

*Document Report*

**ĐIỀU kHIỂN ỨNG DỤNG 3DVIEWER BẰNG CỬ CHỈ TAY CỦA NGƯỜI DÙNG.**

*Mục đích: Tài liệu phân tích và mô tả hệ thống sử dụng cử chỉ tay trong việc điều khiển ứng dụng 3DViewer.*

# Revision history

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| # | Người viết | Phiên bản | Ngày | Ghi chú |
| 1 | Tran Van Quang Huy  [tvq\_huy@brycen.com.vn](mailto:tvq_huy@brycen.com.vn) | 1.0 | 2020/08/10 | Bắt đầu viết tài liệu. |
| 2 | Tran Van Quang Huy  [tvq\_huy@brycen.com.vn](mailto:tvq_huy@brycen.com.vn) | 1.1 | 2020/08/11 | Bố trí lại cấu trúc tài liệu và thêm nội dung cơ bản. |
| 3 | Tran Van Quang Huy  [tvq\_huy@brycen.com.vn](mailto:tvq_huy@brycen.com.vn) | 1.2 | 2020/08/12 | Viết phần giới thiệu chung. |
| 4 | Tran Van Quang Huy  [tvq\_huy@brycen.com.vn](mailto:tvq_huy@brycen.com.vn) | 1.3 | 2020/08/17 | Thêm biểu đồ tuần tự, biểu đồ phân rã chức và và biểu đồ luồng. |
| 5 | Tran Van Quang Huy  [tvq\_huy@brycen.com.vn](mailto:tvq_huy@brycen.com.vn) | 1.4 | 2020/08/18 | Thêm bảng mô tả cử chỉ và mockup giao diện. |
| 6 | Tran Van Quang Huy  [tvq\_huy@brycen.com.vn](mailto:tvq_huy@brycen.com.vn) | 1.5 | 2020/09/14 | Bổ sung mục công nghệ sử dụng. |
| 7 | Tran Van Quang Huy  [tvq\_huy@brycen.com.vn](mailto:tvq_huy@brycen.com.vn) | 1.6 | 2020/09/21 | Bổ sung hướng dẫn sử dụng. |

MỤC LỤC

[MỤC LỤC 1](#_Toc52292362)

[DANH MỤC HÌNH ẢNH VÀ BẢNG 3](#_Toc52292363)

[PHẦN 1: GIỚI THIỆU CHUNG 5](#_Toc52292364)

[1.1 Đặt vấn đề 5](#_Toc52292365)

[1.1.1 Đặt vấn đề 5](#_Toc52292366)

[1.1.2 Lý do 5](#_Toc52292367)

[1.2 Mục tiêu của hệ thống 5](#_Toc52292368)

[1.2.1 Mục tiêu tổng quát 5](#_Toc52292369)

[1.2.2 Mục tiêu cụ thể 5](#_Toc52292370)

[1.3 Phạm vi thực hiện 6](#_Toc52292371)

[1.3.1 Đơn vị chủ quản 6](#_Toc52292372)

[1.3.2 Đối tượng phục vụ 6](#_Toc52292373)

[1.4 Khảo sát, phác thảo hệ thống 6](#_Toc52292374)

[1.5 Khó khăn và thuận lợi của dự án 7](#_Toc52292375)

[1.5.1 Khó khăn 7](#_Toc52292376)

[1.5.2 Thuận lợi 7](#_Toc52292377)

[PHẦN 2: PHÂN TÍCH THIẾT KẾ 9](#_Toc52292378)

[2.1 Yêu cầu chức năng 9](#_Toc52292379)

[2.1.1 Phân tích chức năng 9](#_Toc52292380)

[2.1.2 Biểu đồ phân rã chức năng 10](#_Toc52292381)

[2.1.3 Các phím tắt và hành động chuột muốn điều khiển 11](#_Toc52292382)

[2.2 Kiến trúc hệ thống 15](#_Toc52292383)

[2.3 Kiến trúc thành phần 16](#_Toc52292384)

[2.4 Biểu đồ tuần tự 18](#_Toc52292385)

[2.5 Giao diện ứng dụng 20](#_Toc52292386)

[2.5.1 Giao diện tổng quan 20](#_Toc52292387)

[2.5.2 Giao diện bảng điều khiển kết nối với server 21](#_Toc52292388)

[2.5.3 Giao diện bảng hiển thị cử chỉ trợ giúp 22](#_Toc52292389)

[2.6 Công nghệ sử dụng 23](#_Toc52292390)

[2.6.1 Môi trường và ngôn ngữ sử dụng 23](#_Toc52292391)

[2.6.2 Các thư viện chính 24](#_Toc52292392)

[2.6.3 Mô hình nhận dạng cử chỉ 24](#_Toc52292393)

[PHẦN 3: TRIỂN KHAI VÀ HƯỚNG DẪN SỬ DỤNG 26](#_Toc52292394)

[3.1 Thiết lập môi trường 26](#_Toc52292395)

[3.2 Các cách kích hoạt chức năng 27](#_Toc52292396)

[3.2.1 Kích hoạt chức năng với file 3dviewer.ini 27](#_Toc52292397)

[3.2.2 Kích hoạt chức sử dụng phím tắt “G” 28](#_Toc52292398)

[3.2.3 Kích hoạt chức năng sử dụng chuột 29](#_Toc52292399)

[3.3 Hiển thị thông báo và hình ảnh 32](#_Toc52292400)

[3.4 Bảng trợ giúp cử chỉ 33](#_Toc52292401)

[3.5 Đánh giá và lưu ý cho người dùng 35](#_Toc52292402)

[DANH MỤC TÀI LIỆU THAM KHẢO 37](#_Toc52292403)

DANH MỤC HÌNH ẢNH VÀ BẢNG

**Hình ảnh**

[Hình 1. Mô tả cách thức hoạt động mức tổng quát. 7](#_Toc52292404)

[Hình 2. Biểu đồ thể hiện các chức năng chính có trong hệ thống. 10](#_Toc52292405)

[Hình 3. Bảng mô tả chức năng phím tắt và hành động chuột của ứng dụng 3DViewer. 11](#_Toc52292406)

[Hình 4. Biểu đồ luồng thực thi của hệ thống. 15](#_Toc52292407)

[Hình 5. Module nhận dạng hình ảnh. 16](#_Toc52292408)

[Hình 6. Module điều khiển của hệ thống. 17](#_Toc52292409)

[Hình 7. Biểu đồ tuần tự người dùng giao tiếp với màn điều chỉnh hành vi và hiển thị thông tin. 18](#_Toc52292410)

[Hình 8. Biểu đồ tuần tự người dùng giao tiếp với 19](#_Toc52292411)

[Hình 9. Giao diện 3DViewer được tích hợp thêm chức năng điều khiển bằng cử chỉ tay. 20](#_Toc52292412)

[Hình 10. Bảng thiết lập kết nối với server đang ở trạng thái bật. 21](#_Toc52292413)

[Hình 11. Bảng điều khiển đang ở trạng thái tắt. 21](#_Toc52292414)

[Hình 12. Bảng Help Gesture hiển thị cử chỉ trợ giúp. 22](#_Toc52292415)

[Hình 13. Kiến trúc mạng VGG16 phân loại 1000 đối tượng. 25](#_Toc52292416)

[Hình 14. Thư mục của server Python. 26](#_Toc52292417)

[Hình 15. Cửa sổ hoàn thành sau khi chạy file setup.bat 27](#_Toc52292418)

[Hình 16. Cửa sổ thông báo server đã sẵn sàng và đang đợi kết nối đến. 27](#_Toc52292419)

[Hình 17. Các cờ được bật ở file 3dviewer.ini 28](#_Toc52292420)

[Hình 18. Phím tắt “G”. 29](#_Toc52292421)

[Hình 19. Chức năng điều khiển bằng cử chỉ trong TOOLKIT MENU. 29](#_Toc52292422)

[Hình 20. Bảng điều khiển chức năng ở trạng tái tắt. 30](#_Toc52292423)

[Hình 21. Bảng điều khiển đang ở trạng thái bật và đang kết nối server. 31](#_Toc52292424)

[Hình 22. Bảng điều khiển đang ở trạng thái bật và không có kết nối server. 32](#_Toc52292425)

