How to Run the Multi-Modal Fusion Pose Estimation on Windows Using CPU or GPU

Introduction

This document provides step-by-step instructions on how to set up and run the multi-modal fusion technique for real-time pose estimation on Windows. The system combines 2D and 3D images to predict all joints in the human body. It supports both CPU and GPU for inference.

Prerequisites

- Windows OS
- Anaconda or Miniconda installed
- Git installed
- For GPU: NVIDIA GPU with CUDA support (the environment includes CUDA 11.6)

Steps to Set Up

Step 1: Clone the Repository

1. Open Anaconda Prompt.
2. Run:

```
git clone https://github.com/BadBoy0170/Pose_Estimation.git
cd Pose_Estimation
```

Step 2: Create and Activate the Environment

1. In Anaconda Prompt, create the environment:

```
conda env create -f environment.yml
```

2. Activate it:

```
conda activate posefusion
```

Step 3: Download Pre-trained Weights

1. Download `yolov7-w6-pose.pt` from this link.
2. Place it in the `Pose_Estimation` directory.

   Alternatively, if you have wget installed (or use PowerShell):

```
Invoke-WebRequest -Uri https://github.com/WongKinYiu/yolov7/releases/download/v0.1/yolo
```

Step 4: Run Inference

On CPU

```
python run_pose.py --source 0 --device cpu
```

On GPU Ensure your GPU drivers are up to date. Then:

```
python run_pose.py --source 0 --device 0
```

For video files, replace `0` with the path to the video.

Troubleshooting

- If GPU is not detected, check CUDA installation.
- Ensure the environment is activated before running commands.

Happy Coding!