

# Nhận dạng cử chỉ tay bằng MPU6050 và ESP8266 trong thời gian thực

Vũ Giang Nam\*, Mẫn Bá Sâm<sup>†</sup>, Phạm Thanh Phương<sup>‡</sup>, Nguyễn Quang Tiến<sup>‡</sup>,

\*Khoa Công nghệ Thông tin, Trường Đại học Đại Nam Hà Nội, Việt Nam

Email: nguyenvana@xyz.edu.vn

**Abstract**—Nhận dạng cử chỉ tay là một lĩnh vực quan trọng trong tương tác người-máy, có nhiều ứng dụng trong điều khiển thông minh, y tế và trò chơi. Trong dự án này, chúng tôi sử dụng cảm biến MPU6050 kết hợp với vi điều khiển ESP8266 để thu thập dữ liệu cử chỉ tay, truyền dữ liệu đến server Flask và xử lý bằng mô hình LSTM. Hệ thống hoạt động trong thời gian thực và trực quan hóa kết quả nhằm hỗ trợ nghiên cứu và phát triển ứng dụng điều khiển không tiếp xúc.

## I. GIỚI THIỆU

Trong những năm gần đây, công nghệ nhận dạng cử chỉ tay đã phát triển mạnh mẽ nhờ vào sự kết hợp của cảm biến IMU (Inertial Measurement Unit) và các thuật toán học sâu. Dự án này nhằm mục đích xây dựng một hệ thống nhận diện cử chỉ tay dựa trên dữ liệu gia tốc và con quay hồi chuyển từ cảm biến MPU6050, truyền qua WiFi bằng ESP8266 và xử lý bằng mô hình học sâu.

## II. PHẦN CỨNG SỬ DỤNG

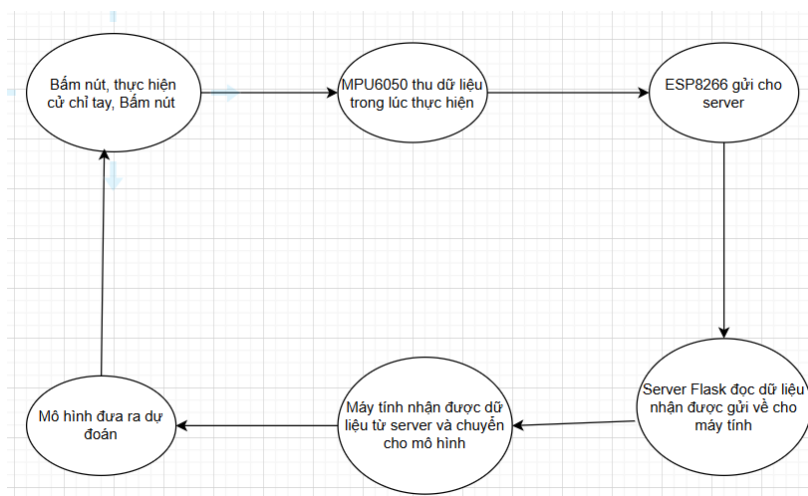
- ESP8266:** Vi điều khiển hỗ trợ WiFi, chịu trách nhiệm thu thập và gửi dữ liệu cảm biến.
- MPU6050:** Cảm biến IMU có tích hợp gia tốc kế và con quay hồi chuyển.
- Nút nhấn:** Dùng để bắt đầu và kết thúc quá trình ghi nhận dữ liệu cử chỉ.
- Máy tính chạy Flask Server:** Xử lý dữ liệu, dự đoán cử chỉ và hiển thị kết quả trực quan.

## III. PHẦN MỀM SỬ DỤNG

- Python 3.x:** Ngôn ngữ lập trình chính để phát triển server và xử lý dữ liệu.
- Flask:** Framework web giúp nhận dữ liệu từ ESP8266 và trả kết quả dự đoán.
- TensorFlow/Keras:** Dùng để xây dựng và huấn luyện mô hình LSTM nhận dạng cử chỉ.
- Matplotlib:** Vẽ biểu đồ trực quan hóa dữ liệu cảm biến để phân tích.
- Arduino IDE:** Soạn thảo và nạp chương trình cho ESP8266.

