

**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 chữ 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**



**BÁO CÁO MÔN NHẬP MÔN XỬ LÝ ẢNH SỐ**

**Bài toán: Nhận dạng bảng điểm,  
Nhận biết chữ 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 hổng, 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 hiểu, 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)) hoặc truy cập trực tiếp vào đường dẫn sau [https://github.com/nguyenhuy158/final\\_dip](https://github.com/nguyenhuy158/final_dip). Và đường dẫn demo ở kênh youtube <https://www.youtube.com/@ntqhuy2k2>.

## MỤC LỤC

<b>LỜI CẢM ƠN .....</b>	<b>i</b>
<b>TÓM TẮT .....</b>	<b>ii</b>
<b>MỤC LỤC .....</b>	<b>iii</b>
<b>DANH MỤC HÌNH VẼ .....</b>	<b>v</b>
<b>PHẦN 1: Lý thuyết và hiện thực phần chung: Nhận dạng bảng điểm .....</b>	<b>1</b>
1.1 Giới thiệu về bài toán .....	1
1.2 Các thư viện được sử dụng.....	1
1.2.a Thư viện img2table .....	1
1.2.b Thư viện Canny .....	2
1.2.c Thư viện HoughLinesP .....	2
1.2.d Thư viện pytesseract .....	2
1.3 Hiện thực hóa và kết quả.....	3
<b>PHẦN 2: Lý thuyết và hiện thực phần riêng: Nhận biết cử chỉ bàn tay .....</b>	<b>5</b>
2.1 Các ứng dụng được thực hiện .....	5
2.1.a Trò chơi kéo - búa - bao .....	5
2.1.b Trò chơi khủng long .....	6
2.1.c AirGesture – nhận diện hình vẽ .....	7
2.1.d Thoát trò chơi, quay trở lại menu .....	8
2.2 Thư viện (framework) được sử dụng .....	9
2.2.a MediaPipe: Google's open-source framework.....	9
2.2.b Quickdraw-dataset .....	9
2.2.c pynput.....	10
2.3 Hiện thực và kết quả.....	10
<b>TÀI LIỆU THAM KHẢO .....</b>	<b>15</b>

## **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

Hình 1 Dữ liệu ảnh đầu vào.....	3
Hình 2 Dữ liệu sau khi phân tích .....	3
Hình 3 Visually dữ liệu phân tích lên ảnh gốc.....	4
Hình 4 21 điểm trên bàn tay.....	5
Hình 5 Công thức tính slope.....	5
Hình 6 Trò chơi kéo - búa - bao .....	6
Hình 7: Trò chơi khủng long.....	6
Hình 8: Ví dụ vẽ hình trước camera.....	7
Hình 9: Ví dụ kết quả trả về.....	8
Hình 10: Thoát và quay về menu.....	8
Hình 11: Kết quả hòa – Tie .....	10
Hình 12: Kết quả thắng – Player 2 win.....	11
Hình 13: Kết quả thắng – Player 1 win.....	11
Hình 14: Khi để tay bình thường.....	12
Hình 15: Khi nắm tay lại để khủng long nhảy .....	12
Hình 16: Vẽ sách.....	13
Hình 17: Vẽ sét .....	14

## **PHẦN 1: Lý thuyết và hiện thực phần chung: Nhận dạng bảng điểm**

### **1.1 Giới thiệu về bài toán**

- Bài toán nhận dạng bảng điểm này là một trong những dạng bài xử lý hình ảnh máy tính. Bài toán nhận dạng bảng điểm có thể được sử dụng trong nhiều ứng dụng khác nhau, chẳng hạn như phân tích dữ liệu thể thao, phát sóng trực tiếp và tự động hóa.
- Ở đây chúng em sẽ ứng dụng để nhận dạng và trích xuất thông tin từ bảng điểm đã được các thầy cô cung cấp trong phạm vi trường.
- Đầu tiên, chúng ta sẽ phát hiện bảng có kích thước lớn nhất trong dữ liệu là file ảnh đã cho. Sau đó chúng ta sẽ kiểm tra xem có bao nhiêu hàng, cột và bao nhiêu ô ở trong bảng đó.
- Tiếp theo chúng ta sẽ xem xét từng ô. Bằng cách cắt ô đó ra và dùng thuật toán từ thư viện pytesseract, ta sẽ lưu những ký tự đó trong 1 mảng 2 chiều và từ mảng 2 chiều ta sẽ lưu sang file excel để kiểm tra tính chính xác của kết quả.
- Cuối cùng ở phần điểm số, chúng em sẽ xem xét 2 cột điểm nguyên và điểm thập phân. Biết rằng mỗi ô tròn sẽ có một tọa độ nhất định, ta sẽ dựa vào đó để xác định điểm của từng cột, hàng rồi lưu sang file excel bên trên.

### **1.2 Các thư viện được sử dụng**

#### ***1.2.a Thư viện img2table***

- img2table là một thư viện Python được sử dụng để xác định và trích xuất bảng từ hình ảnh và tệp PDF. Nó dựa trên OpenCV, một thư viện máy tính thị giác mã nguồn mở phổ biến.
- img2table hỗ trợ hầu hết các định dạng tệp hình ảnh phổ biến, cũng như tệp PDF.