[Hình 23. Thông báo ở góc phải 3DViewer khi kết nối thất bại. 32](#_Toc52292426)

[Hình 24. Thông báo ở góc phải 3DViewer khi kết nối thành công. 33](#_Toc52292427)

[Hình 25. Bảng trợ giúp cử chỉ trong TOOLKIT MENU. 34](#_Toc52292428)

[Hình 26. Hình ảnh và thông tin về chức năng của cử chỉ. 35](#_Toc52292429)

[Hình 27. Khoảng cách giữa tay và Webcam. 36](#_Toc52292430)

**Bảng**

[Bảng 1. Hình ảnh cử chỉ tương ứng với chữ cái. 12](#_Toc52292431)

[Bảng 2. Phiên bản của ngôn ngữ và công cụ sử dụng. 23](#_Toc52292432)

[Bảng 3. Phiên bản của thư viện. 24](#_Toc52292433)

[Bảng 4. Mô tả hành động bật/ tắt ứng dụng của người dùng. 30](#_Toc52292434)

[Bảng 5. Mô tả hành động đóng/mở kết nối đến server của người dùng. 31](#_Toc52292435)

1. GIỚI THIỆU CHUNG
   1. Đặt vấn đề
      1. Đặt vấn đề

Ứng dụng 3DViewer của dự án FPGA đang sử dụng phím tắt từ bàn phím và chuột cho việc điều khiển chế độ hiển thị của hình ảnh 3D, và hiện tại dự án đang muốn mở rộng thêm một tính năng trong việc điều khiển ứng dụng thông qua cử chỉ tay từ người dùng, được trích xuất thông qua Webcam trong thời gian thực.

* + 1. Lý do
* Phục vụ và hỗ trợ tốt cho công việc của dự án FPGA.
* Mở rộng thêm tính năng, tạo điều kiện thuận lợi, khả thi để triển khai dự án thành công sau này.
  1. Mục tiêu của hệ thống
     1. Mục tiêu tổng quát
* Sử dụng tính năng nhận dạng cử chỉ tay để thay đổi giao diện dữ liệu đám mây điểm – Point Cloud Data – trong 3DViewer thay cho việc sử dụng chuột và bàn phím, nhằm cung cấp thêm một cách thức điều khiển ứng dụng.
  + 1. Mục tiêu cụ thể
* Ứng dụng được mô hình nhận dạng cử chỉ vào thực tế, cụ thể là sử dụng nhận dạng cử chỉ phục vụ cho mục đích điều khiển ứng dụng 3DViewer.
* Phối hợp và kết nối giữa những module thực hiện các nhiệm vụ khác nhau.
* Mở rộng khả năng tương tác giữa người và ứng dụng 3DViewer.
  1. Phạm vi thực hiện
     1. Đơn vị chủ quản

Công ty TNHH MTV Brycen Việt Nam.

* + 1. Đối tượng phục vụ

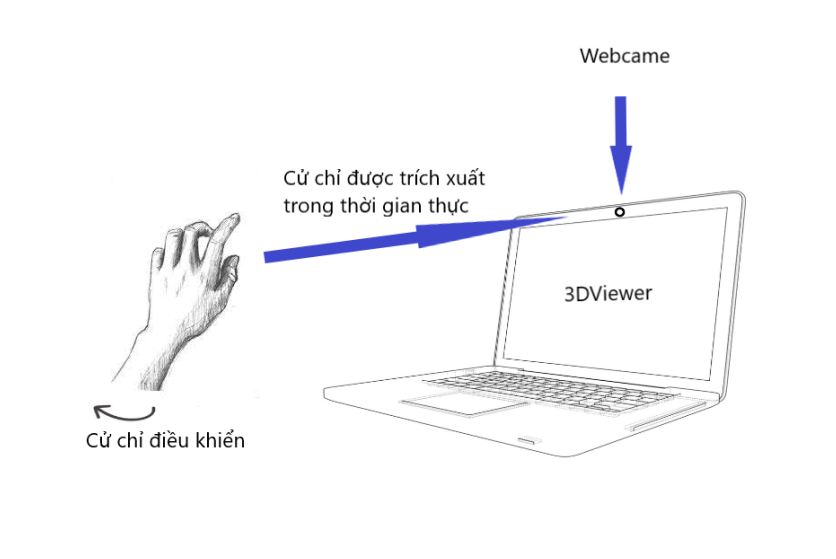
Phục vụ cho ứng dụng của dự án FPGA và các phòng ban có nhu cầu liên quan.

* 1. Khảo sát, phác thảo hệ thống

Đáp ứng được tác vụ thay thế điều khiển ứng dụng 3DViewer bằng cử chỉ tay thông qua hình ảnh cử chỉ tay được thu nhận từ Webcam máy tính, đảm bảo thuận tiện và dễ sử dụng cho người dùng.

Chức năng của hệ thống có thể được mô tả phác thảo như sau:

* Tín hiệu cử chỉ của người dùng được thu nhận trực tiếp, thông qua Webcam máy tính trong vùng được chỉ định (các mô tả về cử chỉ tương ứng với chức năng thực hiện sẽ được thể hiện trên màn hình cho người dùng dễ nắm bắt và sử dụng)
* Hình ảnh nhận được sau khi trích xuất sẽ đi qua một vài bước tiền xử lý cơ bản rồi được gửi đến server chứa mô hình dự đoán cử chỉ để đưa ra dự đoán.
* Dự đoán được ra sẽ tương ứng với một hành vi đã được quy định trước đó, sẽ được hệ thống đọc và tiến hành thực thi tương ứng với chức năng trong ứng dụng 3DViewer.



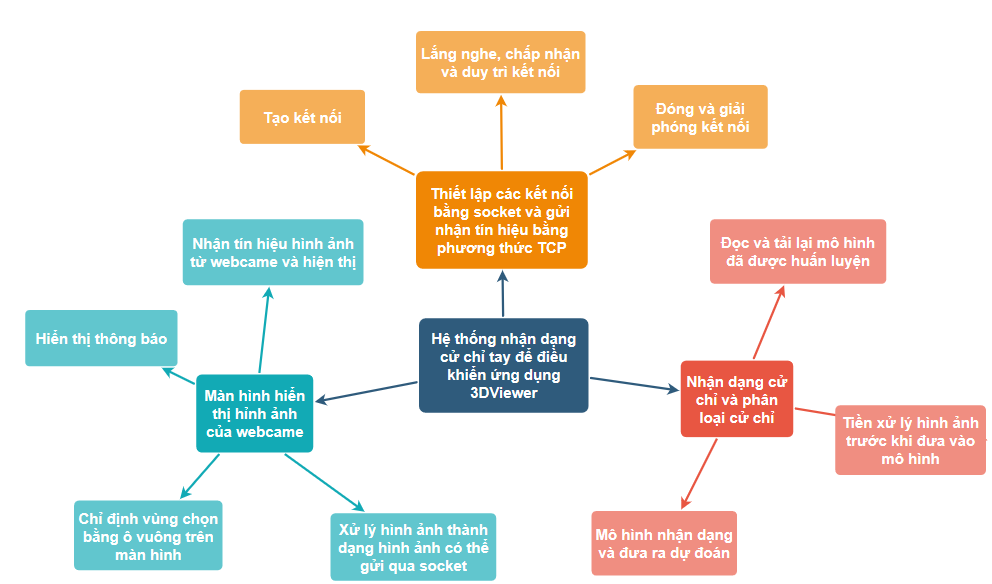
Hình 1. Mô tả cách thức hoạt động mức tổng quát.

* 1. Khó khăn và thuận lợi của dự án
     1. Khó khăn
* Việc tương thích giữa những module khác nhau thường dễ phát sinh lỗi.
* Cần đảm bảo tốc độ gửi nhận tín hiệu hợp lý để không tạo cảm giác mất kiểm soát cho người dùng.
* Nắm bắt và kiểm soát tốt cử chỉ tay là nhiệm vụ tương đối khó, trong việc xác định điểm bắt đầu và điểm kết thúc của một cử chỉ trong thời gian thực.
* Việc truyền tải và thực thi quá nhiều tầng xử lý sẽ dẫn đến độ trễ trong lúc thực thi của hệ thống.
  + 1. Thuận lợi
* Mô hình dự đoán đã làm việc rất tốt trong môi trường thử nghiệm thực tế.
* Có nhiều thư viện hỗ trợ trong việc liên kết gửi nhận tín hiệu điều khiển thông qua kết nối socket.
* Việc phát triển và mở rộng thêm chức năng cho dự án đã được phát triển trước đó bởi nhóm FPGA, sẽ luôn nhận được sự giúp đỡ và hướng dẫn bởi các thành viên trong nhóm.
* Nguồn tài liệu và kiến thức sử dụng đã được nhóm FPGA tổng hợp, sẽ cung cấp đầy đủ các thông tin cần thiết cho việc tích hợp chức năng điều khiển bằng cử chỉ.
* Việc tích hợp vào dự án đang chạy trong thực tế sẽ luôn nhận được những phản hồi và nhận xét tích cực và sát thực.