## IV. KIẾN TRÚC HỆ THỐNG

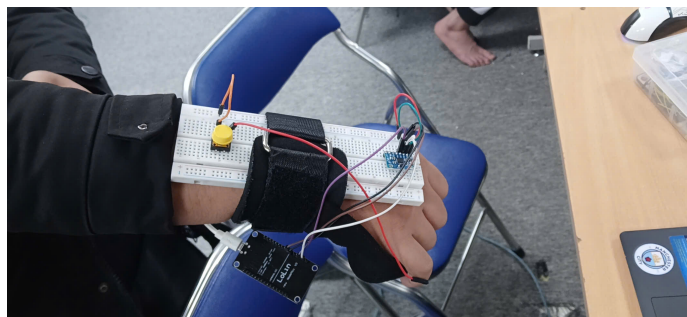
Hệ thống bao gồm ba thành phần chính: cảm biến MPU6050, ESP8266 và server Flask. Cảm biến thu thập dữ liệu cử chỉ, ESP8266 xử lý và gửi dữ liệu đến server, nơi mô hình LSTM thực hiện phân loại cử chỉ.



figureSơ đồ kiến trúc hệ thống

## V. QUY TRÌNH HOẠT ĐỘNG

- ESP8266 đọc dữ liệu từ MPU6050 với tần suất 20Hz và gửi đến server thông qua giao thức HTTP GET.
- Server Flask lưu trữ dữ liệu, tiền xử lý, chuẩn hóa, và truyền vào mô hình LSTM.
- Sau khi dự đoán, kết quả được hiển thị trên giao diện web hoặc dưới dạng đồ thị phân tích tín hiệu.



## Quy trình lắp đặt

### VI. GIỚI THIỆU BỘ DỮ LIỆU ĐỂ HUẤN LUYỆN

#### A. Danh sách các hành động có trong bộ dữ liệu

Bộ dữ liệu được thu thập từ 4 người tham gia, mỗi người thực hiện 10 loại cử chỉ:

- 1) Nâng tay lên
- 2) Hạ tay xuống
- 3) Xoay cổ tay theo chiều kim đồng hồ
- 4) Xoay cổ tay ngược chiều kim đồng hồ
- 5) Vẫy tay từ trái qua phải
- 6) Vẫy tay từ phải qua trái
- 7) Đưa tay ra trước
- 8) Kéo tay ra sau
- 9) Lắc tay nhanh
- 10) Gõ tay (flick)

#### B. bộ dữ liệu sau khi thu

- Sau khi thu, ta sẽ có file như sau

A	B	C	D	E	F	G	H	I	J	K
Time	AccelX	AccelY	AccelZ	GyroX	GyroY	GyroZ	Temp	Marker	Label	
83326	0	0	0	0	0	0	0	0 START	0	
83523	1.66	-0.02	9.83	-0.03	-0.15	-0.08	32.16	DATA	0	
83667	2.33	0.16	9	-0.1	0.01	-0.04	32.16	DATA	0	
83791	2.24	-0.1	9.85	-0.13	-0.03	-0.03	32.14	DATA	0	
83915	2.37	-0.32	10.36	0.18	0.03	0.05	32.16	DATA	0	
84039	2.15	0.13	10.02	0.35	-0.08	-0.03	32.14	DATA	0	
84166	2.51	0.73	9.53	0.54	0.1	0.01	32.14	DATA	0	
84286	2.22	1.41	9.59	0.65	0.05	0	32.16	DATA	0	
84409	2.4	2.3	8.88	0.66	0.06	-0.06	32.15	DATA	0	
84532	2.11	3.16	8.39	0.53	0.1	-0.07	32.15	DATA	0	
84655	1.97	3.65	8.54	0.49	-0.01	-0.07	32.17	DATA	0	
84779	2.03	4.12	8.47	0.52	0.07	-0.02	32.15	DATA	0	
84905	1.92	4.91	8.13	0.51	0.08	0.05	32.17	DATA	0	
85029	1.89	5.4	7.58	0.48	0.11	-0.02	32.15	DATA	0	
85153	1.82	5.86	7.09	0.42	0.14	-0.01	32.16	DATA	0	
85276	1.49	6.36	7.01	0.37	0.16	-0.01	32.19	DATA	0	
85578	1.48	7.04	6.19	0.18	0.06	0	32.16	DATA	0	
85663	1.55	7.25	5.83	0.06	-0.03	0.02	32.18	DATA	0	
85786	1.75	7.07	6.14	-0.05	0.03	-0.02	32.16	DATA	0	
85910	1.51	7.11	6.19	-0.02	0.02	-0.05	32.18	DATA	0	
86035	1.51	7.19	6	-0.12	0.02	-0.01	32.18	DATA	0	
86159	1.69	6.8	6.09	-0.37	0.16	-0.07	32.16	DATA	0	
86283	1.17	6.71	6.47	-0.38	0.11	-0.04	32.16	DATA	0	
86407	0.98	6.35	6.89	-0.6	-0.15	-0.06	32.19	DATA	0	
86530	1.78	5.43	7.26	-0.66	-0.05	0.01	32.16	DATA	0	
86655	1.71	4.76	7.98	-0.56	0.02	-0.11	32.17	DATA	0	