### 1.2.b Thư viện *Canny*

- Thư viện Canny trong Python là một thư viện mã nguồn mở được sử dụng để phát hiện các cạnh trong hình ảnh. Nó dựa trên thuật toán Canny Edge Detector, được phát triển vào năm 1986.
- Bạn có thể sử dụng nó để phát hiện các cạnh trong hình ảnh bằng cách sử dụng hàm `cv2.Canny()`. Hàm có ba tham số:
  - *image*: Hình ảnh đầu vào.
  - *threshold1*: Ngưỡng trên để phát hiện các cạnh.
  - *threshold2*: Ngưỡng dưới để phát hiện các cạnh.

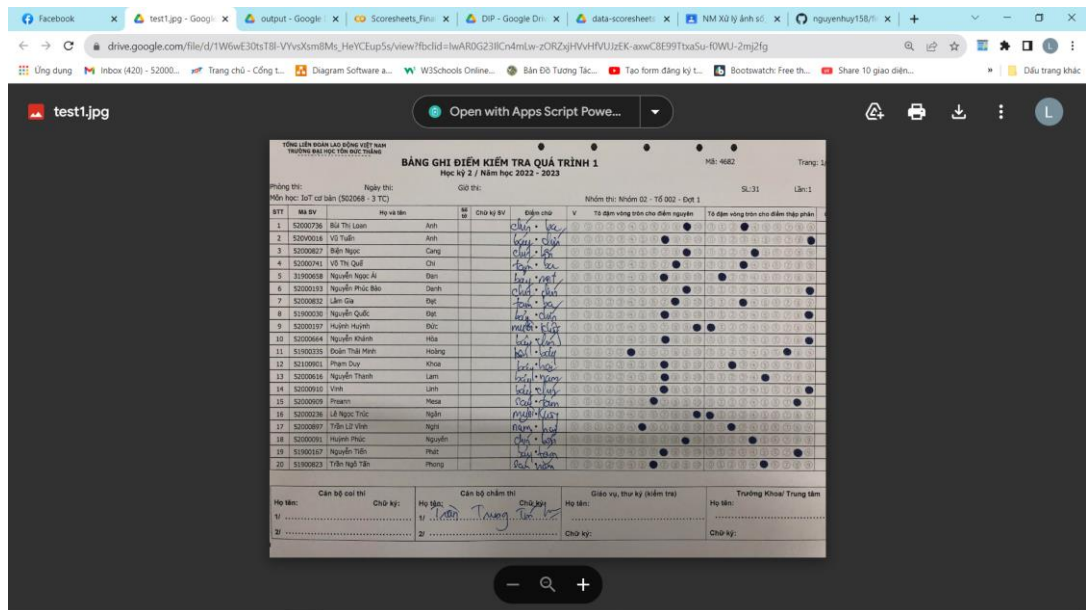
### 1.2.c Thư viện *HoughLinesP*

- Thư viện HoughLinesP trong Python là một thư viện mã nguồn mở được sử dụng để tìm đường thẳng trong hình ảnh.
- Bạn có thể sử dụng nó để tìm đường thẳng trong hình ảnh bằng cách sử dụng hàm `cv2.HoughLinesP()`. Hàm có sáu tham số:
  - *image*: Hình ảnh đầu vào.
  - *rho*: Độ dài của một bước trong không gian phẳng.
  - *theta*: Độ rộng của một bước trong không gian góc.
  - *threshold*: Ngưỡng để xác định đường thẳng.
  - *minLineLength*: Độ dài tối thiểu của một đường thẳng.
  - *maxLineGap*: Khoảng cách tối đa giữa hai điểm trên một đường thẳng.

### 1.2.d Thư viện *pytesseract*

- Pytesseract rất dễ sử dụng và có thể được sử dụng để nhận dạng ký tự trong nhiều loại hình ảnh khác nhau, bao gồm hình ảnh văn bản, hình ảnh chụp văn bản và hình ảnh được quét.
- Bạn có thể sử dụng để nhận dạng ký tự trong ảnh bằng cách sử dụng hàm `pytesseract.image_to_string()`. Hàm `pytesseract.image_to_string()` có một tham số:
  - *image*: Hình ảnh đầu vào.

### 1.3 Hiện thực hóa và kết quả



Hình 1 Dữ liệu ảnh đầu vào

STT	Mã SV	Họ và tên	Mã môn	Điểm	Tổng điểm	Tỷ lệ %
1	52000736	Bùi Thị Loan	Anh			
2	52000016	Vũ Tuấn Anh				
3	52000827	Biển Ngọc Càng				
4	52000741	Vũ Thị Quế Chi				
5	31900658	Nguyễn Ngọc Ai Đan				
6	52000193	Nguyễn Phúc Bảo Danh				
7	52000832	Lâm Gia Đạt				
8	51900030	Nguyễn Quốc Đạt				
9	52000197	Huỳnh Huỳnh Đức				
10	52000664	Nguyễn Khánh Hòa				
11	51900335	Đoàn Thái Minh Hoàng				
12	52100901	Phạm Duy Khoa				
13	52000616	Nguyễn Thanh Lâm				
14	52000910	Vinh Linh				
15	52000909	Preann Mera				
16	52000236	Lê Ngọc Trúc Ngân				
17	52000897	Trần Lê Vinh Nghi				