1. PHÂN TÍCH THIẾT KẾ
   1. Yêu cầu chức năng
      1. Phân tích chức năng

Chức năng chính của hệ thống là đọc và nhận dạng đúng cử chỉ người dùng gửi đến, từ đó suy ngược ra hành động người dùng muốn tác động lên ứng dụng 3DViewer và thực thi hành động tương ứng.

* Ràng buộc:
* Chỉ hoạt động hiệu quả đối với hình ảnh tay trái của người dùng.
* Cử chỉ được đặt trong vùng cho phép và từ góc nhìn trực diện.
* Cử chỉ sử dụng chỉ là cử chỉ đơn, không nên sử dụng một lúc nhiều cử chỉ.
* Chỉ nên sử dụng các cử chỉ trong bảng mô tả, không nên sáng tạo cử chỉ để điều khiển ứng dụng.
* Nên sử dụng trong môi trường đủ ánh sáng.
* Yêu cầu:
* Tốc độ nhận dạng và thực thi hành động phải tương thích với hành động trong thực tế, không quá nhanh và cũng không quá chậm so với cách sử dụng bàn phím và chuột thông thường.
* Các phím hay cử chỉ nên chọn sao cho có thể miêu tả một phần hành động muốn thực thi, làm cho người dùng dễ sử dụng, dễ liên tưởng.
* Không nên sử dụng các cử chỉ phức tạp khó biểu diễn.
  + 1. Biểu đồ phân rã chức năng

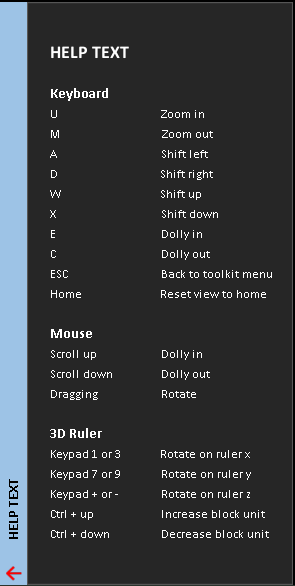


Hình 2. Biểu đồ thể hiện các chức năng chính có trong hệ thống.

Các chức năng của hệ thống gồm:

* Hiển thị và cung cấp thông tin cho người dùng thông qua màn hình và hình ảnh được thu nhận từ Webcam máy tính.
* Mô hình dự đoán sẽ có nhiệm vụ tiếp nhận hình ảnh đã được tiền xử lý và đưa ra dự đoán tương ứng. Mô hình này đã được huấn luyến để có thể nhận biết phát hiện cử chỉ trước đó
* Chức năng kết nối và gửi nhận tín hiệu với ứng dụng 3DViewer sẽ đóng vai trò cầu nối giao tiếp giữa hệ thống và ứng dụng theo mô hình client-Server kết nối thông qua socket sử dụng phương thức TCP để truyền nhận tín hiệu. Trong đó 3DViewer đóng vai trò là client và Server là nơi thực thi phần đưa ra dự đoán từ hình ảnh mà nhận được từ 3DViewer.
* Hiển thị thông báo về trạng thái kết nối và lỗi phát sinh cho người dùng qua màn hình.
  + 1. Các phím tắt và hành động chuột muốn điều khiển

Các phím tắt và hành động chuột được ứng dụng 3DViewer sử dụng trong việc điều khiển hiển thị hình ảnh 3D.



Hình 3. Bảng mô tả chức năng phím tắt và hành động chuột của ứng dụng 3DViewer.

Bộ ngôn ngữ cử chỉ của người câm điếc có thể biểu diễn được tất cả các chữ cái trong bảng chữ cái thành cử chỉ, bảng dưới đây sẽ mô tả chức năng các cử chỉ được sử dụng và hình ảnh biểu diễn của cử chỉ đó.

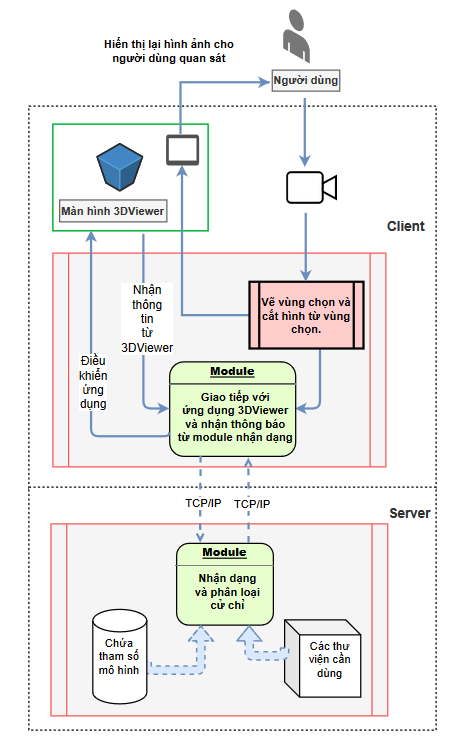
Bảng 1. Hình ảnh cử chỉ tương ứng với chữ cái.

|  |  |
| --- | --- |
| **Phím tắt và hành động chuột** | **Hình ảnh cử chỉ tương ứng** |
| + Phím U: phóng to hình ảnh giữ nguyên góc nhìn. |  |
| + Phím M: thu nhỏ hình ảnh giữ nguyên góc nhìn. |  |
| + Phím A: di chuyển hình ảnh sang trái. |  |
| + Phím D: di chuyển hình ảnh sang phải. |  |
| + Phím W: di chuyển hình ảnh lên trên. |  |
| + Phím X: di chuyển hình ảnh xuống dưới. |  |
| + Phím E: phóng to hình ảnh theo phạm vi quan sát |  |
| + Phím C: thu nhỏ hình ảnh theo phạm vi quan sát. |  |
| + Phím ESC: được thay thế bằng phím B (back), sẽ mở lại lại menu toolkit ban đầu. |  |
| + Phím H: đưa ứng trở về màn hình Home, màn hình khi bắt đầu của ứng dụng 3DViewer. |  |
| + Phím L: nhấn chuột trái và di chuyển qua bên trái một khoảng 5 pixel trên màn hình. |  |
| + Phím R: nhấn trái và di chuyển qua bên phải một khoảng 5 pixel trên màn hình. |  |
| + Phím V: nhấn chuột trái di chuyển xuống một khoảng 5 pixel trên màn hình. |  |
| + Phím Y: nhấn chuột trái di chuyển lên một khoảng 5 pixel trên màn hình. |  |

*Lưu ý :*

* Có hai hành động Dolly in và Dolly out đã được điều khiển tốt thông qua cử chỉ tương tác đối với bàn phím nên không cần phải lặp lại trường hợp đó ở chuột.
* Hành động nhấn và kéo chuột sẽ được tái hiện lại theo bốn hướng cố định là lên xuống, trái, phải tương ứng với bốn phím Y, V, L, R.
  1. Kiến trúc hệ thống

Kiến trúc tổng quan của hệ thống:



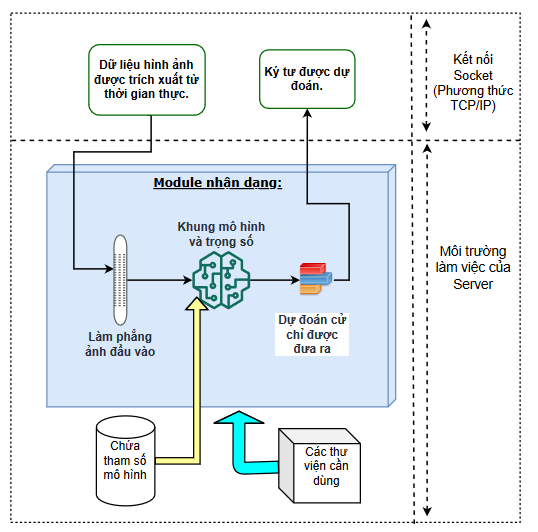
Hình 4. Biểu đồ luồng thực thi của hệ thống.

