A Camera-Based Platform for Patient Ambulation Assessment

Mostafa Habibi, Mehrdad Nourani
Predictive Analytics and Technologies Laboratory
The University of Texas at Dallas, Richardson, Texas, USA

Dennis H. Sullivan
Geriatric Research, Education, and Clinical Center
Central Arkansas Veterans Healthcare System, North Little Rock, Arkansas, USA

THE UNIVERSITY OF TEXAS AT DALLAS
Erik Jonsson School of Engineering and Computer Science
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Research Question (S)

The goal of this study is to develop an ambulation assessment platform in healthcare that employs a deep neural network for body/head detection but further recognizes the position, postures and motion of persons in a video stream.

research question:

- How to calculate the distance of a person from a basic camera (e.g., webcam) and track the same person in video streaming using image processing and deep learning methods?
- How to estimate the 3D position of body landmarks and posture of a person using image processing and deep learning methods?

We present:
- Person & Head Detection and Tracking
- Head Coordinates and Ambulation Measurement
- Body Landmark Estimation and Posture Analysis.

Introduction

Background and Motivation:
- Older patients in hospitals spend a majority of their time in bed, leading to negative outcomes such as falls, blood clots, infections, bowel issues, and pressure ulcers.
- Prolonged bed rest increases hospital stays, burdens caregivers, raises the risk of nursing home placement, and can have fatal consequences.
- Promoting early and progressive ambulation (EPA) in older patients can reduce the risk of functional decline and improve clinical outcomes.
- Current technologies for monitoring ambulation in hospitals are expensive, have low adoption rates, and are often inaccurate, highlighting the need for a low-cost technology to accurately measure patient ambulation.

AI-Driven Image Processing :
- YOLO v8, a cutting-edge object recognition technique in computer vision, utilizes a convolutional neural network (CNN) to identify objects in photos and images. It incorporates an EfficientNet backbone for feature extraction, a neck module for combining features, and a head module for predicting bounding boxes and class probabilities.
- MediaPipe Pose is a machine learning tool that tracks the entire body by predicting the positions of 33 3D landmarks in images or videos, allowing for assessment and correction of human poses. It is based on the BlazePose algorithm, trained to handle heavily occluded scenarios, and provides raw data that requires further processing for classification.

Hardware Requirements:
- Single simple camera

Methods

The proposed algorithm for calculating activity tracking based on head detection consists of three main steps:

I. Person & Head Detection and Tracking:

- At first need to detect person and head, The YOLO object detection model was trained using the CrowdHuman data set.
- We used BoT-SORT method to track patient heads in a video stream.

II. Head Coordinates and Ambulation Measurement:

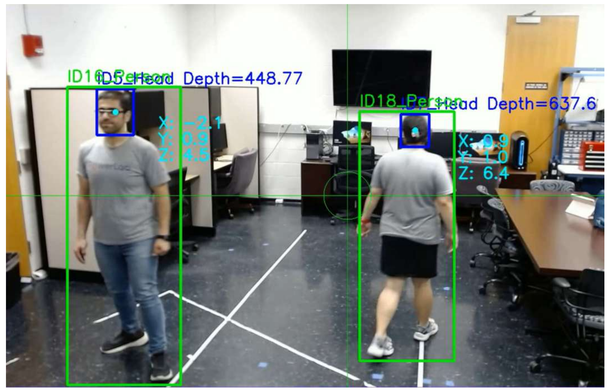
- Using triangulation and camera features we can calculate distance of head from camera.

$$D_{Head} = \frac{W_{Head\_Real} \times F}{W_{Head\_Img}}$$

- Based on homogeneous relation and camera feature we can convert 2D coordinate from image to 3D coordinate in real-world.

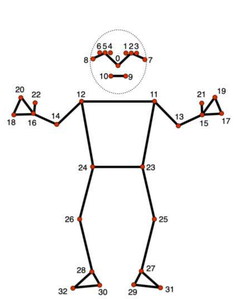
$$\begin{cases} X_{real} = D_{Head} \times \dfrac{X_{img}}{F} \\ Y_{real} = D_{Head} \times \dfrac{Y_{img}}{F} \\ Z_{real} = D_{Head} \end{cases}$$

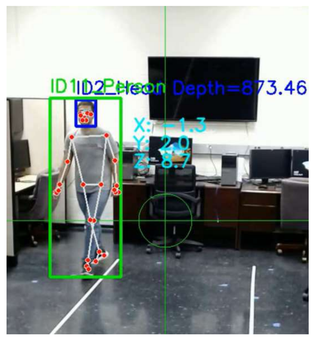
- we calculate the movement in each frame. Euclidean formula is used to calculate total distance in a time window with N frames.



