



## THỰC TẬP CNTT 5: TRIỂN KHAI ỨNG DỤNG AI, IoT GIỚI THIỆU VỀ HỌC PHẦN

**ĐỀ TÀI: Hệ thống nhận diện cử chỉ tay thời gian thực để điều  
khiển trình chiếu PowerPoint**

Giảng viên hướng dẫn: ThS. Lê Trung Hiếu

Ks. Nguyễn Thái Khánh

Nhóm 5: Hoàng Công Sơn

Trịnh Minh Thành

Nguyễn Thị Kiều Hoa

Nguyễn Thế Khải

- 1. GIỚI THIỆU ĐỀ TÀI**
- 2. MỤC TIÊU NGHIÊN CỨU**
- 3. PHẠM VI NGHIÊN CỨU**
- 4. SƠ ĐỒ HỆ THỐNG**
- 5. QUY TRÌNH THỰC HIỆN**
- 6. CÁC CHỨC NĂNG CHÍNH CỦA HỆ THỐNG**
- 7. PHƯƠNG PHÁP THỰC HIỆN**
- 8. KẾT QUẢ ĐẠT ĐƯỢC**
- 9. KHÓ KHĂN VÀ HƯỚNG PHÁT TRIỂN**

## 1. Đặt vấn đề

Nhận diện cử chỉ tay đang trở thành một xu hướng quan trọng trong lĩnh vực **tương tác người-máy (HCI - Human-Computer Interaction)**. Đề tài này hướng đến việc xây dựng **một hệ thống nhận diện cử chỉ tay thời gian thực để điều khiển trình chiếu PowerPoint**, giúp người dùng chuyển đổi slide một cách thuận tiện mà không cần sử dụng chuột hoặc bàn phím.





# MỤC TIÊU NGHIÊN CỨU

---

## Mục tiêu chính của hệ thống:

- **Nhận diện cử chỉ tay** theo thời gian thực bằng camera thông thường.
- **Chuyển đổi dữ liệu hình ảnh bàn tay** thành đặc trưng có thể phân tích.
- **Dự đoán cử chỉ** và gửi lệnh tương ứng để điều khiển PowerPoint



# PHẠM VI NGHIÊN CỨU

Ứng dụng chính của hệ thống là **hỗ trợ thuyết trình thông minh**, đặc biệt trong **giáo dục, hội nghị và hỗ trợ người khuyết tật**. Tuy nhiên, hệ thống **chỉ nhận diện các cử chỉ đã được huấn luyện**, chưa hỗ trợ nhiều bàn tay cùng lúc và có thể bị ảnh hưởng trong điều kiện ánh sáng yếu.

Hệ thống sử dụng **MediaPipe Hands** để trích xuất tọa độ bàn tay, chuyển đổi dữ liệu sang **Trường góc Gramian (GAF)** và nhận diện cử chỉ bằng **GAFormer (Gesture Attention Transformer)**. Lệnh điều khiển PowerPoint được thực hiện thông qua **Pyautogui**.

Phạm vi nghiên cứu bao gồm **tám cử chỉ tay phổ biến** như vuốt trái/phải, mở tay, dừng, OK, v.v. Dữ liệu được thu thập từ camera RGB trong nhiều điều kiện ánh sáng và góc quay khác nhau, với tổng số **43.200 hình ảnh**.