Hệ thống sẽ được phân làm 2 phần chính theo mô hình client-server, giao tiếp thông qua socket bằng phương thức TCP/IP. Phía server sẽ nắm giữ phần lớn những duyết định đưa ra, còn ở phía client (hay 3DViewer) sẽ được tích hợp thêm module điều khiển và giao tiếp, là nơi gửi nhận dữ liệu với server và thực thi quyết định được đưa ra ở server. Đồng thời ở client sẽ lập trình thêm giao diện chức năng tích hợp vào gian diện đã có trước đó.

* 1. Kiến trúc thành phần

Kiến trúc của mỗi module thành phần trong hệ thống:

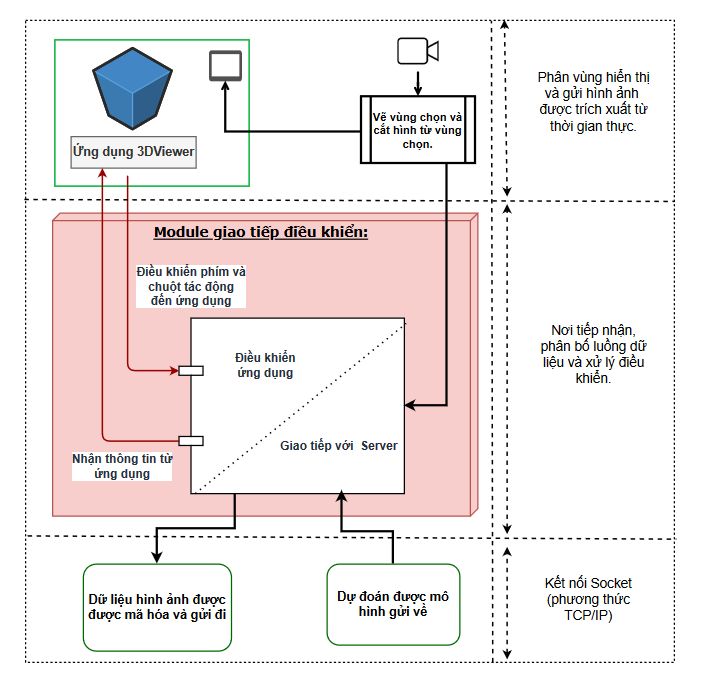
* Module nhận dạng ở server:



Hình 5. Module nhận dạng hình ảnh.

Đây sẽ là module chính của hệ thống, là nơi xử lý các nhiệm vụ phức tạp về nhận dạng và phát hiện các đặc trưng của hình ảnh để có thể đưa ra dự đoán chính xác cho hệ thống. Mọi tín hiệu gửi đến và truyền đi đều được thực hiện trong thời gian thực, nên cần quan tâm đến tốc độ thực thi và sự toàn vẹn của dữ liệu trong suốt quá trình trao đổi.

* Module giao tiếp và điều khiển ở phía client:

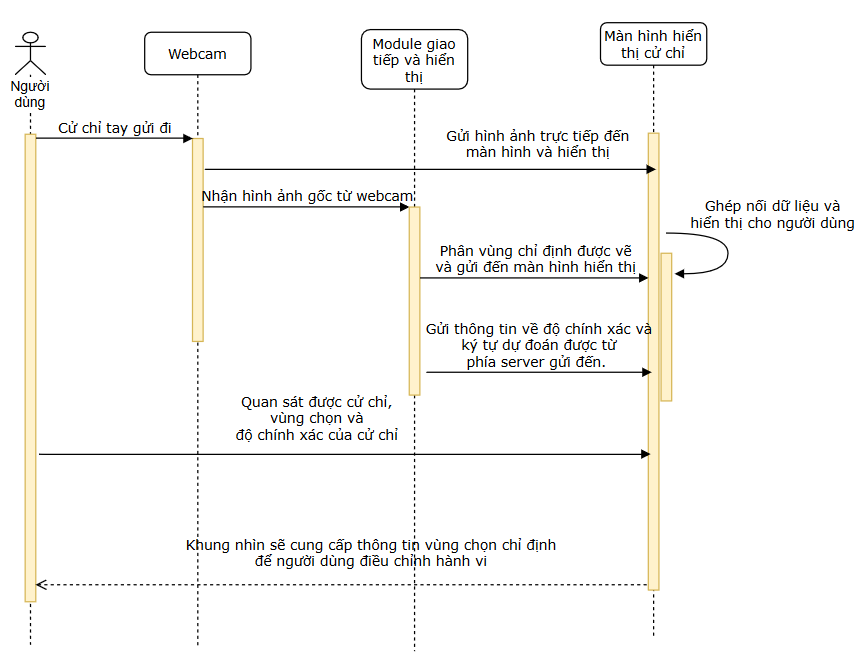


Hình 6. Module điều khiển của hệ thống.

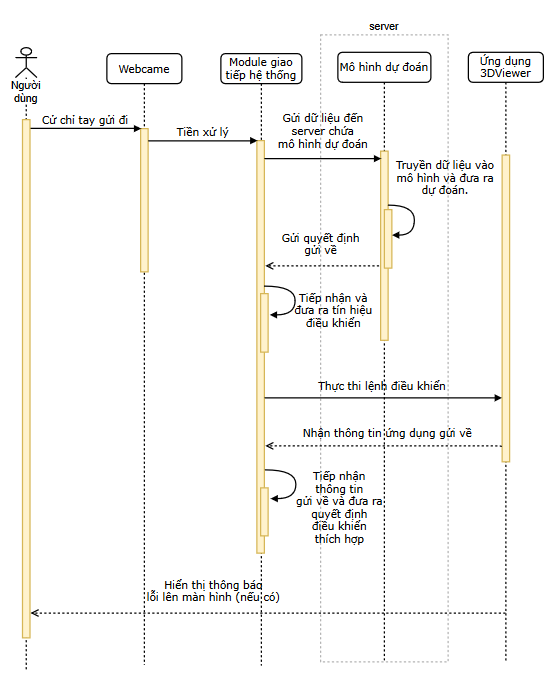
Module giao tiếp điều khiển là nơi xử lý hầu hết các sự kiện điều khiển tác động đến ứng dụng 3DViewer, luồng dữ liệu chính được đi qua, được tiếp nhận và xử lý thay đổi trong thời gian thực.

* 1. Biểu đồ tuần tự

Mô tả hoạt động của người dùng tác động lên hệ thống được hiển thị qua màn hình chính của 3DViewer trong đó có vùng hiển thị cử chỉ người dùng giúp người điều khiển có thể tự điều chỉnh hành vi và vị trí của bàn tay sao cho hợp lý.



Hình 7. Biểu đồ tuần tự người dùng giao tiếp với màn điều chỉnh hành vi và hiển thị thông tin.

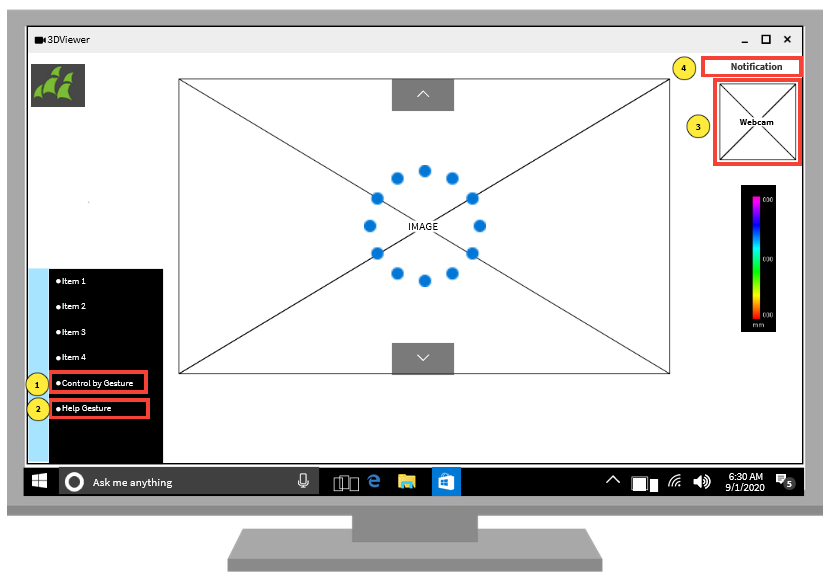


Hình 8. Biểu đồ tuần tự người dùng giao tiếp với

3DViewer thông qua cử chỉ.