#### Thành quả lắp đặt

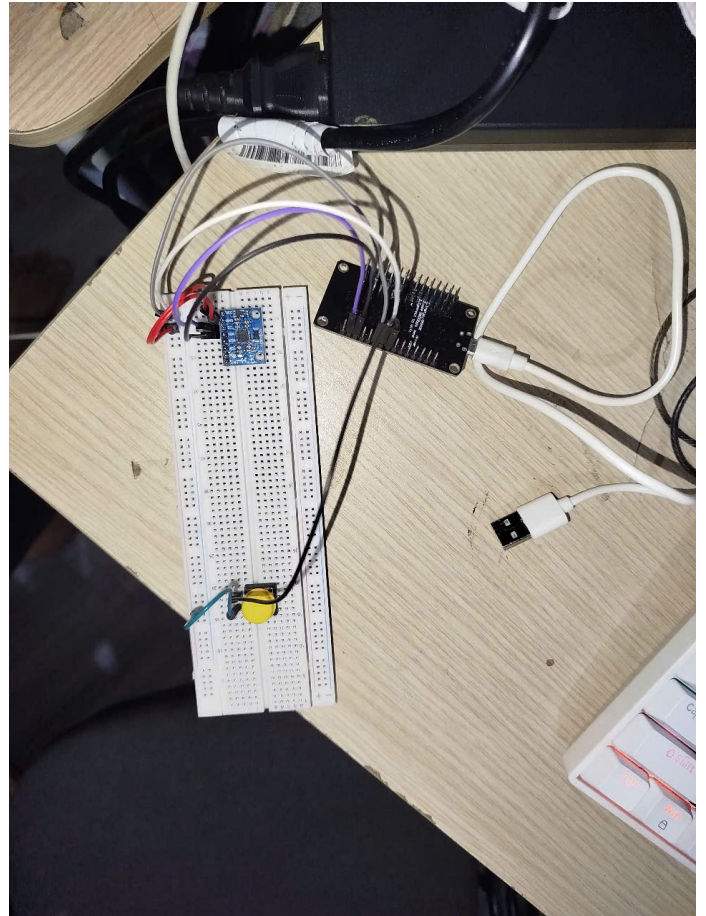
- trong file có 9 cột dữ liệu là: AccelX, AccelY, AccelZ, GyroX, GyroY, GyroZ, Temp, Marker và 1 cột label được tạo ra thêm để gán nhãn cho hành động tương ứng

### VII. MÔ HÌNH HỌC MÁY

Dữ liệu cử chỉ thu thập được sẽ được huấn luyện trên mạng nơ-ron hồi tiếp LSTM với các đặc điểm như sau:

- **Lớp Masking:** Loại bỏ các bước thời gian bị padding.
- **Lớp LSTM 64 đơn vị:** Học đặc trưng cử chỉ từ chuỗi dữ liệu cảm biến.
- **Lớp Dense 32 đơn vị với ReLU:** Học biểu diễn dữ liệu trung gian.
- **Lớp Softmax đầu ra:** Phân loại cử chỉ thành 10 loại.

Mô hình được huấn luyện với thuật toán Adam, hàm mất mát categorical crossentropy, và dữ liệu được chuẩn hóa để tăng độ chính xác.



