Real-time Pose Estimation

Landmark detection with python for Unity Engine

20.07.2023

Andreas Lakus 11012049
Ahmed Hussein 11013671

Applied Artificial Intelligence
Dr. Simon Ziegler | Maximilian Schuhmacher

Project description

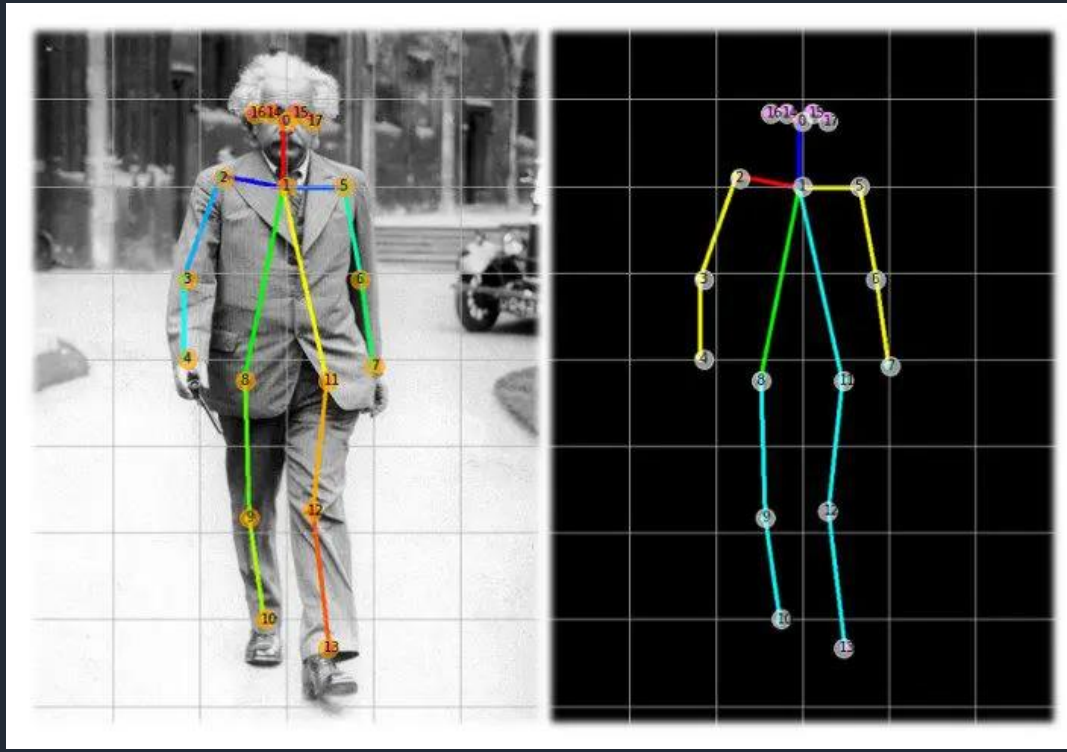
What?

Real-time pose estimation for Unity Engine with AI (ML)

Why?

Record animations

Control games with movements



MediaPipe

- Flexible Framework developed by Google for multimedia processing Pipelines
- Consists of:
  - Graph-based Pipeline -> define data flow yourself
  - Processing components -> pre-built components (modular units)
    - Feature-extractors, video decoders, machine learning components
  - Custom Components -> Develop needed components yourself
  - Connect different sources to MediaPipe
  - Synchronization and Threading
  - Optimizing and Hardware Acceleration

