

ĐẠI HỌC ĐÀ NẴNG  
TRƯỜNG ĐẠI HỌC BÁCH KHOA  
KHOA ĐIỆN TỬ – VIỄN THÔNG

---

**BÁO CÁO**  
**PBL4: TRÍ TUỆ NHÂN TẠO**

**ĐỀ TÀI:**  
**DỰ ĐOÁN Ý ĐỊNH ĐIỀU KHIỂN THIẾT BỊ CHO**  
**BỆNH NHÂN HẠN CHẾ VẬN ĐỘNG DỰA TRÊN**  
**GIAO TIẾP BẰNG MẮT**

**Sinh viên thực hiện:**  
Trần Anh Toàn – 106220237 – 22KTMT1

**Giảng viên hướng dẫn:**  
TS. Trần Thị Minh Hạnh

Đà Nẵng, 2026

# 1 Giới thiệu

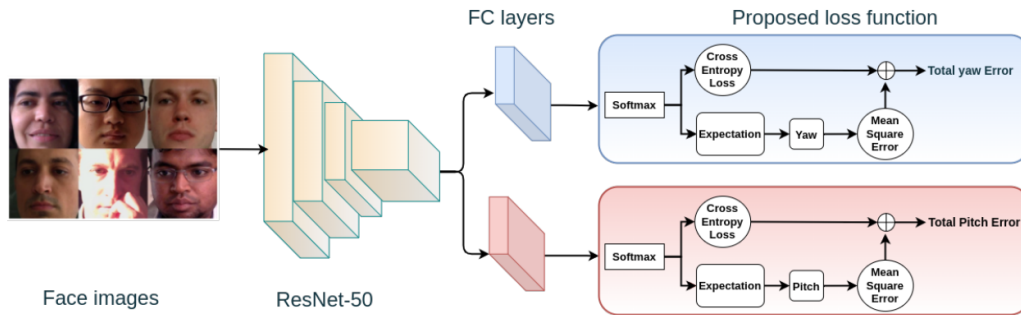
Hướng nhìn của mắt là một tín hiệu quan trọng, được khai thác trong nhiều lĩnh vực ứng dụng khác nhau. Thông tin này phản ánh mức độ tập trung và sự tương tác của người dùng trong các hệ thống tương tác người – robot cũng như các hệ thống hội thoại mở. Bên cạnh đó, trong lĩnh vực thực tế tăng cường, hướng nhìn của người dùng được sử dụng để dự đoán vùng chú ý, từ đó nâng cao khả năng nhận thức của thiết bị và giảm mức tiêu thụ năng lượng. Chính vì những lợi ích này, nhiều nghiên cứu đã được đề xuất nhằm ước lượng chính xác hướng nhìn của con người. Các phương pháp hiện có có thể được phân thành hai nhóm chính: phương pháp dựa trên mô hình và phương pháp dựa trên đặc trưng hình ảnh. Trong đó, các phương pháp dựa trên mô hình thường yêu cầu phần cứng chuyên dụng, khiến chúng khó triển khai trong các môi trường không bị ràng buộc. Ngược lại, các phương pháp dựa trên đặc trưng hình ảnh ước lượng trực tiếp hướng nhìn từ hình ảnh thu được bởi các camera thông dụng, nhờ đó dễ dàng áp dụng trong nhiều bối cảnh và điều kiện khác nhau.

Dựa vào những yếu tố trên, dự án này sẽ sử dụng phương pháp dựa trên đặc trưng hình ảnh để đưa ra các dự đoán về hướng nhìn của mắt nhằm nhận biết được ý định điều khiển thiết bị cho bệnh nhân hạn chế vận động. Mô hình sẽ được thực thi trên phần cứng Raspberry Pi (hoặc Jetson Nano) để ứng dụng cho xe lăn.

## 2 Phương pháp

Hiện tại có 2 phương pháp đang được xem xét:

### 2.1 L2CS-NET



Hình 1: Kiến trúc của phương pháp L2CS-NET

Hình 1 là kiến trúc của phương pháp L2CS-NET với phần chính là lớp trích xuất đặc trưng ResNet-50, lớp Fully-Connected và cuối cùng là hàm mất mát được đề xuất

Ảnh đầu vào được ResNet-50 thực hiện trích xuất đặc trưng hướng nhìn theo không gian. Tiếp theo đặc trưng sẽ được chia thành hai lớp Fully-Connected, và cũng áp dụng hai hàm mất mát khác nhau để tính góc của hướng nhìn theo phương ngang và phương dọc. Với mỗi đầu ra của lớp Fully-Connected, sử dụng softmax để chuyển từ mạng đầu ra của nơ-ron thành phân bố xác suất. Sau đó, hàm mất mát cross-entropy được áp dụng để tính toán mất mát phân loại nhị phân giữa xác suất đầu ra và nhãn nhị phân mong đợi.

