

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA CÔNG NGHỆ THÔNG TIN

BÁO CÁO HƯỚNG DẪN CHẠY CHƯƠNG TRÌNH

Môn: Thị giác máy tính

Lớp: 22TGMT

Sinh viên:

Đinh Nguyễn Gia Bảo

(22127027)

Hoàng Bảo Khanh (22127183)

Giáo viên hướng dẫn:

TS. Võ Hoài Việt

ThS. Phạm Minh Hoàng



Mục lục

1	Chức năng huấn luyện và dự đoán cả bộ dữ liệu	2
2	Chức năng dự đoán một video bất kỳ (Inference)	3
3	Video demo hướng dẫn chạy mã nguồn	4

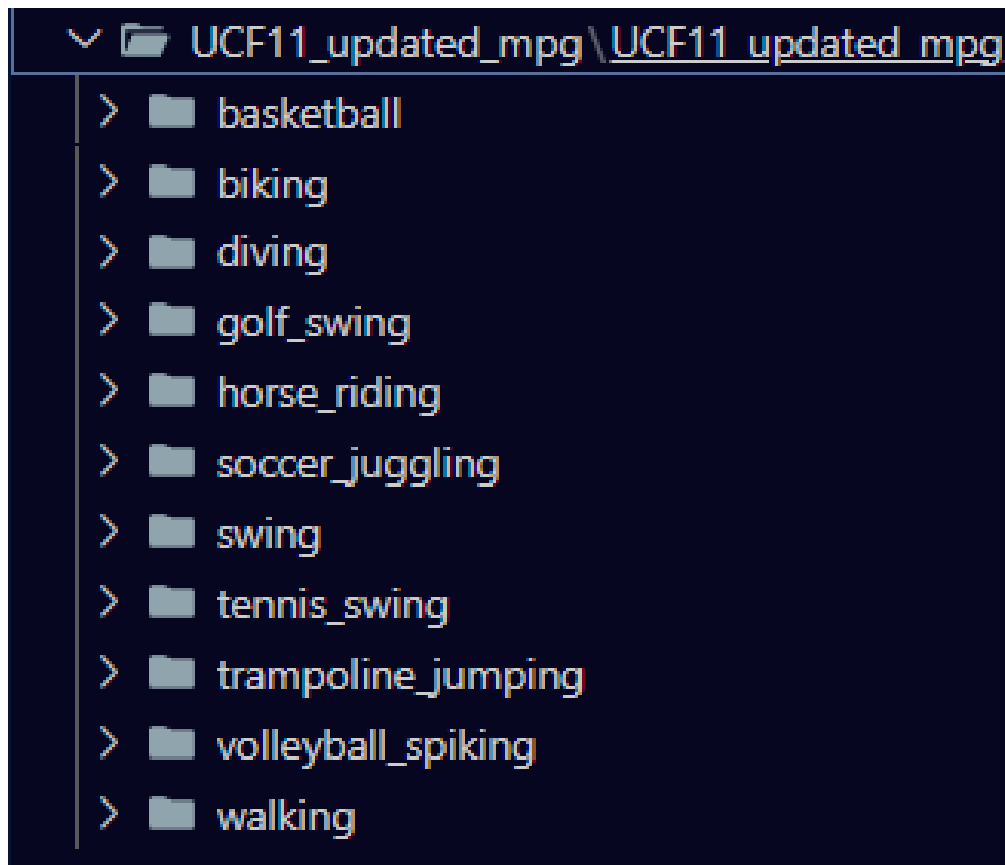
1 Chức năng huấn luyện và dự đoán cả bộ dữ liệu

Để thực hiện huấn luyện mô hình, chúng ta sẽ chạy file `main.py`. Để chạy chương trình, trước tiên ta phải di chuyển địa chỉ đến nơi chứa file `main.py`, sau đó gõ lệnh command line theo cú pháp như sau:

```
python main.py <type-of-dataset> <dataset-path>
```