Hình 2 Dữ liệu sau khi phân tích

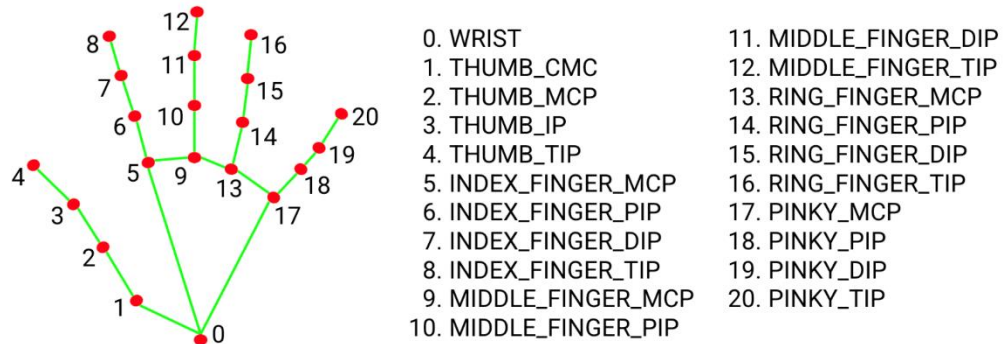
**BẢNG ĐIỂM KIỂM TRA QUÁ TRÌNH 1**  
Học kỳ 2 / Năm học 2022 - 2023

STT	Mã SV	Họ và tên	Môn	Chức vụ	Điểm số	Tổng điểm	Trung bình
1	S2000706	Phạm Thị Loan	Anh	Chức vụ	1.0	1.0	1.0
2	S2000063	Phạm Thị Loan	Anh	Chức vụ	1.0	1.0	1.0
3	S2000063	Phạm Thị Loan	Anh	Chức vụ	1.0	1.0	1.0
4	S2000063	Phạm Thị Loan	Anh	Chức vụ	1.0	1.0	1.0
5	S2000063	Phạm Thị Loan	Anh	Chức vụ	1.0	1.0	1.0
6	S2000063	Phạm Thị Loan	Anh	Chức vụ	1.0	1.0	1.0
7	S2000063	Phạm Thị Loan	Anh	Chức vụ	1.0	1.0	1.0
8	S2000063	Phạm Thị Loan	Anh	Chức vụ	1.0	1.0	1.0
9	S2000063	Phạm Thị Loan	Anh	Chức vụ	1.0	1.0	1.0
10	S2000063	Phạm Thị Loan	Anh	Chức vụ	1.0	1.0	1.0
11	S2000063	Phạm Thị Loan	Anh	Chức vụ	1.0	1.0	1.0
12	S2000063	Phạm Thị Loan	Anh	Chức vụ	1.0	1.0	1.0
13	S2000063	Phạm Thị Loan	Anh	Chức vụ	1.0	1.0	1.0
14	S2000063	Phạm Thị Loan	Anh	Chức vụ	1.0	1.0	1.0
15	S2000063	Phạm Thị Loan	Anh	Chức vụ	1.0	1.0	1.0
16	S2000063	Phạm Thị Loan	Anh	Chức vụ	1.0	1.0	1.0
17	S2000063	Phạm Thị Loan	Anh	Chức vụ	1.0	1.0	1.0
18	S2000063	Phạm Thị Loan	Anh	Chức vụ	1.0	1.0	1.0
19	S2000063	Phạm Thị Loan	Anh	Chức vụ	1.0	1.0	1.0
20	S2000063	Phạm Thị Loan	Anh	Chức vụ	1.0	1.0	1.0

Hình 3 Visually dữ liệu phân tích lên ảnh gốc

## PHẦN 2: Lý thuyết và hiện thực phần riêng: Nhận biết cử chỉ bàn tay

### 2.1 Lý thuyết về landmark trong bài toán nhận diện cử chỉ bàn tay



Hình 4 21 điểm trên bàn tay

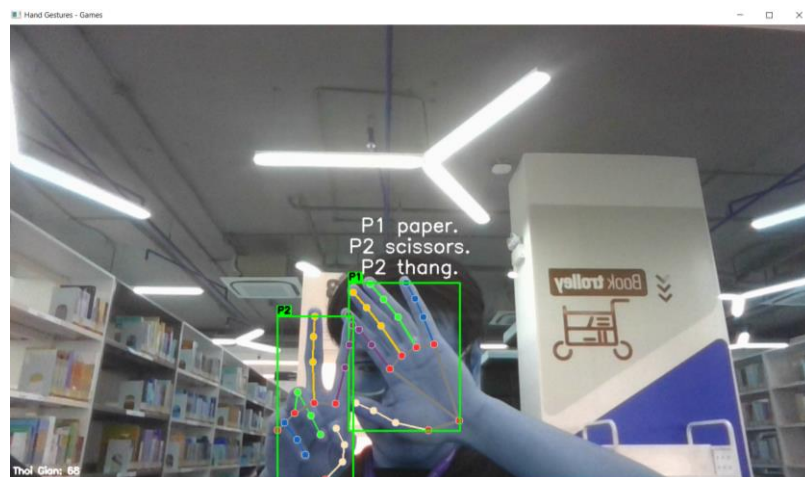
### 2.2 Lý thuyết về tính độ nghiêng của tay