Biểu đồ trên miêu tả những sự kiện và hiển thị mà người dùng giao tiếp với màn hình 3DViewer, luồng đi của dữ liệu và việc các quyết định điều khiển được đưa ra và điều chỉnh trong suốt quá trình thực thi của hệ thống.

* 1. Giao diện ứng dụng
     1. Giao diện tổng quan

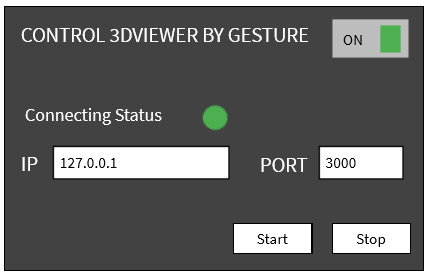


Hình 9. Giao diện 3DViewer được tích hợp thêm chức năng điều khiển bằng cử chỉ tay.

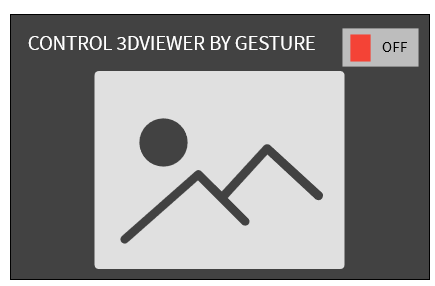
Giao diện của ứng dụng 3DViewer trước đó sẽ được thiết kế và tích hợp thêm các giao diện ở phần bôi đỏ, gồm:

1. Mục chức năng điều khiển bằng cử chỉ - Control By Gesture ở TOOLKIT MENU.
2. Mục chức năng trợ giúp cử chỉ - Help Gesture ở TOOLKIT MENU.
3. Màn hình hiển thị cử chỉ tay được trích xuất từ Webcam.
4. Nơi hiển thị thông báo cho người dùng.
   * 1. Giao diện bảng điều khiển kết nối với server

Đây là bảng điều khiển xuất hiện sao khi nhấn vào chức năng **Control by Gesture:**

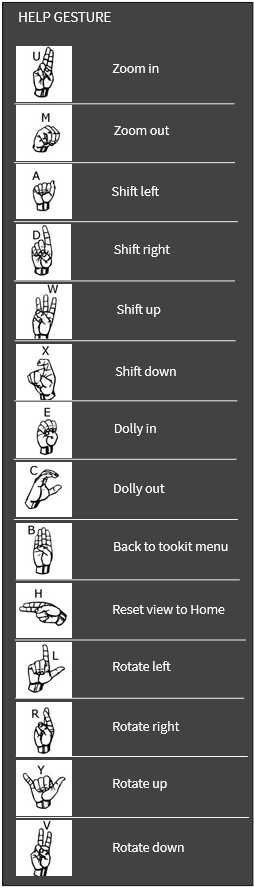


Hình 10. Bảng thiết lập kết nối với server đang ở trạng thái bật.



Hình 11. Bảng điều khiển đang ở trạng thái tắt.

* + 1. Giao diện bảng hiển thị cử chỉ trợ giúp



Kéo thả chuột

Hình 12. Bảng Help Gesture hiển thị cử chỉ trợ giúp.

* 1. Công nghệ sử dụng
     1. Môi trường và ngôn ngữ sử dụng

Việc phát triển và tích hợp chức năng được phát triển dựa trên 2 ngôn ngữ lập trình Python và C# cùng với các công cụ được nêu ở bảng bên dưới:

Bảng 2. Phiên bản của ngôn ngữ và công cụ sử dụng.

|  |  |
| --- | --- |
| **Ngôn ngữ/ Công cụ** | **Phiên bản** |
| Python | 3.7.6 (x64) |
| C# | 7.3 (x64) |
| Hệ điều hành | Window 10 Pro (x64) |
| Unity | 2019.3.9f1 (x64) |
| Microsoft Visual Studio Community 2019 | 16.7.2 (x64) |
| Visual Studion Code | 1.49.0 (x64) |
| Anconda – công cụ cho khoa học dữ liệu | 4.8.4 |
| Jupyter NoteBook | 6.1.1 |
| Spyder | 4.1.4 |

Ngoài ra, trong quá trình huấn luyện mô hình có sử dụng Colab – là một dịch đám mây miễn phí, cho phép viết và thực thi code Python ngay trên trình duyệt và có hỗ trợ GPU miễn phí.

Ở phía giao diện 3DViewer, được viết và mở rộng thêm tính năng điều khiển bằng cử chỉ dựa trên phần mềm đã được phát triển trước đó.

* + 1. Các thư viện chính

Một số thư viện chính được sử dụng cho mô hình nhận dạng cử chỉ thông qua hình ảnh:

Bảng 3. Phiên bản của thư viện.

|  |  |
| --- | --- |
| **Thư viện** | **Phiên bản** |
| Tensorflow | 2.1.0 |
| Keras | 2.3.1 |
| Numpy | 1.18.5 |
| OpenCV (opencv-python) | 3.4.2.16 |

Các thư viện trên được sử dụng chủ yếu cho việc xây dựng và phát triển mô hình nhận dạng cử chỉ ở phía server, được sử dụng ngay từ quá trình thu thập dư liệu, huấn luyện cho đến việc triển khai ở phía server.

* + 1. Mô hình nhận dạng cử chỉ

Mô hình nhận dạng hình hay mô hình phân loại hình ảnh được xây dựng trên mô hình mạng nơ-ron tích hợp chập và triển khai bằng kiến trúc VGG16 – một trong những kiến trúc mạng nơ-ron hiện đại và hoạt động hiệu quả đối với nhiệm vụ phân loại ảnh.



Hình 13. Kiến trúc mạng VGG16 phân loại 1000 đối tượng.

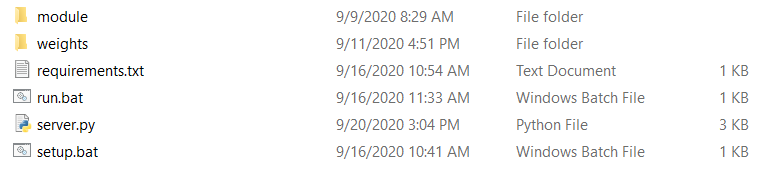
Quá huấn luyện luyện sử dụng bộ dữ liệu từ Kaggle và một số hình tự chụp, được điều chỉnh theo nhiệm vụ và mục tiêu mục tiêu khi dùng cử chỉ để điều khiển ứng dụng 3DViewer.

1. TRIỂN KHAI VÀ HƯỚNG DẪN SỬ DỤNG
   1. Thiết lập môi trường

**Lưu ý:** Ngôn ngữ và phần mềm hỗ trợ đã nêu ở mục 2.6, với các thông số về phiên bản và các phần mềm và ngôn ngữ đã sử dụng trong quá trình xây dựng và tích hợp ứng dụng vào 3DViewer.

