

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

Hoàng Thế Khải, Nguyễn Thị Kiều Hoa, Trịnh Minh Thành, Hoàng Công Sơn

Trường Đại học Đại Nam, Hà Nội, Việt Nam

Giới thiệu

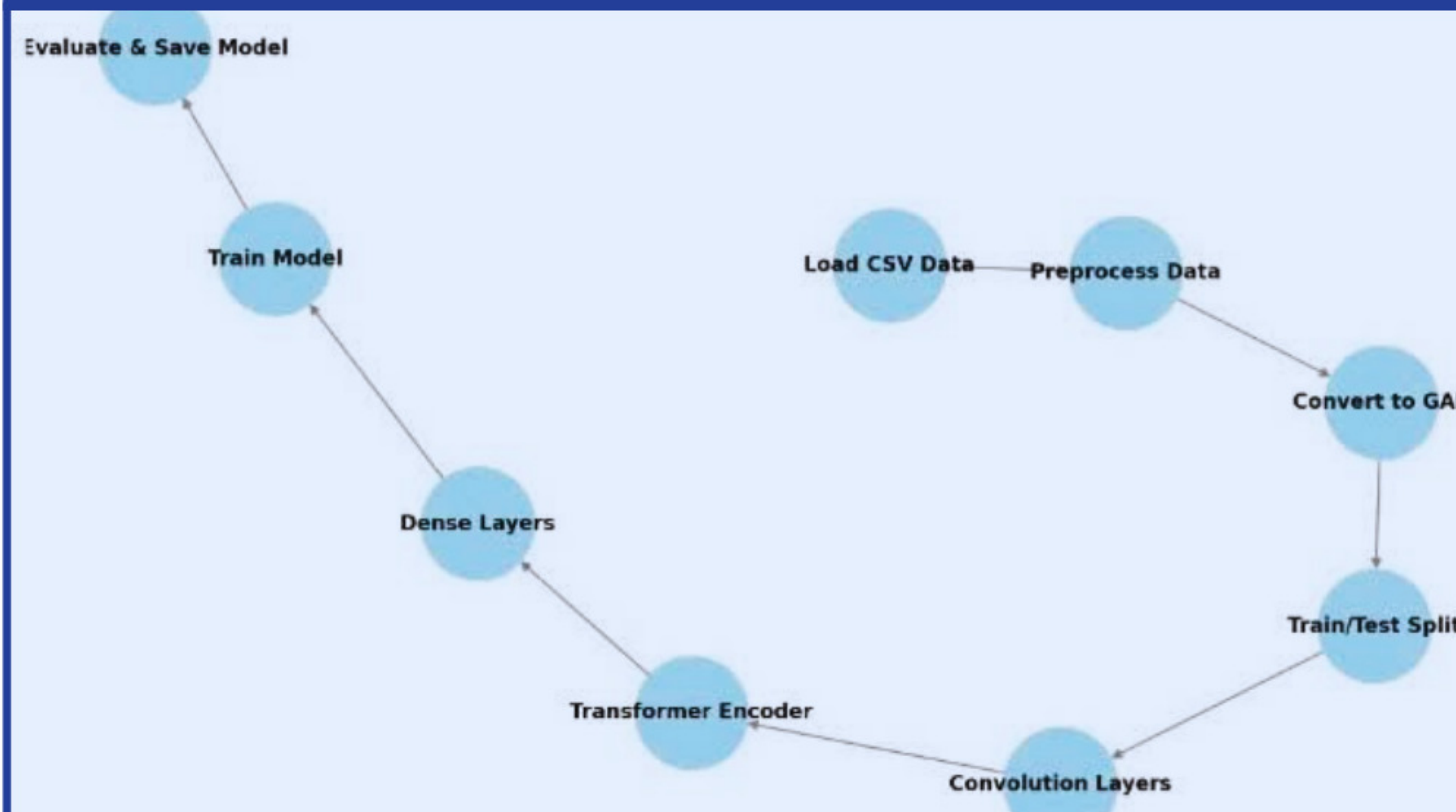
Nhận dạng cử chỉ tay động:

- Nhằm mục đích phân loại cử chỉ theo trình tự chuỗi thời gian nhất định
- Có tiềm năng ứng dụng lớn trong các lĩnh vực khác nhau: chăm sóc sức khỏe, điều khiển thiết bị gia đình, v.v.