$$\text{slope} = \frac{y_2 - y_1}{x_2 - x_1}$$

Hình 5 Công thức tính slope

### 2.3 Các ứng dụng được thực hiện

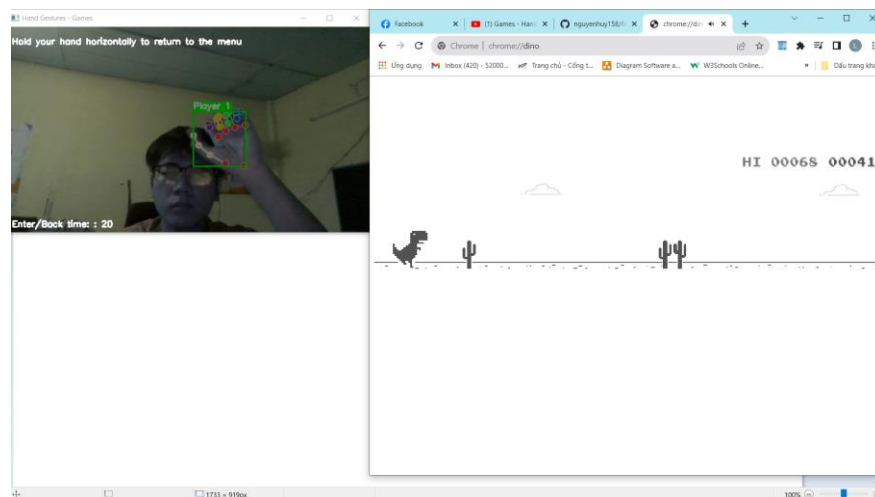
#### 2.3.a Trò chơi kéo - búa - bao



Hình 6 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.

### 2.3.b Trò chơi khủng long



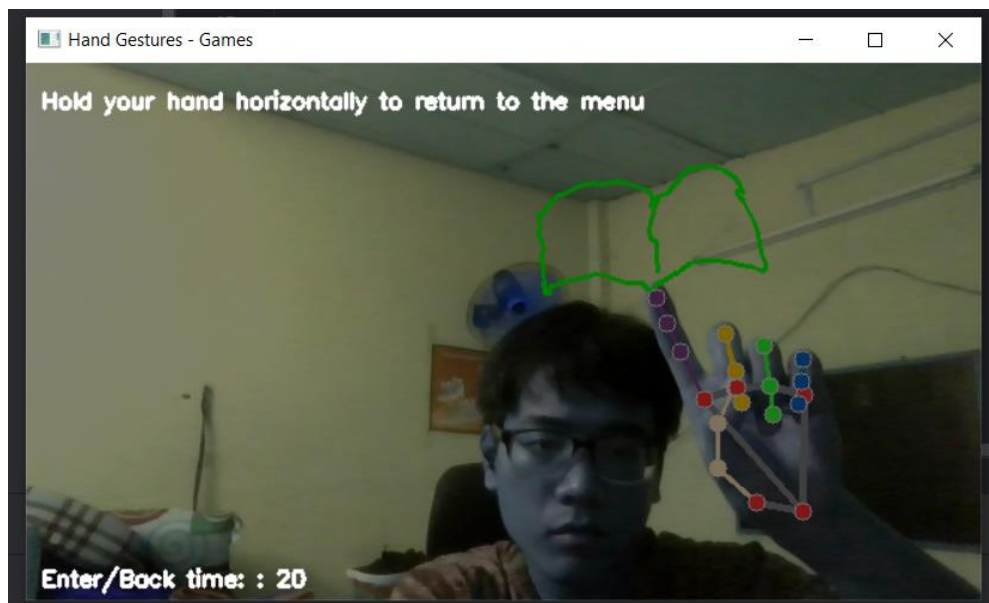
Hình 7: Trò chơi khủng long

- 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 dừng 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.

### 2.3.c 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.



Hình 8: Ví dụ vẽ hình trước camera

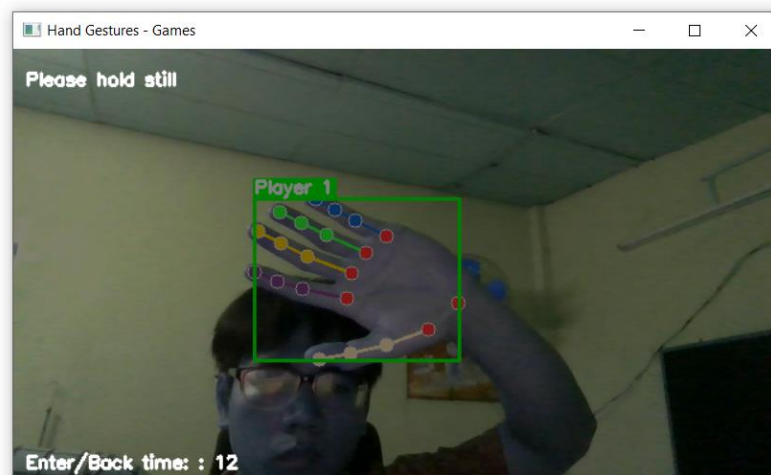
- Khi vẽ chúng ta sẽ nắm tay lại và chỉ đưa ra ngón trỏ dùng để vẽ.
- Ở đây em sẽ vẽ ra một hình dạng như là một cuốn sách, sổ,...
- Hình dạng được vẽ có màu xanh lục như hình ở bên trên.



Hình 9: 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.

### 2.3.d Thoát trò chơi, quay trở lại menu



Hình 10: 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.