Trong đó, *type-of-dataset* là loại dữ liệu muốn huấn luyện. Trong chương trình sẽ cho phép người dùng huấn luyện ba bộ dữ liệu chính là `ucf11`, `ucf50` và `ucf101`. Và đây cũng sẽ là tên thư mục được tạo trong benchmark sau khi thực hiện đánh giá mô hình.

dataset-path là địa chỉ dẫn đến bộ dataset. Ở đây chúng ta nên nhập đầy đủ địa chỉ để tránh trường hợp chương trình không tìm thấy dataset. Ở đây chúng ta sẽ nhập địa chỉ sao cho đến thư mục chứa các thư mục nhân của dataset.



Hình 1: Cấu trúc thư mục của dataset UCF11

Dựa vào cấu trúc thư mục trên, chúng ta sẽ nhập địa chỉ cho đến

```
../UCF11_updated_mpg/UCF11_updated_mpg.
```

Dưới đây là hình ảnh minh họa cách chạy chương trình huấn luyện. Ở đây ta sẽ ví dụ với bộ dataset ucf11.

```
PS E:\Computer_Vision\Project\src> python main.py ucf11 E:\Computer_Vision\Project\UCF11_updated_mpg\UCF11_updated_mpg
E:\Computer_Vision\Project\UCF11_updated_mpg\UCF11_updated_mpg\basketball
E:\Computer_Vision\Project\UCF11_updated_mpg\UCF11_updated_mpg\biking
E:\Computer_Vision\Project\UCF11_updated_mpg\UCF11_updated_mpg\diving
E:\Computer_Vision\Project\UCF11_updated_mpg\UCF11_updated_mpg\golf_swing
E:\Computer_Vision\Project\UCF11_updated_mpg\UCF11_updated_mpg\horse_riding
E:\Computer_Vision\Project\UCF11_updated_mpg\UCF11_updated_mpg\soccer_juggling
E:\Computer_Vision\Project\UCF11_updated_mpg\UCF11_updated_mpg\swing
E:\Computer_Vision\Project\UCF11_updated_mpg\UCF11_updated_mpg\tennis_swing
E:\Computer_Vision\Project\UCF11_updated_mpg\UCF11_updated_mpg\trampoline_jumping
E:\Computer_Vision\Project\UCF11_updated_mpg\UCF11_updated_mpg\volleyball_spiking
E:\Computer_Vision\Project\UCF11_updated_mpg\UCF11_updated_mpg\walking
Extracting features using ResNet50...
```

Hình 2: Giao diện CMD khi thực hiện chạy huấn luyện chương trình

Về các công đoạn, chương trình sẽ trải qua các bước giống như những gì đã đề cập ở trong báo cáo. Sau khi chạy thì chương trình sẽ lưu lại model dưới dạng file .pt để sử dụng cho việc dự đoán video bất kỳ. Ngoài ra các thông số như thời gian trích xuất đặc trưng, thời gian huấn luyện, tỷ lệ chính xác và lỗi cũng như biểu đồ liên quan sẽ được lưu tất cả vào thư mục Benchmark với tên thư mục tương ứng với tên bộ dữ liệu. Các thông số này cũng sẽ được in ra màn hình CMD sau khi hoàn thành.

```
Epoch: [6/100] Train Loss: 0.0245, Train Accuracy: 99.90% Val Loss: 0.1739, Val Accuracy: 93.75%
Epoch: [7/100] Train Loss: 0.0159, Train Accuracy: 100.00% Val Loss: 0.1777, Val Accuracy: 94.14%
Epoch: [8/100] Train Loss: 0.0113, Train Accuracy: 100.00% Val Loss: 0.1836, Val Accuracy: 94.14%
Epoch: [9/100] Train Loss: 0.0090, Train Accuracy: 100.00% Val Loss: 0.1761, Val Accuracy: 94.92%
Epoch: [10/100] Train Loss: 0.0077, Train Accuracy: 100.00% Val Loss: 0.1796, Val Accuracy: 94.14%
Epoch: [11/100] Train Loss: 0.0063, Train Accuracy: 100.00% Val Loss: 0.1729, Val Accuracy: 94.92%
Epoch: [12/100] Train Loss: 0.0054, Train Accuracy: 100.00% Val Loss: 0.1764, Val Accuracy: 94.92%
Epoch: [13/100] Train Loss: 0.0047, Train Accuracy: 100.00% Val Loss: 0.1797, Val Accuracy: 94.92%
Epoch: [14/100] Train Loss: 0.0041, Train Accuracy: 100.00% Val Loss: 0.1799, Val Accuracy: 94.92%
Epoch: [15/100] Train Loss: 0.0035, Train Accuracy: 100.00% Val Loss: 0.1798, Val Accuracy: 94.92%
Early stopping triggered
Training completed in 12.87 seconds
Test Loss: 0.1631, Test Accuracy: 96.25%
PS E:\Computer_Vision\Project\src>
```