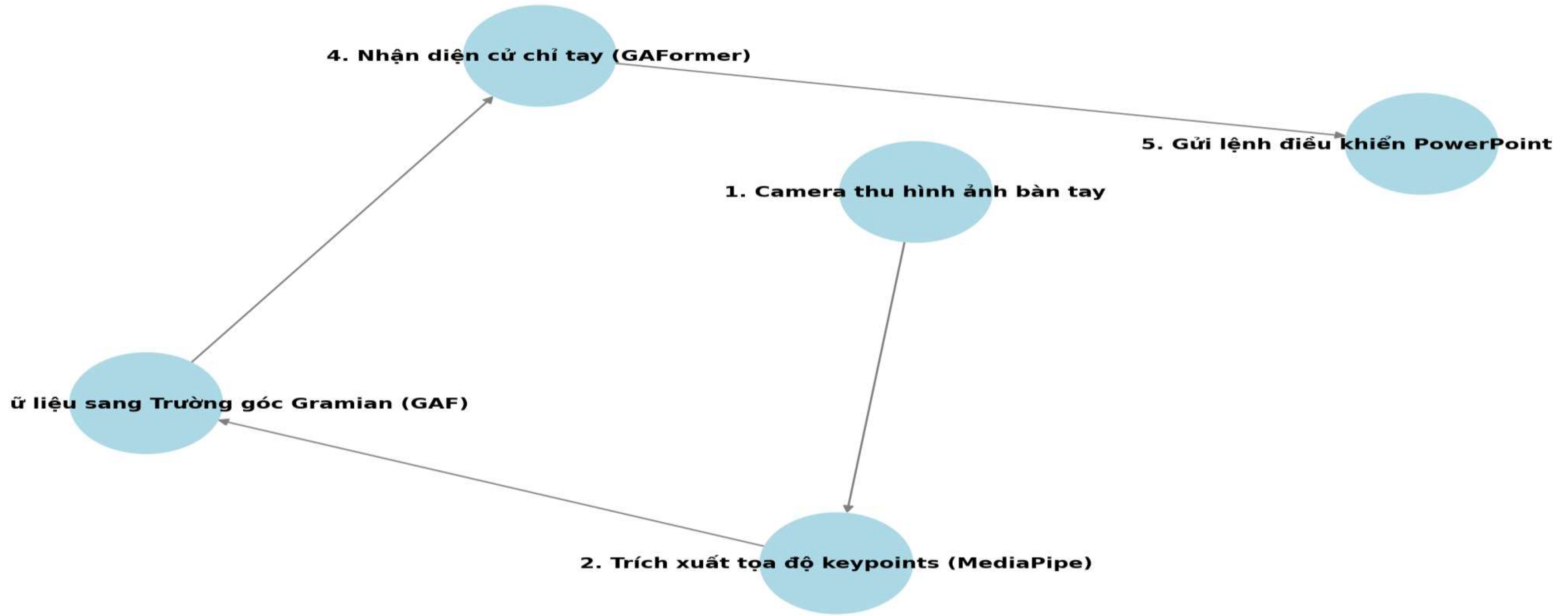
# SƠ ĐỒ HỆ THỐNG





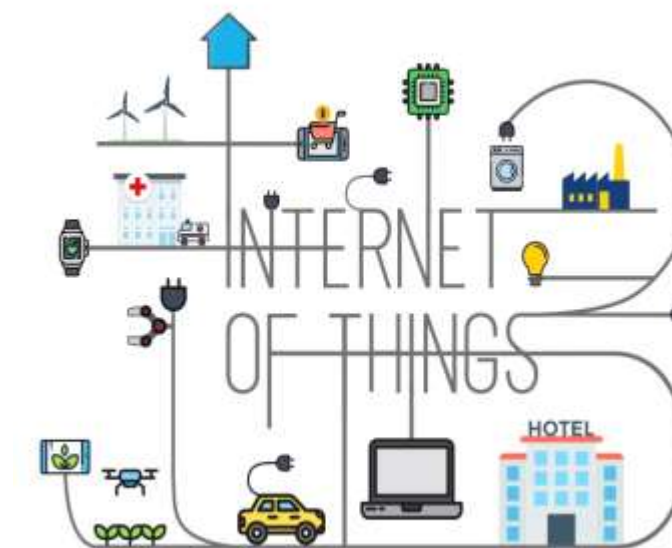
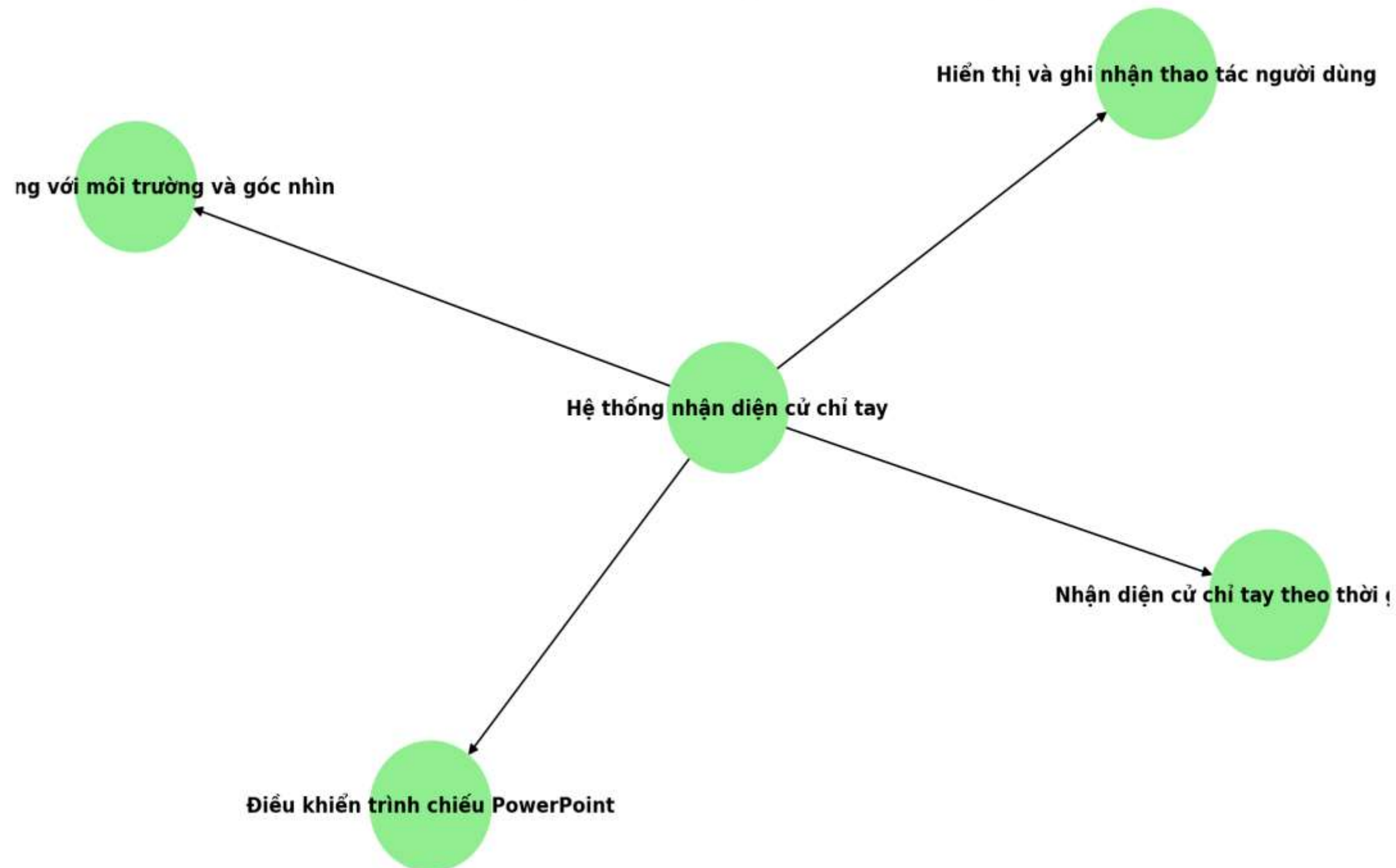
# QUY TRÌNH THỰC HIỆN