III. Body Landmark Estimation and Posture Analysis:

- Mediapipe estimate 33 landmark points.





Results

- **Patient's Tracking Result:** We asked four volunteers to walk along a specific path of 13 meters with three speeds for comparison (Slow, Normal, Fast). To avoid error accumulation, we averaged position of head's center.

Size		Distance Error (%)			
Window	Overlap	Fast	Normal	Slow	Avg.
30	30	-13.23	-9.00	-3.69	8.64
30	15	-11.46	-7.92	-2.23	7.20
15	15	-9.46	-4.38	+3.08	5.65
15	10	-9.08	-3.62	+5.23	5.97
10	5	-5.54	-0.77	+13.92	6.74
10	1	-4.31	+0.15	+18.31	7.59
5	5	-0.92	+5.00	+29.92	11.94

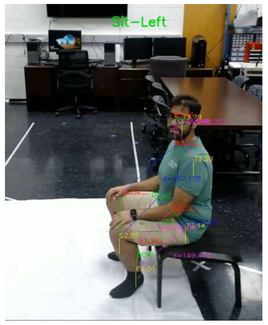
- **Human Posture Detection:** We conducted training on five distinct classifiers using 33 features extracted from main body and customized vectors, consisting of 27 angles and 6 lengths. Our posture classification involved three main classes: standing, sitting, and laying, along with 12 subcategories.

$$Sitting \begin{cases} Front \\ Back \\ Left\ Side \\ Right\ Side \end{cases} \quad Standing \begin{cases} Front \\ Back \\ Left\ Side \\ Right\ Side \end{cases} \quad laying \begin{cases} Supine \\ Prone \\ Left\ Side \\ Right\ Side \end{cases}$$

- To begin the process, we carefully select eleven specific landmarks which include the right and left eyes, shoulders, hips, knees, ankles. Utilizing these selected landmarks enables us to construct eleven vectors. Each vector consists of three dimensions.

Vector Name	Start Point	End Point
Eye Line	4	1
Shoulder Line	12	11
Hip Line	24	23
Knee Line	26	25
Ankle Line	28	27
Nose → Left Shoulder	0	11
Nose → Right Shoulder	0	12
Left Shoulder → left Knee	11	25
Right Shoulder → Right Knee	12	26
Left Hip → Left Ankle	23	27
Right Hip → Right Ankle	24	28

- As a result, we obtained a total of 33 features, including 6 lengths and 27 angles. We utilized this output to implement various classification algorithms.



Model	Accuracy	Mean Error	Avg. F1 Score
Random Forest Tree	99.8	0.3	99.6
SVM Linear	98.0	8.3	96.4
SVM Polynomial	96.1	27.3	94.8
KNN	98.2	2.1	97.5
Gradient Boosted Tree	99.2	14.1	99.1

Conclusion

- We have enhanced the accuracy of human detection and tracking by converting 2D images to 3D real world measurements using homogeneous coordinates and camera specifications.
- By tracking movement, we can accurately determine the patient's ambulation and postures, which can be useful for enhancing patient care.

Reference

1. A. M. Steele, M. Nourani, M. M. Bopp, and D. H. Sullivan, "A system to facilitate early and progressive ambulation using fiducial markers," in 2021 IEEE Biomedical Circuits and Systems Conference (BioCAS). IEEE, 2021, pp. 1-6.
2. A. Latreche, R. Kelaiaia, A. Chemori, and A. Kerboua, "Reliability and validity analysis of mediapipe-based measurement system for some human rehabilitation motions," Measurement, vol. 214, p. 112826, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0263224123003901
3. N. Aharon, R. Orfaig, and B.-Z. Bobrovsky, "Bot-sort: Robust associations multi-pedestrian tracking," 2022