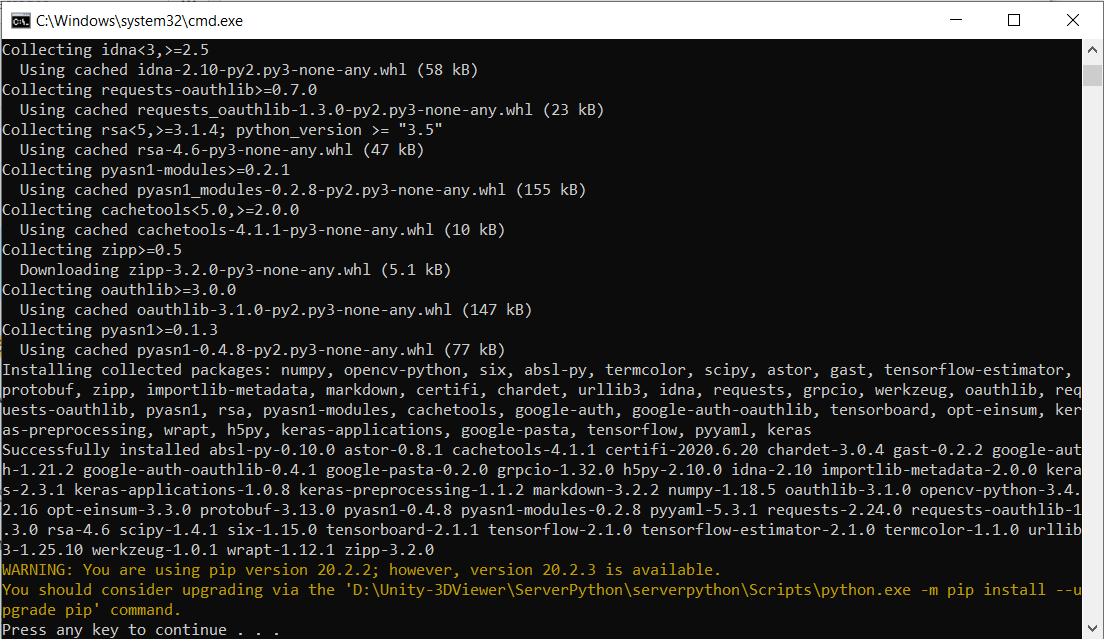
Phiên bản server Python và ứng dụng 3DViewer hiện tại đang chạy ổn định trên hệ điều hành Window. Đối với server Python thì cần cài đặt Python và thiết lập biến môi trường hoặc sử dụng Anaconda đã được tích hợp sẵn Python trong lúc cài đặt.

Để có thể chạy server Python, đầu tiên ta chạy file **setup.bat** trong môi trường đã có cài đặt Python và module pip, để thiết lập các thư viện cần thiết cho server. Sau đó chỉ cần chạy file **run.bat** để khởi động server Python.



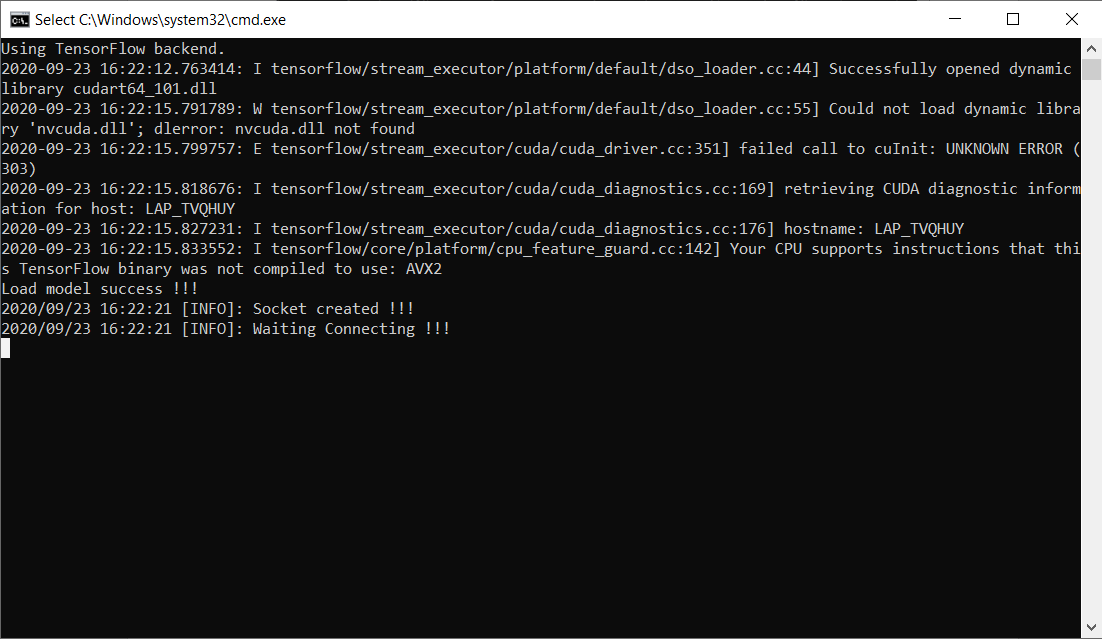
Hình 14. Thư mục của server Python.

Quá trình thiết lập môi trường diễn ra hơi chậm, do phải cài đặt nhiều thư viện liên qua phục vụ cho server nhận dạng, bên dưới hình ảnh cửa sổ sau khi hoàn tất việc thiết lập các thư viện:



Hình 15. Cửa sổ hoàn thành sau khi chạy file setup.bat

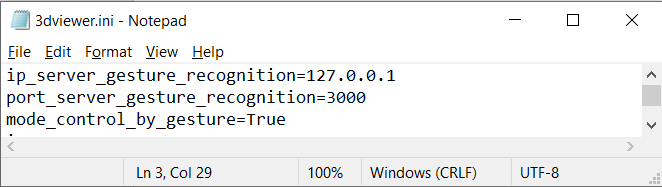
Sau khi quá trình cài đặt thư viện hoàn tất thì có thể chạy file run.bat để khởi động server, sẵn sàng cho nhiệm vụ nhận dạng cử chỉ.



Hình 16. Cửa sổ thông báo server đã sẵn sàng và đang đợi kết nối đến.

* 1. Các cách kích hoạt chức năng
     1. Kích hoạt chức năng với file 3dviewer.ini

Kích hoạt chức năng bằng file 3dviewer.ini, dưới đây là một số thuộc tính nhằm kích hoạt chức năng cùng lúc với 3DViewer được khỏi chạy.

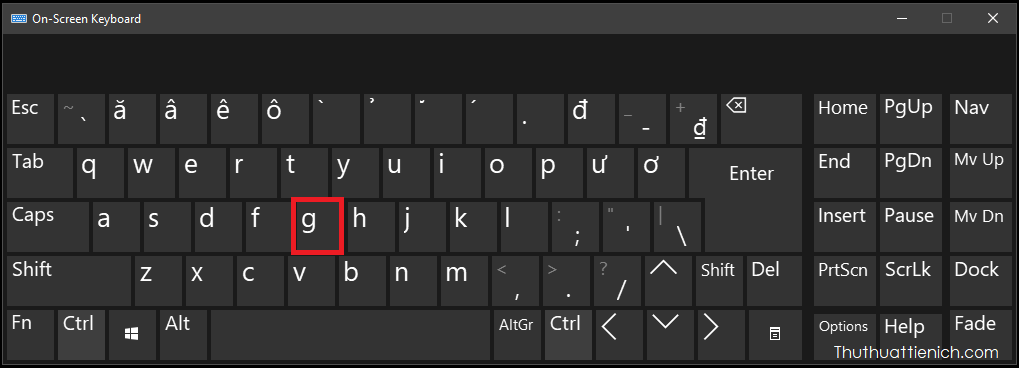


Hình 17. Các cờ được bật ở file 3dviewer.ini

*Trong đó:*

* ip\_server\_gesture\_recognition=127.0.0.1: là địa chỉ IP của server nhận dạng.
* port\_server\_gesture\_recognition=3000: là cổng mặc định của server nhận dạng có giá trị trong khoảng 0-65535.
* mode\_control\_by\_gesturre: là cờ bật/tắt chế độ điều khiển bằng cử chỉ, chỉ nhận hai giá trị True/False tương ứng.
  + 1. Kích hoạt chức sử dụng phím tắt “G”

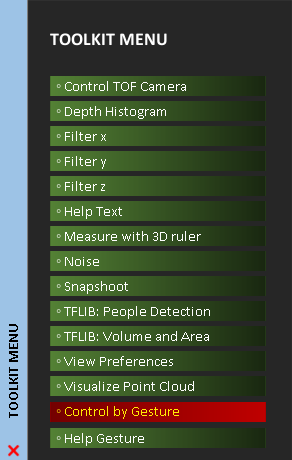
Chức năng điều khiển 3DViewer bằng cử chỉ có thể bật/tắt nếu sử dụng phím “G”, rất tiện sử dụng và linh hoạt cho người dùng trong quá trình sử dụng. Có thể dễ dàng bật/tắt ứng dụng trong trường hợp không có chuột và muốn chuyển qua sử dụng điều khiển bằng cử chỉ ngay lập tức.