MediaPipe

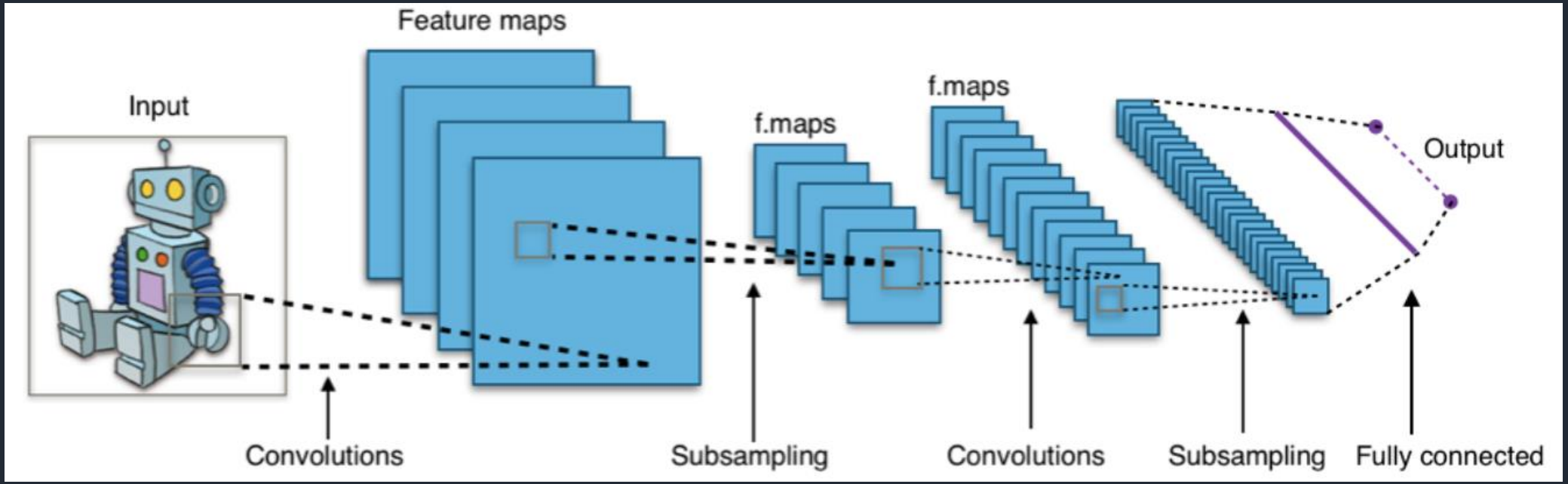
YOLOv7 vs MediaPipe Pose Features

Features	YOLOv7 Pose	MediaPipe Pose
Topology	17 Keypoints **COCO**	33 Keypoints **COCO** + **Blaze Palm** + **Blaze Face**
Workflow	Detection runs for all frames	Detection runs once followed by tracker until occlusion occurs
GPU support	Support for both CPU and GPU	Only CPU
Segmentation	Segmentation not integrated to pose directly	Segmentation integrated
Number of persons	Multi-person	Single person

MediaPipe

BlazePose

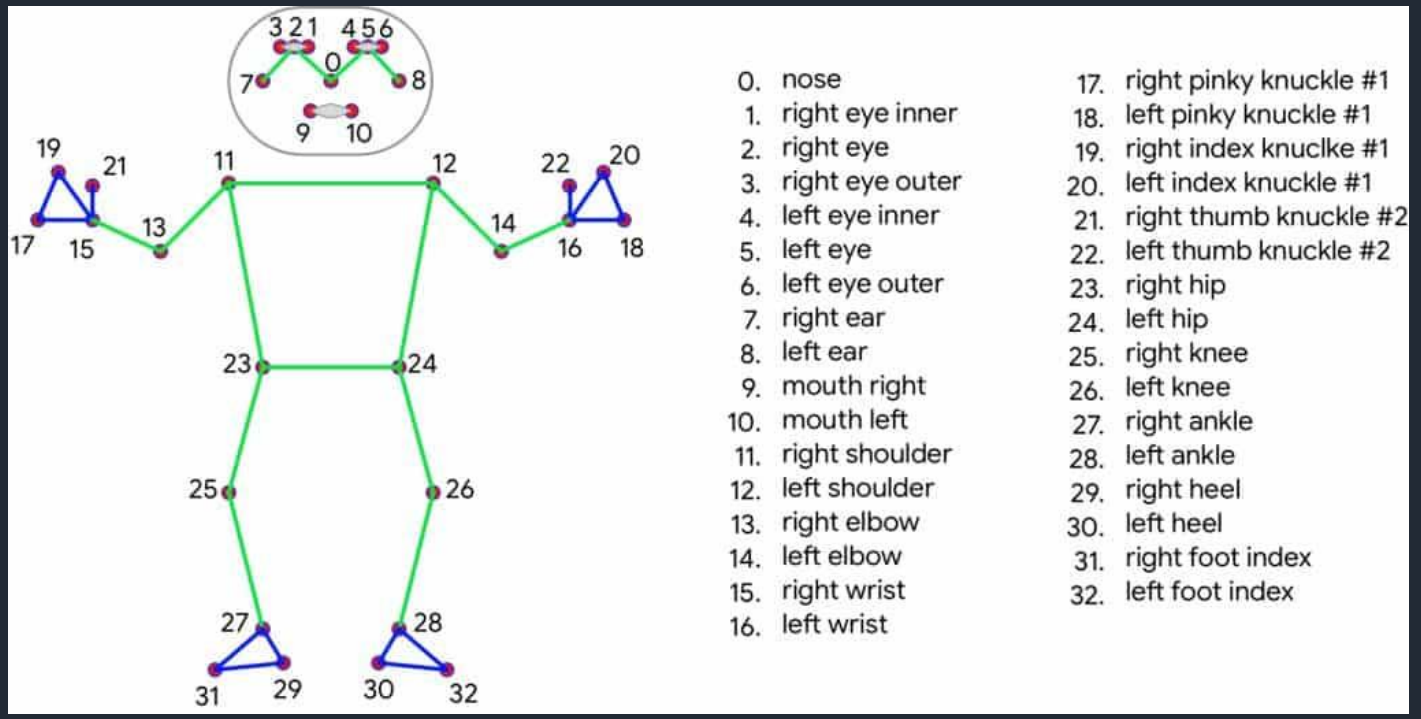
- Convolutional Neural Network-based model (CNN)
  - Convolutional Layer (Matrices with illumination values)
  - Pooling Layer (clean data)
  - Full-connected Layer (classification)



