[2.2.b Quickdraw-dataset 5](#_Toc142661528)

[2.2.c pynput 6](#_Toc142661529)

[2.3 Hiện thực và kết quả 6](#_Toc142661530)

[PHẦN 3: Phân công nhiệm vụ 11](#_Toc142661531)

[TÀI LIỆU THAM KHẢO 12](#_Toc142661532)

**ĐỒ ÁN ĐƯỢC HOÀN THÀNH**

**TẠI TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

Tôi xin cam đoan đây là sản phẩm đồ án của riêng tôi và được sự hướng dẫn của TS. Phạm Văn Huy. Các nội dung nghiên cứu, kết quả trong đề tài này là trung thực và chưa công bố dưới bất kỳ hình thức nào trước đây. Những số liệu trong các bảng biểu phục vụ cho việc phân tích, nhận xét, đánh giá được chính tác giả thu thập từ các nguồn khác nhau có ghi rõ trong phần tài liệu tham khảo.

Ngoài ra, trong đồ án còn sử dụng một số nhận xét, đánh giá cũng như số liệu của các tác giả khác, cơ quan tổ chức khác đều có trích dẫn và chú thích nguồn gốc.

**Nếu phát hiện có bất kỳ sự gian lận nào tôi xin hoàn toàn chịu trách nhiệm về nội dung đồ án của mình.** Trường đại học Tôn Đức Thắng không liên quan đến những vi phạm tác quyền, bản quyền do tôi gây ra trong quá trình thực hiện (nếu có).

*TP. Hồ Chí Minh, ngày 11 tháng 08 năm 2023*

*Đại diện nhóm*

*(ký tên và ghi rõ họ tên)*

*Nguyễn Trần Quang Huy*

DANH MỤC HÌNH VẼ

**No table of figures entries found.**

1. Lý thuyết và hiện thực phần chung: Nhận dạng bảng điểm
   1. Giới thiệu về bài toán
   2. Các thư viện được sử dụng
   3. Hiện thực hóa và kết quả
2. Lý thuyết và hiện thực phần riêng: Nhận biết cử chỉ bàn tay
   1. Các ứng dụng được thực hiện
      1. Trò chơi kéo - búa - bao

A computer screen shot of a person touching their hands

Description automatically generated

*Hình 2.1.1: Trò chơi kéo - búa - bao*

* Ở ứng dụng đầu tiên này, chúng em sẽ hiện thực hóa trò chơi kéo búa bao bằng cách nhận diện bàn tay của 2 người chơi.
* Với tạo hình từ bàn tay của người chơi, máy sẽ hiển thị người chơi đó ra hình gì và sẽ đưa ra kết quả người chơi nào thắng.
  + 1. Trò chơi khủng long

A screenshot of a computer

Description automatically generated

*Hình 2.1.2: Trò chơi khủng long*

* Trong trò chơi này, chúng em sẽ chơi trò khủng long (Project Bolan) của google bằng cách nắm bàn tay lại.
* Xác định nút Space sẽ là nút dùng để điều khiển khủng long nhảy lên, chúng em sẽ gán hành động nắm bàn tay như là hành động nhấn nút Space, để từ đó có thể chơi trò chơi này mà không cần phải chạm vào bàn phím.
  + 1. AirGesture – nhận diện hình vẽ
* Trò chơi này sẽ được thực hiện bằng cách dùng ngón trỏ của chúng ta để vẽ lên trước camera. Sau khi vẽ ra được hình dạng mong muốn, chúng ta sẽ thả tay ra và mô hình học máy sẽ đưa ra kết quả dự đoán. Kết quả trả về sẽ là một hình ảnh gần giống nhất với hình vẽ của mình.

A screenshot of a video game

Description automatically generated

*Hình 2.1.3: Ví dụ vẽ hình trước camera*

* Khi vẽ chúng ta sẽ nắm tay lại và chỉ đưa ra ngón trỏ dùng để vẽ.
* Ở đây em sẽ vẽ ra một hình dạng như là một cuốn sách, sổ,…
* Hình dạng được vẽ có màu xanh lục như hình ở bên trên.