Hình 3: Kết quả sau khi chạy xong huấn luyện và dự đoán bộ UCF11

2 Chức năng dự đoán một video bất kỳ (Inference)

Sau khi huấn luyện mô hình thành công, người dùng có thể sử dụng mô hình để dự đoán hành động trong một video bằng câu lệnh sau:

```
python src/inference.py <đường_dẫn_video_test> <đường_dẫn_model>
```

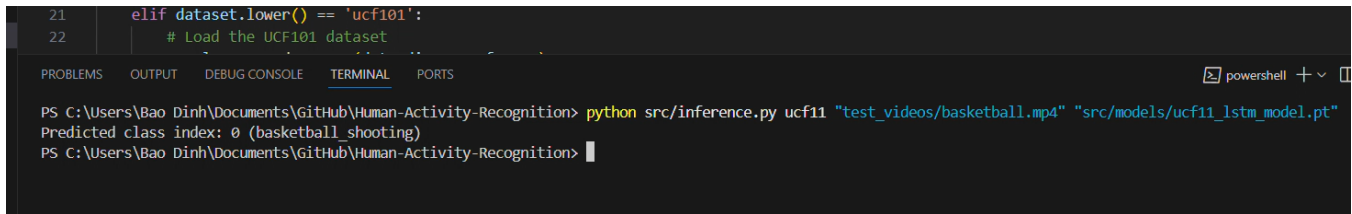
Trong đó:

- <đường_dẫn_video_test> là đường dẫn đến video cần được phân loại, ví dụ:

"test_videos/walking.mp4"

- <đường_dẫn_model> là đường dẫn đến tệp trọng số của mô hình đã huấn luyện, ví dụ:

"src/models/ucf101_lstm_model.pt"



```

21 elif dataset.lower() == 'ucf101':
22     # Load the UCF101 dataset
...
PS C:\Users\Bao Dinh\Documents\GitHub\Human-Activity-Recognition> python src/inference.py ucf11 "test_videos/basketball.mp4" "src/models/ucf11_lstm_model.pt"
Predicted class index: 0 (basketball_shooting)
PS C:\Users\Bao Dinh\Documents\GitHub\Human-Activity-Recognition>

```

Hình 4: ví dụ câu lệnh chạy inference UCF11

Sau khi thực hiện, hệ thống sẽ trả về nhãn hành động tương ứng mà mô hình dự đoán được từ nội dung video. Như trong hình thì ta thấy kết quả trả về label 0 (Basketball shooting) tức là trong video thì human đang chơi basketball.

3 Video demo hướng dẫn chạy mã nguồn

Để chi tiết hơn thì nhóm có quay một video hướng dẫn demo chương trình với nội dung gồm:

- Huấn luyện dữ liệu từ mã nguồn
- Chạy inference model train được

Thời lượng video kéo dài khoảng 9 phút. Video có thể xem và tải tại Google Drive:
<https://drive.google.com/drive/folders/1vwbIYs8KR0iCNaLyh3lEZLzyQU1vWH7q?>