## 2.4 Thư viện (framework) được sử dụng

### 2.4.a *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,...

### 2.4.b *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.

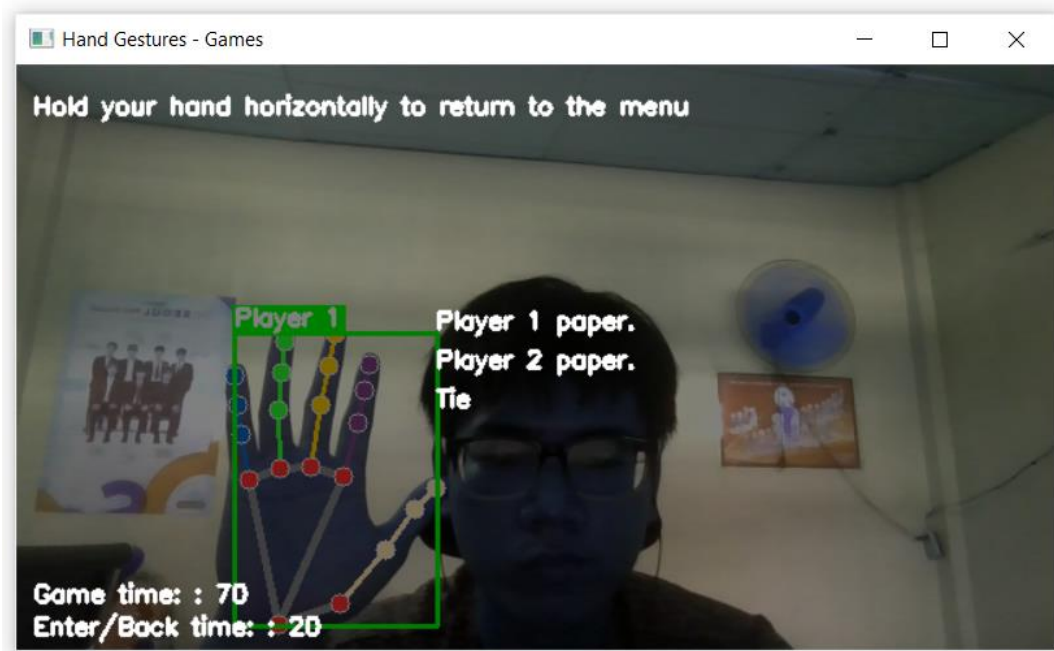


### 2.4.c pynput

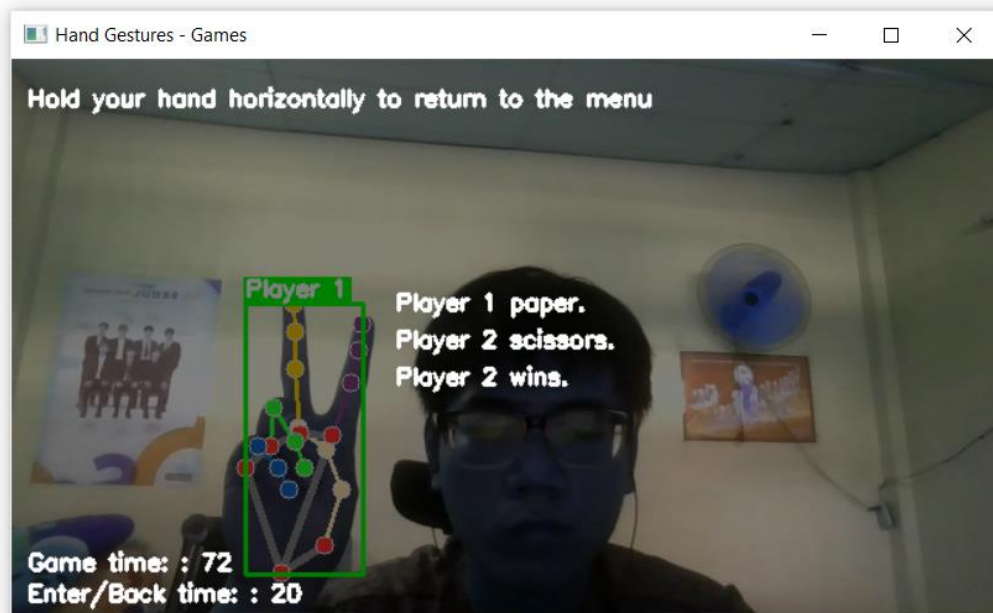
- Pynput là một thư viện Python cho phép bạn tương tác với các thiết bị đầu vào của người dùng, chẳng hạn như bàn phím, chuột và màn hình cảm ứng.
- Pynput dễ sử dụng và có thể được sử dụng để tạo một loạt các chương trình khác nhau, chẳng hạn như trò chơi, trình soạn thảo văn bản và trình mô phỏng.
- 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.