Mất mát cross-entropy được tính như sau:

$$H(\mathbf{y}, \mathbf{p}) = - \sum_i y_i \log p_i \quad (1)$$

Tiếp theo, tính toán độ mong đợi của phân bố xác suất để dự đoán được hướng nhìn chi tiết. Cuối cùng tính sai số bình phương trung bình cho dự đoán và cộng nó với mất mát phân loại. Sai số bình phương trung bình được tính như sau:

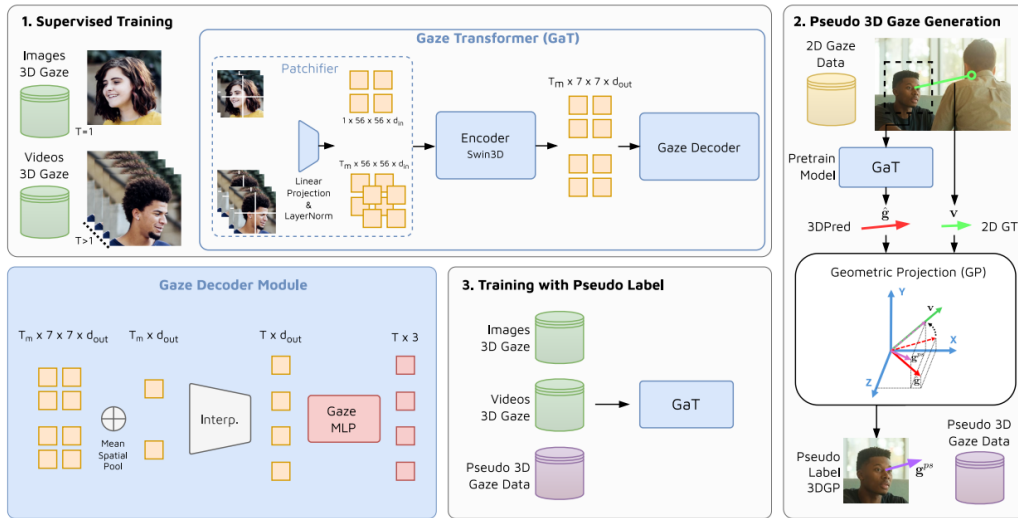
$$\text{MSE}(\mathbf{y}, \mathbf{p}) = \frac{1}{N} \sum_0^N (y - p)^2 \quad (2)$$

Hàm mất mát được đề xuất cho mỗi góc hướng mắt là một hàm kết hợp tuyến tính của mất mát cross-entropy và sai số bình phương trung bình, được tính như sau:

$$\text{CLS}(\mathbf{y}, \mathbf{p}) = H(\mathbf{y}, \mathbf{p}) + \beta \cdot \text{MSE}(\mathbf{y}, \mathbf{p}) \quad (3)$$

Trong đó, CLS là hàm mất mát kết hợp,  $\mathbf{p}$  là giá trị dự đoán,  $\mathbf{y}$  là giá trị nhãn thực, và  $\beta$  là hệ số hồi quy.

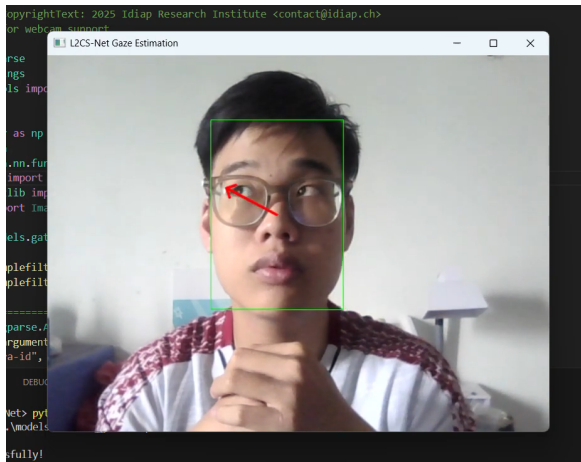
## 2.2 ST-WSGE



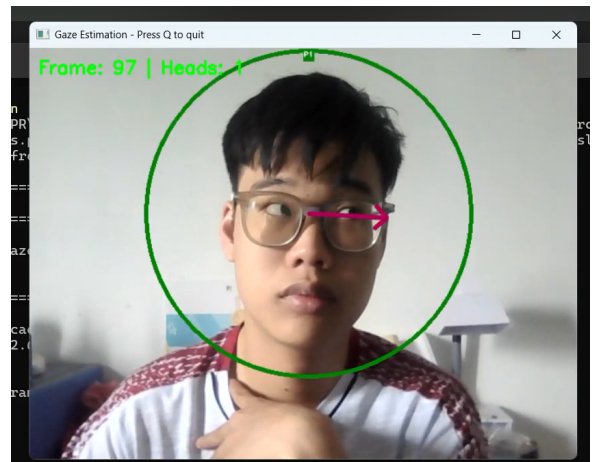
Hình 2: Kiến trúc của phương pháp ST-WSGE

## 3 Kết quả thực hiện (đến thời điểm hiện tại)

Tính đến thời điểm hiện tại, đã thu thập được tập dữ liệu thô Gaze360 và thực thi mô hình pretrained của các phương pháp trên máy tính và thu được kết quả khá chính xác. Tập dữ liệu thô Gaze360 cung cấp rất nhiều chú thích hướng nhìn trong khoảng 360 độ. Bao gồm 238 đối tượng trải dài nhiều độ tuổi, giới tính và sắc tộc, có thể xem chi tiết ở liên kết dưới phần phụ lục. Hình 3 và Hình 4 thể hiện kết quả thử nghiệm cho mô hình của hai phương pháp được đề xuất.



Hình 3: Kết quả của phương pháp L2CS-NET



Hình 4: Kết quả của phương pháp ST-WSGE

Để xem chi tiết kết quả có thể truy cập vào liên kết dưới phần phụ lục

## **A Phụ lục**

### **A.1 Tập dữ liệu**

Để xem chi tiết tập dữ liệu có thể truy cập vào liên kết sau:

<https://drive.google.com/drive/folders/1eUz45L3aqO9PD9rvhOTsmGzdH6Nb7Jrh>

### **A.2 Kết quả chạy mô hình pretrained**

Để xem chi tiết kết quả có thể truy cập vào liên kết sau:

<http://youtube.com/watch?v=2bAZR3xpuA8>