Quy trình thực hiện hệ thống nhận diện cử chỉ tay



# CHỨC NĂNG CHÍNH CỦA HỆ THỐNG

Chức năng chính của hệ thống nhận diện cử chỉ tay



Smart IoT Solutions Shaping  
the Business and Future



## Công nghệ sử dụng

### Phần cứng



Laptop chạy ứng dụng  
tạo QR

### MediaPipe Hands và OpenCV



**OpenCV**

### Ngôn ngữ lập trình



Python và các thư  
viện tạo giao diện

## Một số thư viện được sử dụng

- mediapipe
- OpenCV
- numpy
- pyautogui
- Tensorflow, keras
- Torch, transformers

# KẾT QUẢ ĐẠT ĐƯỢC

## 1. Độ chính xác nhận diện cử chỉ

**Độ chính xác tổng thể: 98.85%**

**Độ thu hồi (Recall): 97.40%**

**Thời gian phản hồi: < 0.5 giây**, đảm bảo nhận diện gần như thời gian thực.

## 2. Hiệu suất trong các điều kiện khác nhau

Hệ thống hoạt động ổn định ngay cả khi thay đổi ánh sáng (phòng sáng, ánh sáng yếu).

**Nhận diện chính xác cử chỉ tay từ nhiều góc quay ( $\pm 30^\circ$ ).**

**Khoảng cách nhận diện tốt nhất: 30 - 80 cm từ camera.**

## 3. Hiệu quả điều khiển PowerPoint

Hệ thống có thể thực hiện **chuyển slide, quay lại, bắt đầu/dừng trình chiếu** một cách chính xác.

**Tương tác mượt mà, không bị gián đoạn**, giúp người dùng thuyết trình chuyên nghiệp hơn.



# KHÓ KHĂN VÀ HƯỚNG PHÁT TRIỂN



## Khó khăn

- Ảnh hưởng của điều kiện môi trường
  - Giới hạn về phạm vi nhận diện
- Hệ thống hiện tại chỉ nhận diện **8 cử chỉ tay cố định**, chưa hỗ trợ cử chỉ tùy chỉnh theo người dùng.

Khoảng cách hiệu quả từ camera đến bàn tay nằm trong **30 - 80 cm**, xa hơn có thể làm giảm độ chính xác.

## Hạn chế trong nhận diện nhiều bàn tay

Hệ thống chỉ nhận diện **một bàn tay tại một thời điểm**, chưa hỗ trợ đa người dùng cùng lúc.

- Cải thiện độ chính xác và độ bền vững
- Hỗ trợ nhiều cử chỉ tay hơn
- Nhận diện nhiều bàn tay cùng lúc
- Tối ưu hóa hiệu suất trên thiết bị yếu
- Hỗ trợ nhiều nền tảng trình chiếu





*Thank You*