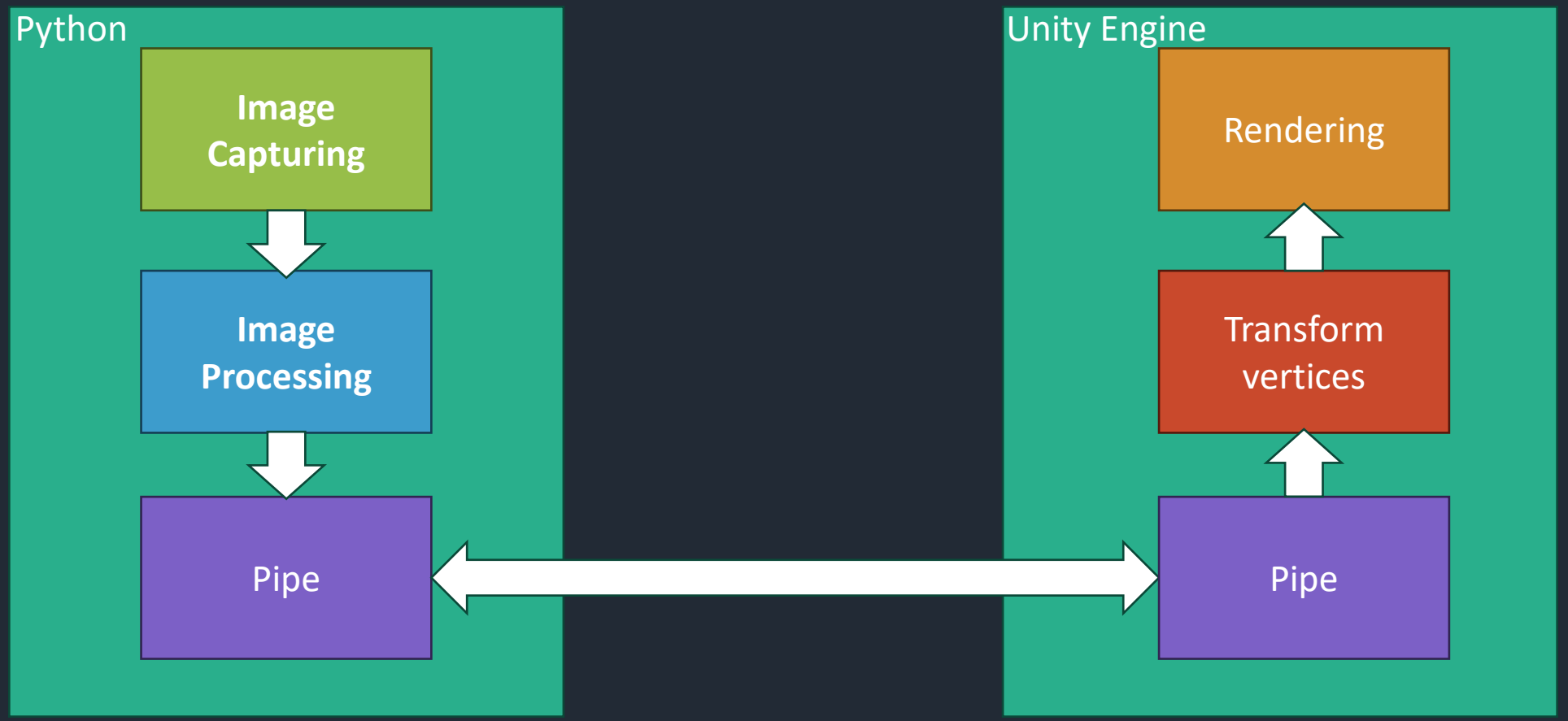
MediaPipe

BlazePose

- extracts 33 2D landmarks on the human body
- BlazePose is a lightweight machine learning architecture that achieves real-time performance on mobile phones and PCs with CPU