## 2.5 Hiện thực và kết quả

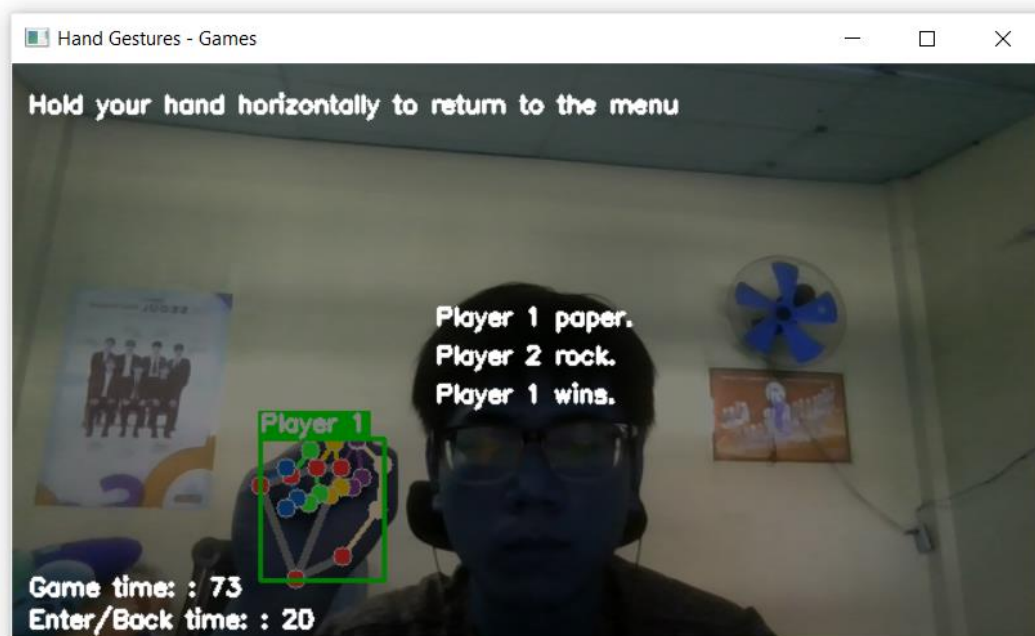
- *Kéo – búa – bao:*



Hình 11: Kết quả hòa – Tie

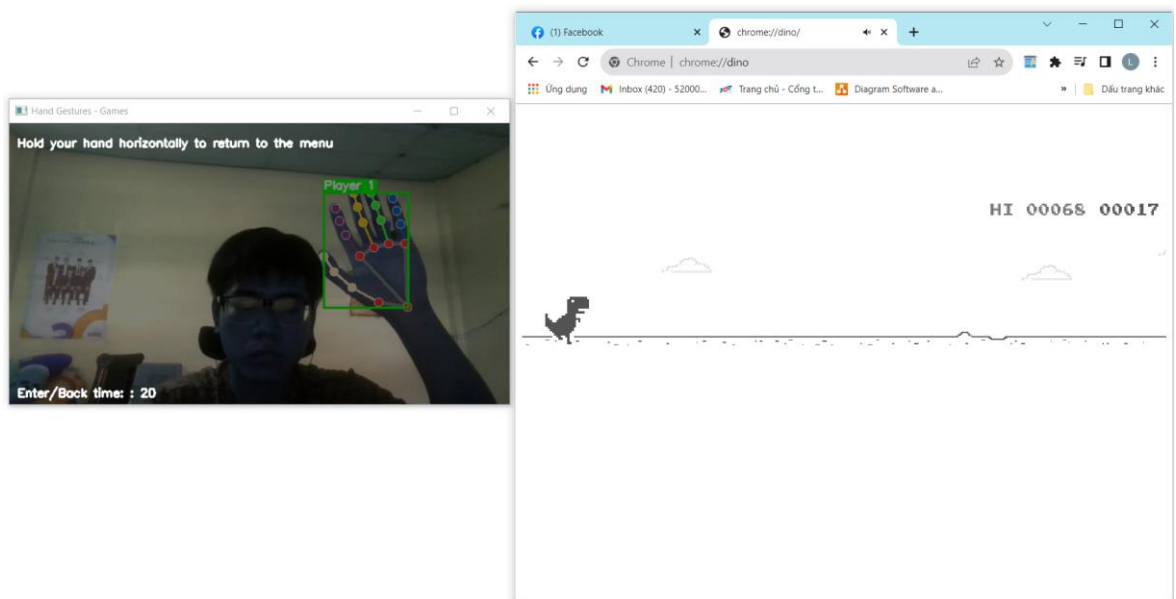


Hình 12: Kết quả thắng – Player 2 win