Hình 18. Phím tắt “G”.

* + 1. Kích hoạt chức năng sử dụng chuột

Chức năng điều khiển bằng cử chỉ tay cũng có thể tìm thấy và kích hoạt như những chắc năng đã có trong 3DViewer thông qua chuột, hình dưới mô tả chức năng điều khiển bằng cử chỉ - Control By Gesture trong TOOLKIT MENU:



Hình 19. Chức năng điều khiển bằng cử chỉ trong TOOLKIT MENU.

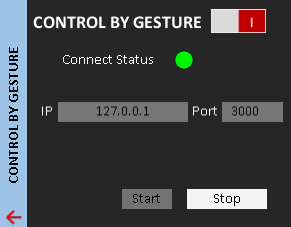
Sau khi chọn chức năng một bảng điều khiển của chức năng sẽ được hiện lên như hình bên dưới để người dùng bắt đầu.



Hình 20. Bảng điều khiển chức năng ở trạng tái tắt.

Bảng 4. Mô tả hành động bật/ tắt ứng dụng của người dùng.

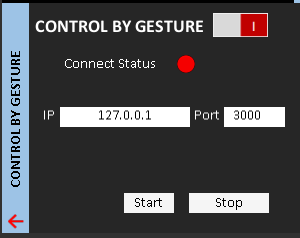
|  |  |  |  |
| --- | --- | --- | --- |
| **Tên hành động** | **Mô tả** | **Thành công** | **Thất bại** |
| Bật chức năng điều khiển bằng cử chỉ | Webcam sẽ được bật và hiển thị ở góc phải màn hình 3DViewer, sau khi kết nối thành công đến server thì quá trình. | Hiển thị bảng điều khiển như *Hình 21* với đèn trạng thái bật xanh. | Hiển thị bảng điều khiển như *Hình 22* với đèn trạng thái bật đỏ. |
| Tắt chức năng điều khiển bằng cử chỉ | Sẽ đồng thời đóng kết nối và tắt Webcam. | Trở về giao diện như *Hình 20* |  |



Hình 21. Bảng điều khiển đang ở trạng thái bật và đang kết nối server.

Bảng 5. Mô tả hành động đóng/mở kết nối đến server của người dùng.

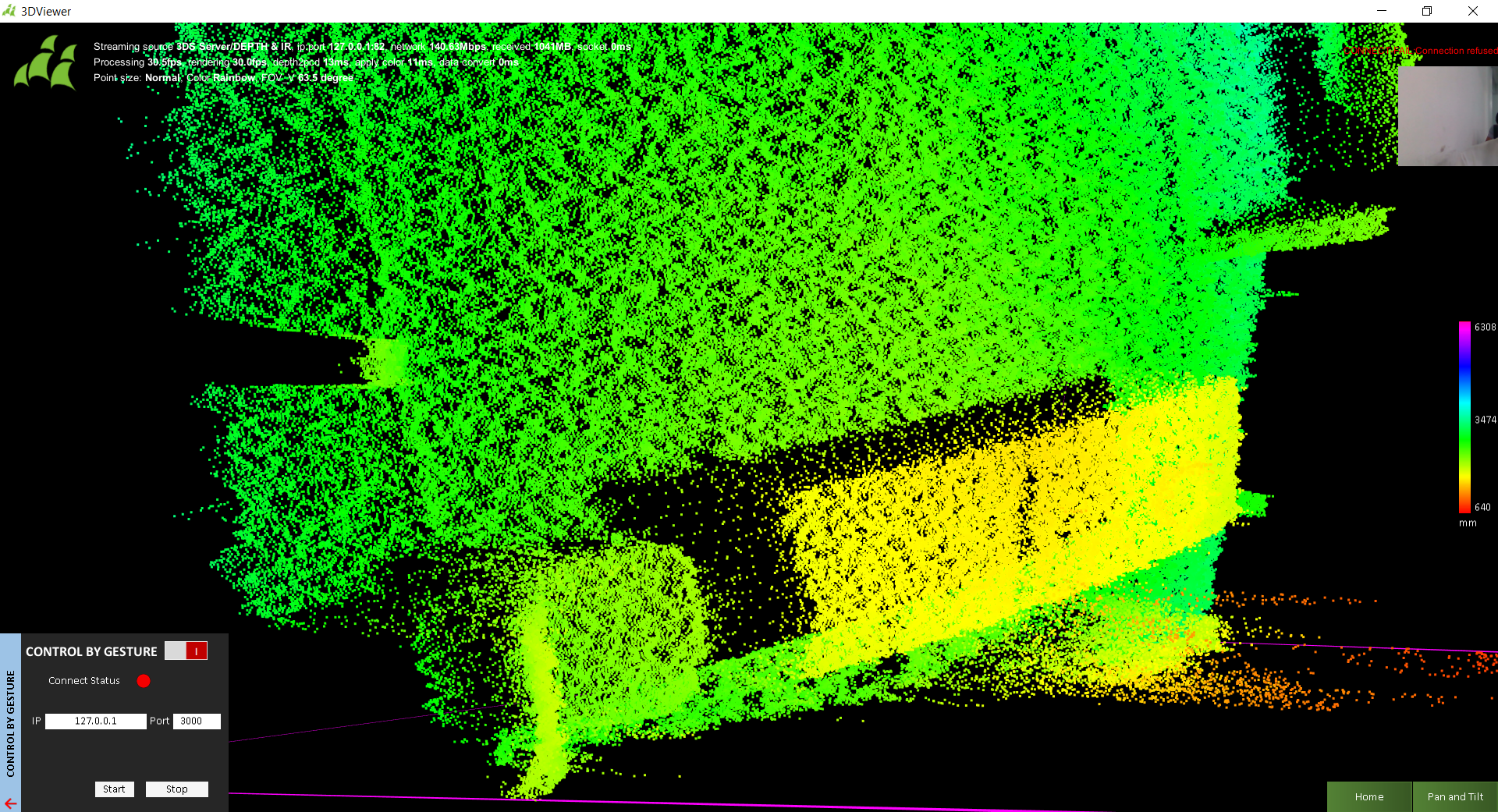
|  |  |  |  |
| --- | --- | --- | --- |
| **Tên hành động** | **Mô tả** | **Thành công** | **Thất bại** |
| Dừng kết nối đến server. | Nhấn nút “Stop” để đóng kết nối đến server. | Đèn báo kết nối chuyển đỏ, nút “Start” và các ô nhập IP và Port được kích hoạt. |  |
| Mở kết nối đến server. | Sau qua trình kiểm tra IP và Port hợp lệ thì nút “Start” được kích hoạt, cho phép người dùng mở kết nối. | Đèn báo kết nối chuyển xanh và nhấp nháy, nút “Start” cũng như ô nhập IP và Port bị vô hiệu hóa. | Đèn báo kết nối chuyển đỏ và lỗi được thông báo ở góc phải màn hình. |