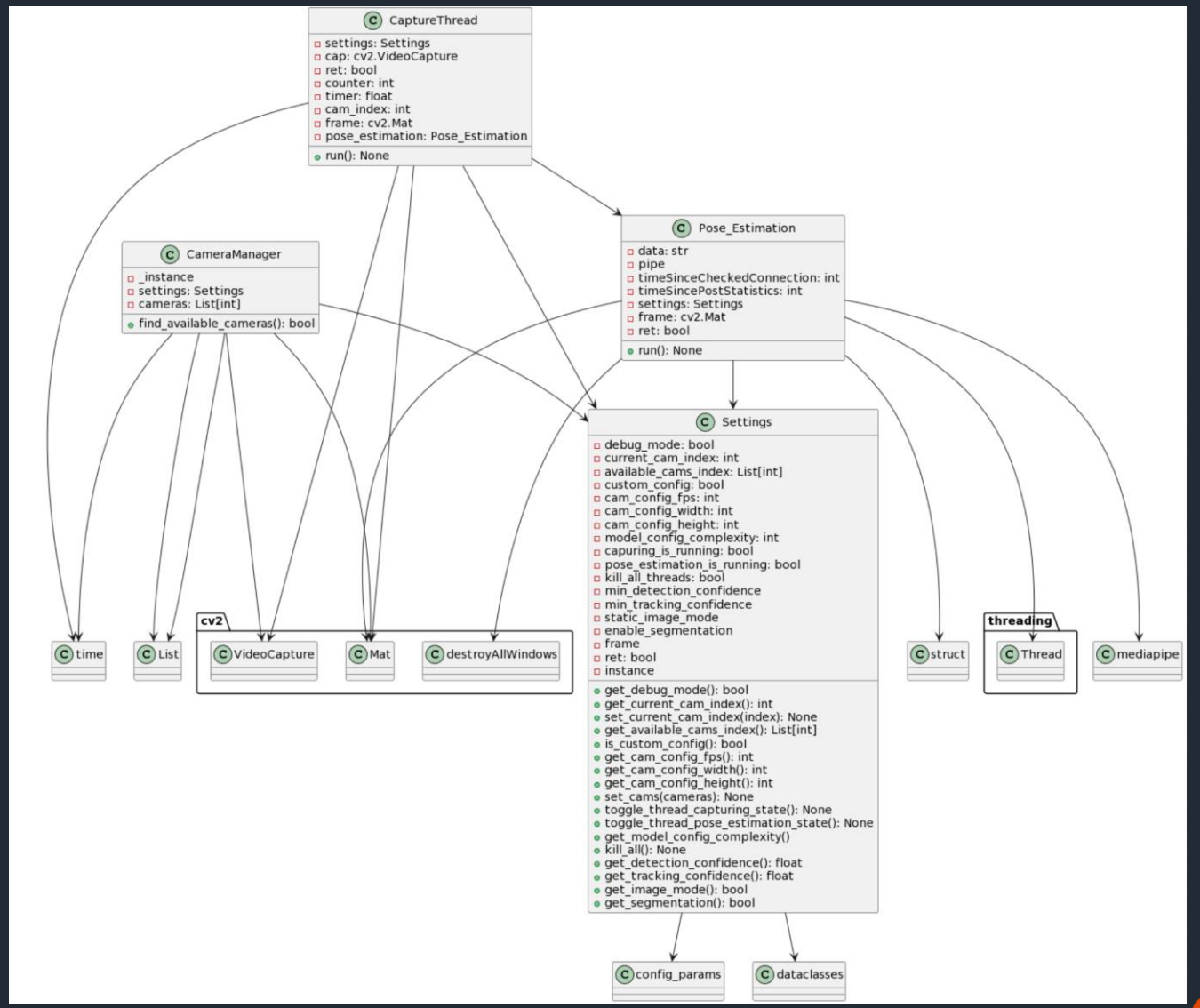
System Overview

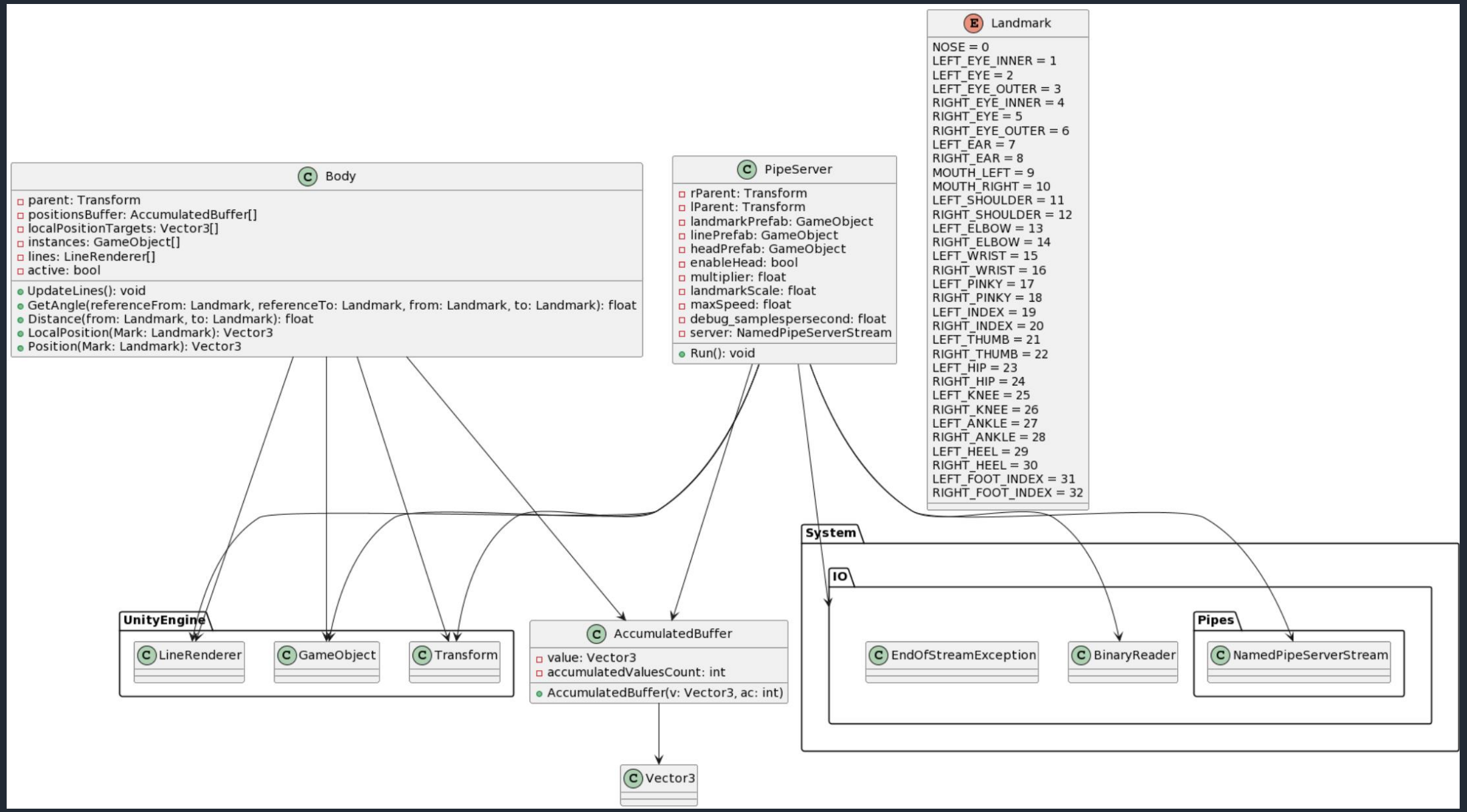


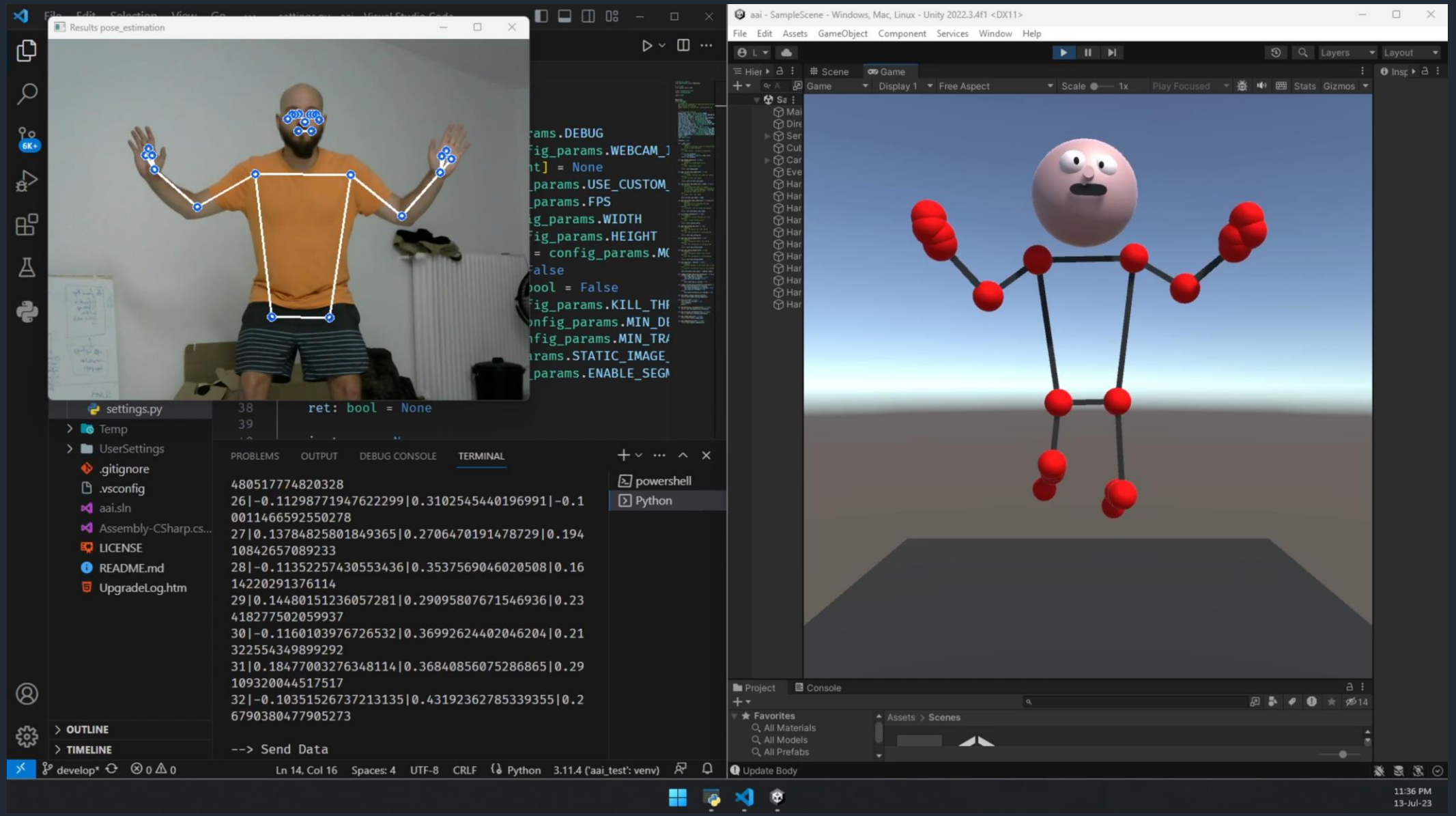
Architecture (py)

- Libs: opencv, mediapipe, time …

- Classes
  - CameraManager
    - Find and manage Cams
  - CaptureThread
    - Get frames from cam with opencv
  - Settings
    - Dataclass
  - Pose_Estimation (Thead)
    - Predict Pose