Đóng góp:

- Triển khai GAFormer để chuyển đổi các điểm chính của bàn tay thành hình ảnh Trường góc Gramian (GAF)
- Sử dụng mô hình biến áp nhẹ để nhận dạng cử chỉ theo thời gian thực
- Nâng cao độ chính xác và khả năng ứng dụng thực tế của công nghệ nhận diện cử chỉ tay

SƠ ĐỒ HỆ THỐNG



Phương pháp đề xuất

Framework bao gồm bốn thành phần quan trọng bao gồm:

- Tiền xử lý dữ liệu thô (Làm mịn và Chuẩn hóa dữ liệu)
- Chuyển đổi Gramian Angular Difference Field (GADF)
- Trích xuất đặc trưng nhúng 2D (mạng nơ-ron tích chập CNNs)
- Tính toán Self-Attention (Học đặc trưng dựa trên Transformer)

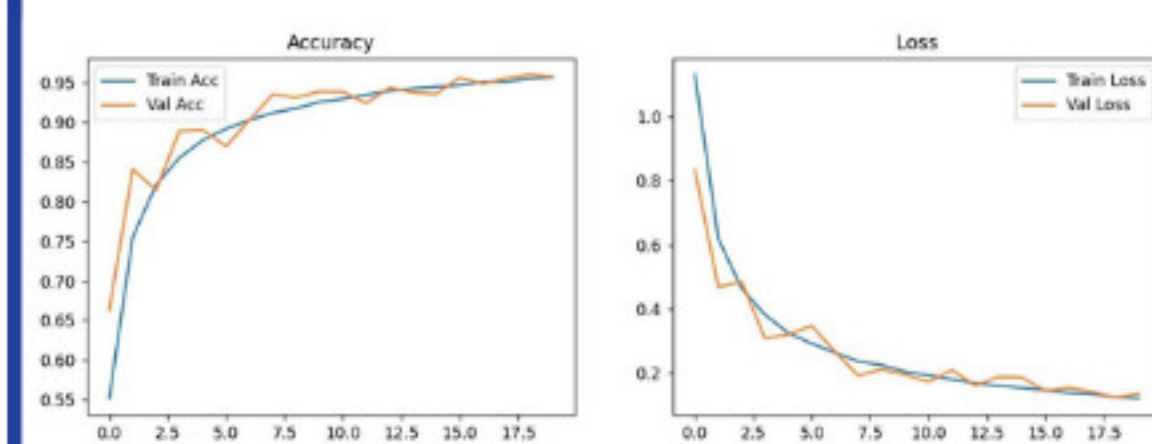
Dataset

Bộ dữ liệu:

- Bộ dữ liệu thu thập cử chỉ tay để điều khiển Google Maps thông qua nhận diện cử chỉ thời gian thực.
- Sử dụng tọa độ 3D của 21 điểm trên bàn tay được trích xuất từ MediaPipe.



Kết quả



Thuộc tính	Số lượng
Số lớp (nhân cử chỉ)	8
Số người thu dữ liệu	4
Số video mỗi nhân	80
Tổng số video	640
Tần số khung hình (FPS)	30
Thời lượng mỗi video	3 giây
Tổng số khung hình	~57,600

Kết luận và công việc tương lai

Kết luận:

- Mô hình GAFormer có khả năng học các mẫu không gian-thời gian của cử chỉ tay, đạt độ chính xác cao trên tập dữ liệu thu thập được.

Công việc trong tương lai:

- Kết hợp cử chỉ tay với lệnh giọng nói để mang lại trải nghiệm tương tác trực quan hơn. Sử dụng cảm biến IMU, camera độ sâu để tăng độ chính xác của nhận diện.