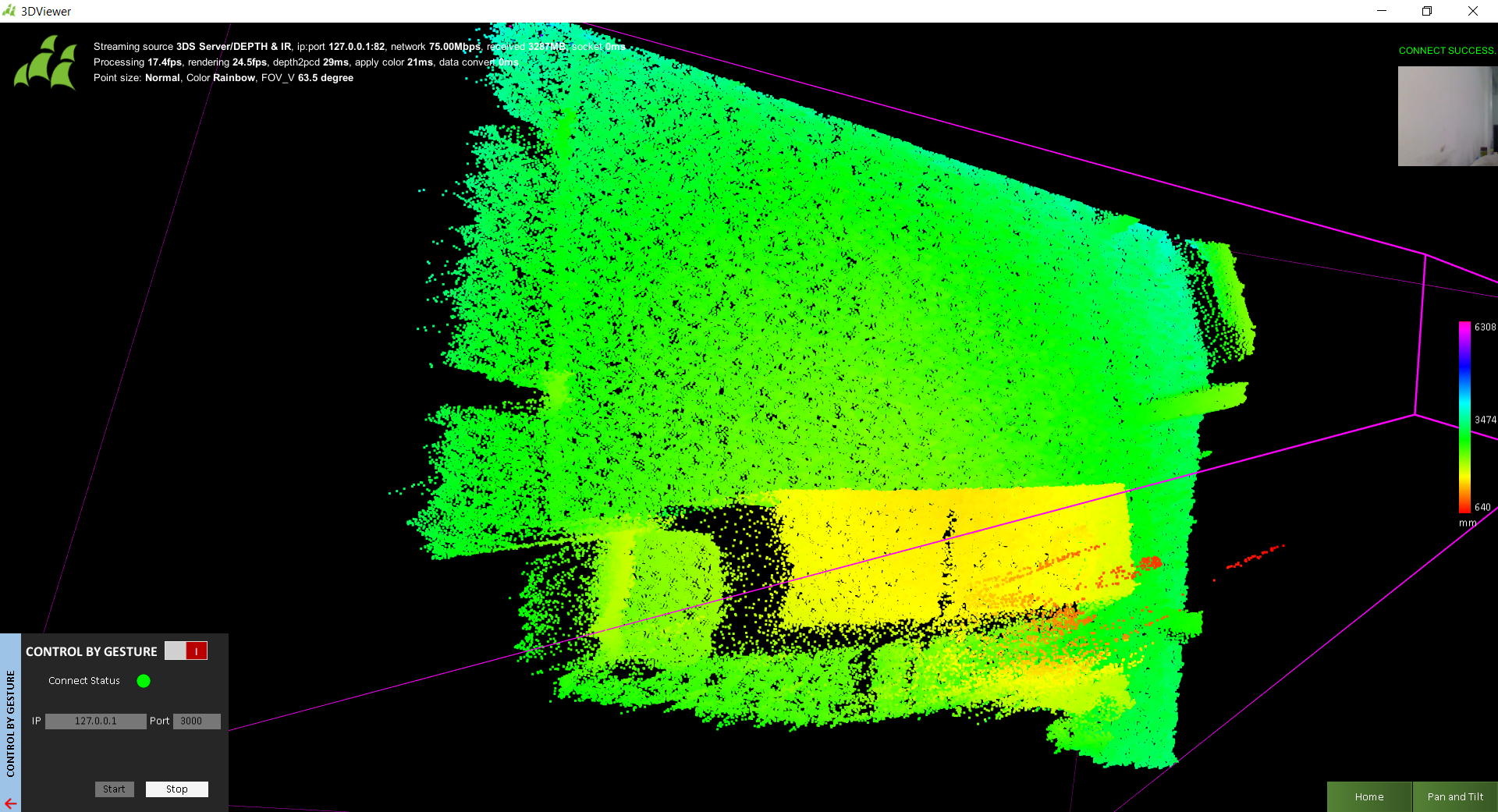
Hình 22. Bảng điều khiển đang ở trạng thái bật và không có kết nối server.

* 1. Hiển thị thông báo và hình ảnh

Trong quá trình bật chức năng điều khiển bằng cử chỉ thì Webcam luôn được bật và hình ảnh trích xuất từ Webcam được hiển thị ở trên bên phải màn hình 3DViewer, để người dùng có thể quan sát và điều chỉnh cử chỉ của mình sao cho hợp lý nhất. Tất cả các thông báo thành công hoặc thất bại cũng được hiển thị ở góc trên bên phải màn hình 3DViewer.



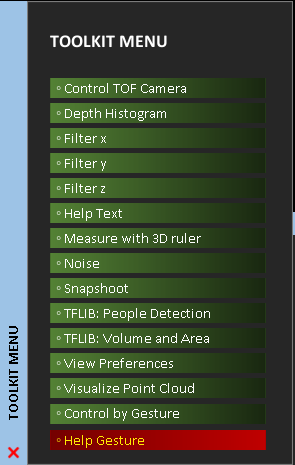
Hình 23. Thông báo ở góc phải 3DViewer khi kết nối thất bại.



Hình 24. Thông báo ở góc phải 3DViewer khi kết nối thành công.

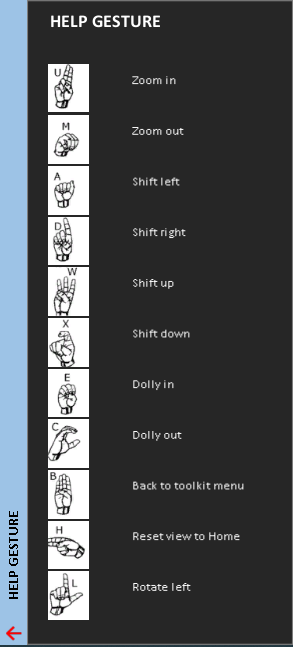
* 1. Bảng trợ giúp cử chỉ

Đây là bảng tổng hợp tất cả các cử chỉ cùng với chức năng được mô tả, dùng để nhắc nhở hay xem lại cử chỉ nếu quên hoặc chưa quen trong lúc sử dụng.



Hình 25. Bảng trợ giúp cử chỉ trong TOOLKIT MENU.

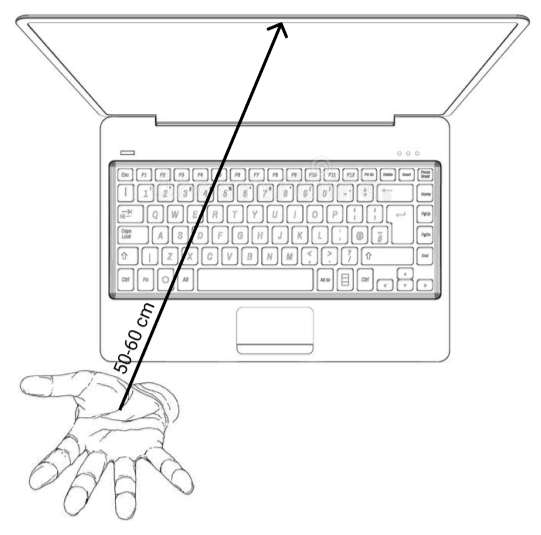
Bảng trợ giúp cử chỉ sẽ được mở như hình bên dưới, có hình ảnh miêu tả cử chỉ và mô tả về chức năng của cử chỉ đó. Kéo thả chuột ở bảng để xem theo nhiều thông tin hơn về các cử chỉ, hiện tại đang sử dụng 15 cử chỉ cho các chức năng của 3DViewer. Và có thể xem bảng trong quá trình chức năng điều khiển bằng cử chỉ được thực thi.



Hình 26. Hình ảnh và thông tin về chức năng của cử chỉ.

* 1. Đánh giá và lưu ý cho người dùng

Người dùng nên sử dụng tay trái để điều khiển ứng dụng 3DViewer vì mô hình được huấn luyện và hoạt động tốt đối với cử chỉ được biểu diễn bằng tay trái. Khoảng cách giữa tay và Webcam nên trong khoảng 50-60 cm, và có thể quan sát hình ảnh được hiển thị ở góc trên bên phải màn hình 3Dviewer để có thể điều chỉnh sao cho hợp lý, để hình ảnh nhận về được thể hiện rõ ràng với kích thước đầy đủ, không quá gần cũng không quá xa so với Webcam, gây nhầm lẫn và không chính xác trong quá trình đưa ra dự đoán hình ảnh ở phía server.



Hình 27. Khoảng cách giữa tay và Webcam.

Mô hình vẫn có thể bị kích hoạt nếu gặp những hình ảnh có các đặc trưng thích hợp, và đủ nhiều để có thể kích hoạt một cử chỉ trong khi đưa ra dự đoán, nên cần tránh các đối tượng không cần thiết xuất hiện trong khung hình của Webcam, để giảm độ nhiễu và những kích hoạt không đáng có.

DANH MỤC TÀI LIỆU THAM KHẢO

1. *Convolutional Neural Networks for Visual Recognition*, URL: <https://cs231n.github.io/neural-networks-1/>.
2. Phân tích và thiết kế phần mềm “Hệ thống thông tin Quản lý cấp giấy phép Kinh doanh”.
3. *Socket Programing in Python*, URL: <https://realpython.com/python-sockets/>.
4. Nguyễn Thanh Tuấn, *Deep Learning cơ bản*, version2, August 2020, URL: <https://drive.google.com/file/d/1lNjzISABdoc7SRq8tg-xkCRRZRABPCKi/view>.
5. *Socket TCP/IP sử dụng C#*, URL: <https://docs.microsoft.com/en-us/dotnet/api/system.net.sockets.tcpclient?view=netcore-3.1>