A screenshot of a video game

Description automatically generated

*Hình 2.1.4: Ví dụ kết quả trả về*

* Sau khi vẽ xong, chúng ta thả tay ra để mô hình lưu hình vẽ của mình.
* Tiếp theo mô hình sẽ đưa ra kết quả dự đoán. Kết quả được hiển thị lên góc trên của camera.
* Như ví dụ trong hình thì em đã vẽ ra hình dạng một cuốn sách và kết quả trả về là hình ảnh của cuốn sách.
  + 1. Thoát trò chơi, quay trở lại menu

A screenshot of a video game

Description automatically generated

*Hình 2.1.5: Thoát và quay về menu*

* Để quay trở về menu lựa chọn trò chơi, chúng ta sẽ đưa tay nghiêng qua 90 độ và giữ yên trong khoảng thời gian ở góc dưới trái.
  1. Thư viện (framework) được sử dụng
     1. MediaPipe: Google's open-source framework
* MediaPipe là một thư viện lập trình mã nguồn mở miễn phí cho phép chúng ta nhanh chóng tích hợp các tính năng xử lý ảnh và video được tối ưu hóa cho thiết bị di động vào ứng dụng của mình. MediaPipe cung cấp một bộ tính năng rộng lớn, bao gồm:
  + Theo dõi khuôn mặt
  + Theo dõi mắt
  + Theo dõi cử chỉ ….
* MediaPipe được xây dựng trên TensorFlow Lite và được tối ưu hóa cho các thiết bị di động như điện thoại thông minh và máy tính bảng. Dễ sử dụng và được tối ưu hóa cho các thiết bị di động khiến nó trở thành một lựa chọn tuyệt vời cho các ứng dụng AR và VR.
* Trong bài này, MediaPipe được dùng để phát hiện, theo dõi và phân loại cử chỉ bàn tay như là nhận biết hình dạng kéo búa bao, nắm tay lại có nghĩa là nhấn nút,…
  + 1. Quickdraw-dataset
* QuickDraw! Dataset là một tập dữ liệu bao gồm hơn 50 triệu bản vẽ từ trò chơi Quick, Draw!. Quick, Draw! là một trò chơi trực tuyến được tạo ra bởi Google AI, cho phép người dùng thử sức mình trong việc vẽ các đối tượng trong thời gian giới hạn. Các bản vẽ sau đó được gửi đến máy chủ và được xếp loại bởi các mô hình học máy.
* Tập dữ liệu được chia thành 345 lớp đối tượng, mỗi lớp bao gồm hàng triệu bản vẽ. Tập dữ liệu được sử dụng để đào tạo các mô hình học máy có thể phân biệt các đối tượng trong các bản vẽ.
* Tập dữ liệu này được sử dụng trong phần AirGesture để thực hiện việc nhận biết hình vẽ từ camera.
  + 1. pynput
* Pynput là một thư viện Python cho phép bạn tương tác với các thiết bị đầu vào của người dùng, chẳng hạn như bàn phím, chuột và màn hình cảm ứng.
* Pynput dễ sử dụng và có thể được sử dụng để tạo một loạt các chương trình khác nhau, chẳng hạn như trò chơi, trình soạn thảo văn bản và trình mô phỏng.
* Thư viện này được sử dụng trong phần trò chơi khủng long – mỗi khi nắm bàn tay lại thì ấn nút để khủng long nhảy lên.
  1. Hiện thực và kết quả
* ***Kéo – búa – bao:***

A screenshot of a computer

Description automatically generated

*Hình 2.3.1: Kết quả hòa – Tie*

A screenshot of a computer

Description automatically generated

*Hình 2.3.2: Kết quả thắng – Player 2 win*

A screenshot of a computer

Description automatically generated

*Hình 2.3.3: Kết quả thắng – Player 1 win*

* ***Khủng long:***

A screenshot of a computer

Description automatically generated

*Hình 2.3.4: Khi để tay bình thường*

A screenshot of a computer

Description automatically generated

*Hình 2.3.5: Khi nắm tay lại để khủng long nhảy*

* ***AirGestures:*** Một số hình vẽ đơn giản

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

*Hình 2.3.6: Vẽ sách*

A screenshot of a computer

Description automatically generated

A screenshot of a computer screen

Description automatically generated

*Hình 2.3.7: Vẽ sét*

1. Phân công nhiệm vụ

Bảng Bảng phân công nhiệm vụ

|  |  |  |
| --- | --- | --- |
| Họ và tên (MSSV) | Nhiệm vụ | Tự đánh giá |
| Nguyễn Trần Quang Huy (52000668) | * Task 1 * Task 2 | Hoàn thành |
| Lương Gia Bảo (52000630) | * Task 1 * Task 2 | Hoàn thành |

TÀI LIỆU THAM KHẢO

**Tiếng Anh**

* <https://developers.google.com/mediapipe>