Body
parent: Transform
positionsBuffer: AccumulatedBuffer[]
localPositionTargets: Vector3[]
instances: GameObject[]
lines: LineRenderer[]
active: bool
UpdateLines(): void
GetAngle(referenceFrom: Landmark, referenceTo: Landmark, from: Landmark, to: Landmark): float
Distance(from: Landmark, to: Landmark): float
LocalPosition(Mark: Landmark): Vector3
Position(Mark: Landmark): Vector3
PipeServer
rParent: Transform
lParent: Transform
landmarkPrefab: GameObject
linePrefab: GameObject
headPrefab: GameObject
enableHead: bool
multiplier: float
landmarkScale: float
maxSpeed: float
debug_samplespersecond: float
server: NamedPipeServerStream
Run(): void
Landmark
NOSE = 0
LEFT_EYE_INNER = 1
LEFT_EYE = 2
LEFT_EYE_OUTER = 3
RIGHT_EYE_INNER = 4
RIGHT_EYE = 5
RIGHT_EYE_OUTER = 6
LEFT_EAR = 7
RIGHT_EAR = 8
MOUTH_LEFT = 9
MOUTH_RIGHT = 10
LEFT_SHOULDER = 11
RIGHT_SHOULDER = 12
LEFT_ELBOW = 13
RIGHT_ELBOW = 14
LEFT_WRIST = 15
RIGHT_WRIST = 16
LEFT_PINKY = 17
RIGHT_PINKY = 18
LEFT_INDEX = 19
RIGHT_INDEX = 20
LEFT_THUMB = 21
RIGHT_THUMB = 22
LEFT_HIP = 23
RIGHT_HIP = 24
LEFT_KNEE = 25
RIGHT_KNEE = 26
LEFT_ANKLE = 27
RIGHT_ANKLE = 28
LEFT_HEEL = 29
RIGHT_HEEL = 30
LEFT_FOOT_INDEX = 31
RIGHT_FOOT_INDEX = 32
UnityEngine
LineRenderer
GameObject
Transform
AccumulatedBuffer
value: Vector3
accumulatedValuesCount: int
AccumulatedBuffer(v: Vector3, ac: int)
Vector3
System
IO
EndOfStreamException
BinaryReader
Pipes
NamedPipeServerStream


Results pose_estimation
settings.py
Temp
UserSettings
.gitignore
.vsconfig
aai.sln
Assembly-CSharp.cs...
LICENSE
README.md
UpgradeLog.htm
PROBLEMS
OUTPUT
DEBUG CONSOLE
TERMINAL
powershell
Python
480517774820328
26|-0.11298771947622299|0.3102545440196991|-0.1
0011466592550278
27|0.13784825801849365|0.2706470191478729|0.194
10842657089233
28|-0.11352257430553436|0.3537569046020508|0.16
14220291376114
29|0.14480151236057281|0.29095807671546936|0.23
418277502059937
30|-0.1160103976726532|0.36992624402046204|0.21
322554349899292
31|0.18477003276348114|0.36840856075286865|0.29
109320044517517
32|-0.10351526737213135|0.43192362785339355|0.2
6790380477905273
--> Send Data
OUTLINE
TIMELINE
develop*
Ln 14, Col 16
Spaces: 4
UTF-8
CRLF
Python
3.11.4 ('aai_test': venv)
aai - SampleScene - Windows, Mac, Linux - Unity 2022.3.4f1 <DX11>
File Edit Assets GameObject Component Services Window Help
Scene
Game
Display 1
Free Aspect
Scale
Play Focused
Stats
Gizmos
Project
Console
Favorites
All Materials
All Models
All Prefabs
Assets > Scenes
Update Body
11:36 PM
13-Jul-23

References

- https://developers.google.com/mediapipe/solutions/vision/pose_landmarker
- https://learnopencv.com/yolov7-pose-vs-mediapipe-in-human-pose-estimation/
- Kim, J.-W.; Choi, J.-Y.; Ha, E.-J.; Choi, J.-H. Human Pose Estimation Using MediaPipe Pose and Optimization Method Based on a Humanoid Model. *Appl. Sci.* **2023**, *13*, 2700. https://doi.org/10.3390/app13042700
- https://de.wikipedia.org/wiki/Convolutional_Neural_Network
- https://mobidev.biz/blog/human-pose-estimation-technology-guide