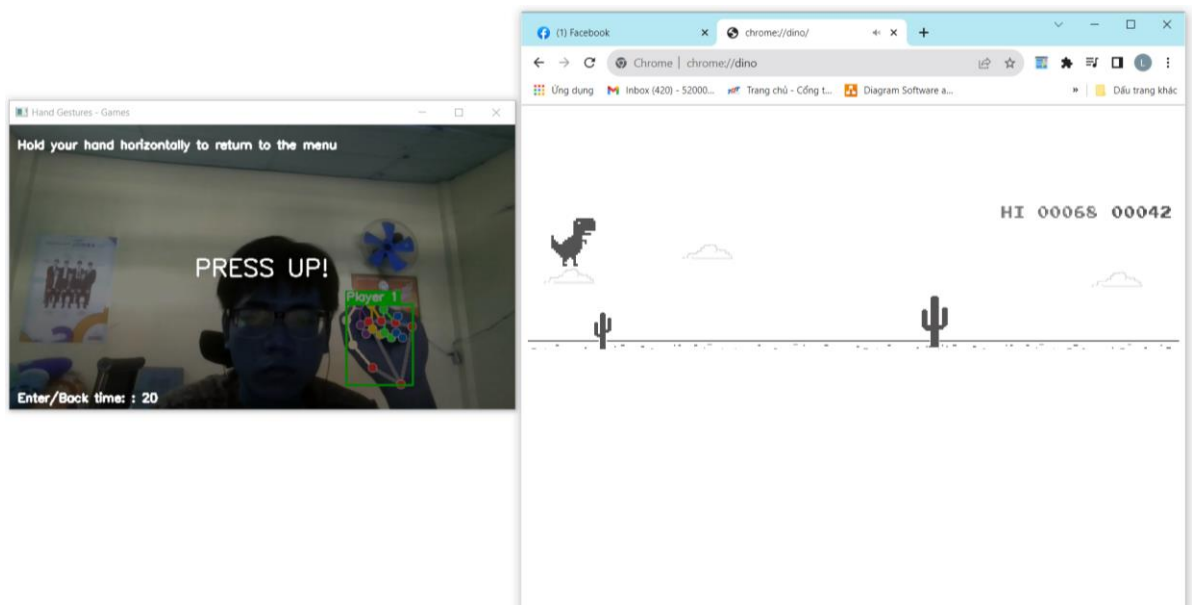


Hình 13: Kết quả thắng – Player 1 win

- **Khủng long:**

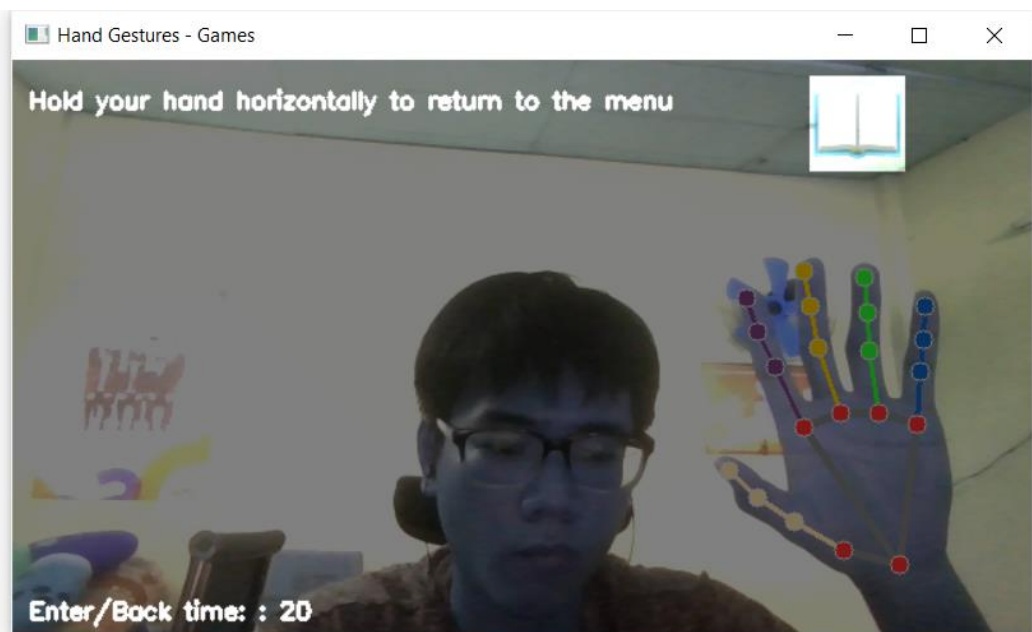
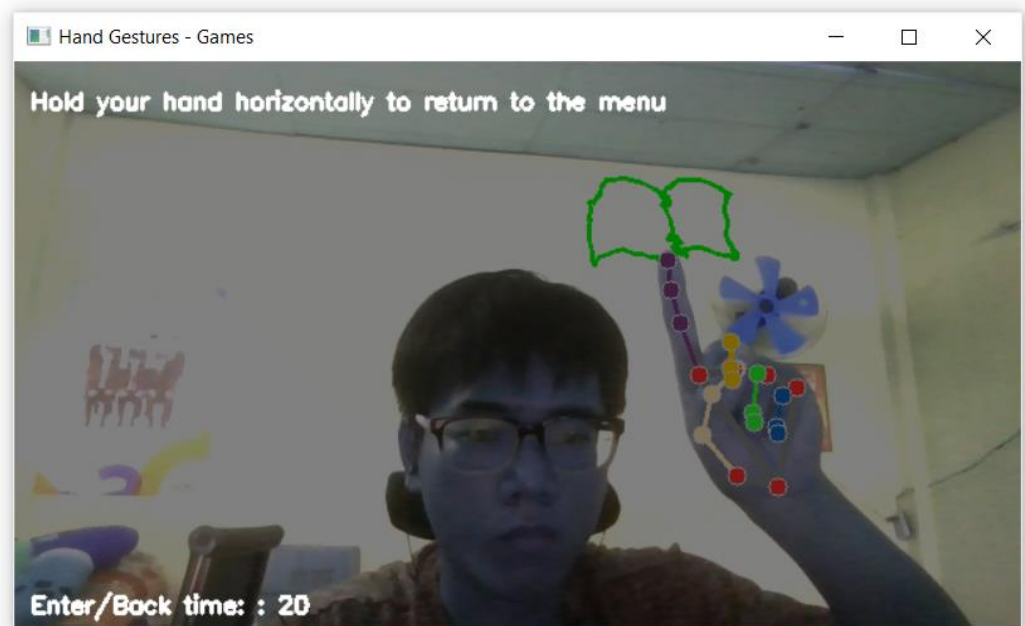