#### Thành quả lắp đặt

### VIII. KẾT QUẢ VÀ TRỰC QUAN HÓA

Sau khi thực hiện cử chỉ, hệ thống dự đoán nhãn tương ứng và hiển thị biểu đồ dữ liệu cảm biến:

```
172.16.18.25 -- [11/Mar/2025 08:00:27] "GET /predict?t=62001&ax=0.00&ay=0.00&az=0.00&gx=0.00&gy=0.00&gz=0.00&temp=0.00&marker=END HTTP/1.1" 200 -
172.16.18.25 -- [11/Mar/2025 08:00:28] "GET /predict?t=66544&ax=0.00&ay=0.00&az=0.00&gx=0.00&gy=0.00&gz=0.00&temp=0.00&marker=START HTTP/1.1" 200 -
172.16.18.25 -- [11/Mar/2025 08:00:28] "GET /predict?t=66643&ax=4.27&ay=0.52&az=9.90&gx=-0.30&gy=-0.20&gz=0.05&temp=27.42&marker=DATA HTTP/1.1" 200 -
172.16.18.25 -- [11/Mar/2025 08:00:29] "GET /predict?t=66763&ax=3.38&ay=0.47&az=9.49&gx=-0.09&gy=-0.26&gz=-0.25&temp=27.42&marker=DATA HTTP/1.1" 200 -
172.16.18.25 -- [11/Mar/2025 08:00:29] "GET /predict?t=66901&ax=5.83&ay=0.92&az=8.50&gx=-0.19&gy=0.28&gz=-0.54&temp=27.43&marker=DATA HTTP/1.1" 200 -
172.16.18.25 -- [11/Mar/2025 08:00:29] "GET /predict?t=67010&ax=6.13&ay=0.27&az=9.00&gx=-0.32&gy=0.08&gz=-1.37&temp=27.42&marker=DATA HTTP/1.1" 200 -
172.16.18.25 -- [11/Mar/2025 08:00:29] "GET /predict?t=67134&ax=3.98&ay=-0.13&az=9.37&gx=-0.50&gy=-0.05&gz=-1.91&temp=27.42&marker=DATA HTTP/1.1" 200 -
172.16.18.25 -- [11/Mar/2025 08:00:29] "GET /predict?t=67258&ax=2.60&ay=-0.57&az=8.87&gx=-0.58&gy=0.15&gz=-1.94&temp=27.42&marker=DATA HTTP/1.1" 200 -
172.16.18.25 -- [11/Mar/2025 08:00:29] "GET /predict?t=67384&ax=1.04&ay=-0.56&az=9.27&gx=-0.54&gy=0.37&gz=-1.64&temp=27.40&marker=DATA HTTP/1.1" 200 -
172.16.18.25 -- [11/Mar/2025 08:00:29] "GET /predict?t=67522&ax=1.22&ay=-0.75&az=9.20&gx=-0.51&gy=0.18&gz=-1.46&temp=27.43&marker=DATA HTTP/1.1" 200 -
172.16.18.25 -- [11/Mar/2025 08:00:29] "GET /predict?t=67647&ax=-1.36&ay=-0.89&az=10.44&gx=-0.67&gy=0.24&gz=-1.08&temp=27.43&marker=DATA HTTP/1.1" 200 -
172.16.18.25 -- [11/Mar/2025 08:00:30] "GET /predict?t=67776&ax=0.30&ay=-0.71&az=10.02&gx=-0.03&gy=0.13&gz=-0.11&temp=27.44&marker=DATA HTTP/1.1" 200 -
[1ml/1'[0m * [32m .....[1ml/1'[0m * [32m .....[0m * [37m'[0m * [1m0s'[0m 16m
s/step.....[0m * [37m'[0m * [1m0s'[0m 49ms/step
Predicted Label: 4
```

#### kết quả trực quan hóa

### IX. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Dự án đã chứng minh tính hiệu quả của việc sử dụng MPU6050 và ESP8266 để nhận diện cử chỉ tay trong thời

gian thực. Trong tương lai, hệ thống có thể mở rộng để:

- Hỗ trợ nhiều cử chỉ hơn với mô hình phức tạp hơn.
- Cải thiện độ chính xác bằng cách tối ưu tham số mô hình.
- Tích hợp vào các hệ thống điều khiển thông minh hoặc thiết bị hỗ trợ người khuyết tật.

#### X. TÀI LIỆU THAM KHẢO

- <https://www.tensorflow.org/>
- <https://flask.palletsprojects.com/>
- <https://docs.espressif.com/projects/esp8266-rtos-sdk/en/latest/>