**Giới thiệu (Introduction)**

Bài báo giới thiệu **CSI-DeepNet**, một hệ thống nhận diện cử chỉ tay sử dụng **Wi-Fi Channel State Information (CSI)** và mạng nơ-ron tích chập sâu (**Deep CNN**).Hệ thống nhận diện cử chỉ đang là một lĩnh vực nghiên cứu phát triển nhanh chóng, với nhiều ứng dụng trong **nhà thông minh, chăm sóc sức khỏe, tối ưu hóa tài nguyên, an ninh và tiết kiệm năng lượng**. Các hệ thống nhận diện cử chỉ truyền thống sử dụng:

* **Cảm biến hình ảnh (Image Sensors)**
* **Cảm biến hồng ngoại (Infrared Sensors)**
* **Cảm biến siêu âm (Ultrasonic Sensors)**
* **Cảm biến đeo trên người (Wearable Sensors)**
* **Nhận diện dựa trên tần số vô tuyến RFID**
* **Hệ thống radar**

Tuy nhiên, các hệ thống này có một số nhược điểm lớn:

1. **Cảm biến hình ảnh**:
   * Gây lo ngại về quyền riêng tư.
   * Phụ thuộc vào **tầm nhìn thẳng (line-of-sight - LOS)**, điều kiện ánh sáng và góc nhìn.
2. **Cảm biến đeo trên người**:
   * Yêu cầu người dùng **luôn phải đeo thiết bị**, gây bất tiện.
3. **Cảm biến siêu âm và radar**:
   * **Chi phí cao**, **phạm vi hoạt động hạn chế**.

**DATA COLLECTION**

**-Thiết bị**: ESP-32 SoC được sử dụng làm bộ thu/phát Wi-Fi.

**-Dữ liệu**: Thu thập 20 cử chỉ tay alphanumeric (10 chữ cái: A, B, C, D, E, F, G, J, O, P và 10 chữ số: 0-9). 10 người tham gia, mỗi người thực hiện nhiều lần, tổng cộng 1.800 lần thử nghiệm.

**-Quy trình**: Dữ liệu được ghi lại ở khoảng cách **1m, 1.2m, và 2m** giữa bộ phát và bộ thu Wi-Fi. Sử dụng **bộ lọc Butterworth và Gaussian smoothing** để loại bỏ nhiễu. Noise Reduction Giam tieng on.

**RESULTS trang 11**

**-Độ chính xác trung bình**: **96.31%** gồm 21 lớp (20 cử chỉ + trạng thái đứng yên).

**-F1-score: 0.97, Cohen’s Kappa Score: 0.96.**

-So sánh với các mô hình khác: (cuối 12)

* ResNet-50: 70.38%
* DenseNet-121: 69.13%
* E2EDLF: 84.30%
* CSI-IANet: 90.50%
* CSI-DeepNet đạt hiệu suất cao nhất với ít tham số hơn (119,273 trainable parameters so với 516,321 của CSI-IANet).
* **Thời gian huấn luyện và nhận diện**:
  + CSI-DeepNet có thời gian huấn luyện ngắn hơn (**966 giây**) và thời gian nhận diện nhanh nhất (**18µs**) so với các mô hình khác.

**5. Hướng phát triển (Trang 12)**

* **Mở rộng tập dữ liệu**: Thêm nhiều kiểu cử chỉ khác nhau.
* **Xử lý môi trường phức tạp**: Kiểm tra hiệu suất khi có nhiễu từ môi trường thực tế.
* **Áp dụng vào hệ thống IoT thực tế**: Phát triển ứng dụng chạy trên các thiết bị có tài nguyên hạn chế.
* **Học bán giám sát**: Giảm bớt công việc gán nhãn dữ liệu bằng phương pháp học không cần giám sát.

**6. Khoảng trống (GAP) (Tổng hợp từ nội dung bài báo - Trang 12-13)**

1. **Chi tiết xử lý dữ liệu**: Bài báo đề cập đến xử lý dữ liệu nhưng không cung cấp đủ chi tiết về các thuật toán sử dụng.
2. **Hiệu suất trong môi trường nhiễu**: Không kiểm tra khả năng hoạt động của hệ thống trong môi trường Wi-Fi thực tế với nhiều nhiễu.