Hình 14: Khi để tay bình thường

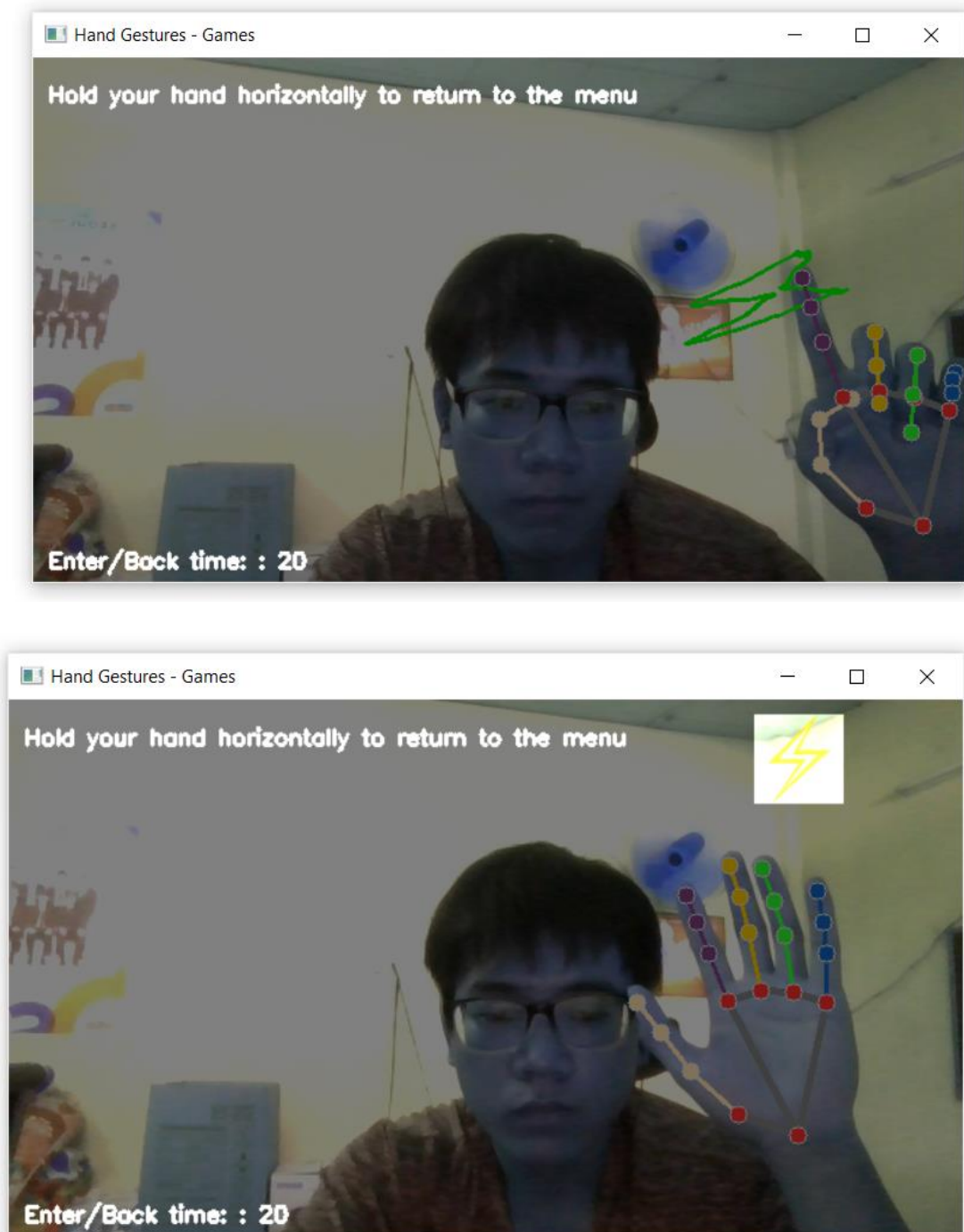


Hình 15: Khi nắm tay lại để khủng long nhảy

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



Hình 16: Vẽ sách



Hình 17: Vẽ sét

## TÀI LIỆU THAM KHẢO

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

### Link drive bài báo cáo:

- [https://drive.google.com/drive/folders/1KKDyVfomSxNiPiYdVloIgT69tkY-ULfM?fbclid=IwAR0sWAS\\_zDJUFtom2PKjAR8qt09xv1DEOW3HZnbMvNXnNJDEgd8IByc0gVk](https://drive.google.com/drive/folders/1KKDyVfomSxNiPiYdVloIgT69tkY-ULfM?fbclid=IwAR0sWAS_zDJUFtom2PKjAR8qt09xv1DEOW3HZnbMvNXnNJDEgd8IByc0gVk)

### Link video youtube:

- <https://www.youtube.com/watch?v=P-jQf5q2kJM>

### Link github:

- [https://github.com/nguyenhuy158/final\\_dip](https://github.com/nguyenhuy158/final_dip)

### Link colab:

- <https://colab.research.google.com/drive/1I0oRkMVnOprXf4iH4-mqefgZwqAQjCkU?usp=sharing&fbclid=IwAR1OWztuF8kp4En7tFbPdgsqjm5ESKtxLQyElkQqzft45TohrzHgPIbJdO0>