IIIT Bhubaneswar

Department Of Computer Secience & Engineering

MAJOR PROJECT PRESENTATION

Project Name: Real Time Human Pose Estimation

Project Members Name

Shrey Sahay
B421047

Piyush Kumar Nayak
B521044

Hemant Sah
B421025

INDEX

S.No	Topics	Page No
1.	Constraints	4
2.	Introduction	5 - 6
3.	Types Of Pose Estimation	7 - 9
4.	Approach and Architecture	10 - 15
5.	Libraries	16 - 19
6.	Flow of Project	20 - 21
7.	Website Home Page	22
8.	Use Cases	23
9.	The Future Scope	24
10.	Conclusion	25

CONSTRAINTS

- Pose estimation is a vast topic and considering their types (2D pose estimation and 3D pose estimation) we had restricted our project to the 2D pose estimation which is the estimation of keypoints of object or image along x and y coordinate.

- On the other hand the implementation can be done in several ways. But we will estimate the single and multi pose with the help of Browser.

- Here we are not going to train the Machine Learning models, instead we had used java script libraries which have inbuilt capabilities of Machine Learning and just by using some import commands we will have access to those Machine Learning models.

INTRODUCTION

- The Human Pose Estimation is the task of using a machine learning model to estimate the approximate pose of a person from an image or a video by estimating the spatial locations of key body joints that is called keypoints.

- The Aim of this project is to deliver the basic use cases of the Pose Net model for real-time human pose estimation using a webcam feed as the data.

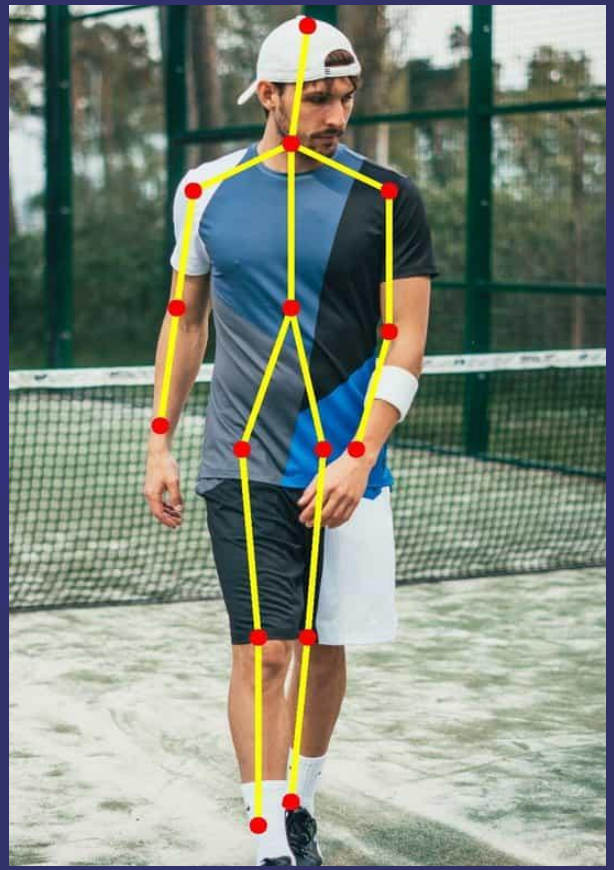
- TensorFlow.js is a JavaScript library that enables the training and deployment of machine learning models in the browser.

POSE ESTIMATION

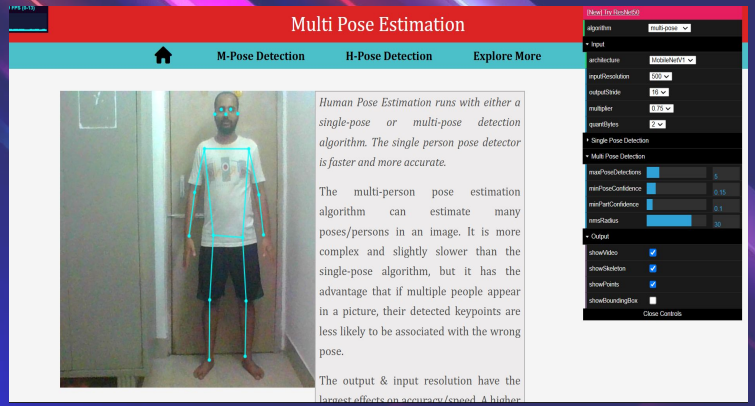
Definition: Real-time human pose estimation is a technique that involves detecting and tracking human body parts in real-time video or image data.

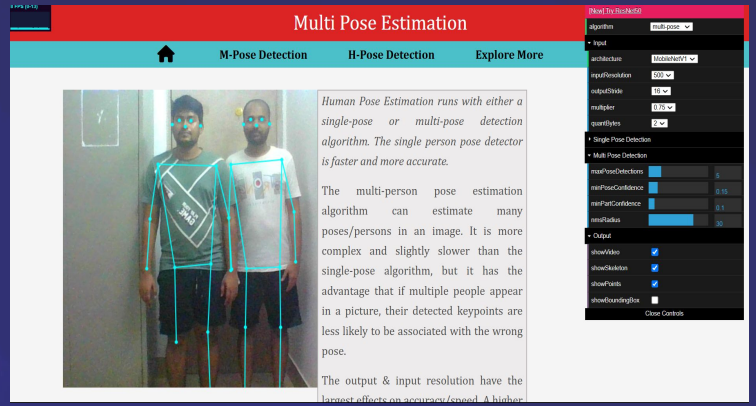
Explanation: Real-time human pose estimation uses computer vision and machine learning techniques to identify the location and orientation of each body part.

Use cases: Human action recognition, human-computer interaction, video surveillance, fitness, health tracking, and gesture recognition.



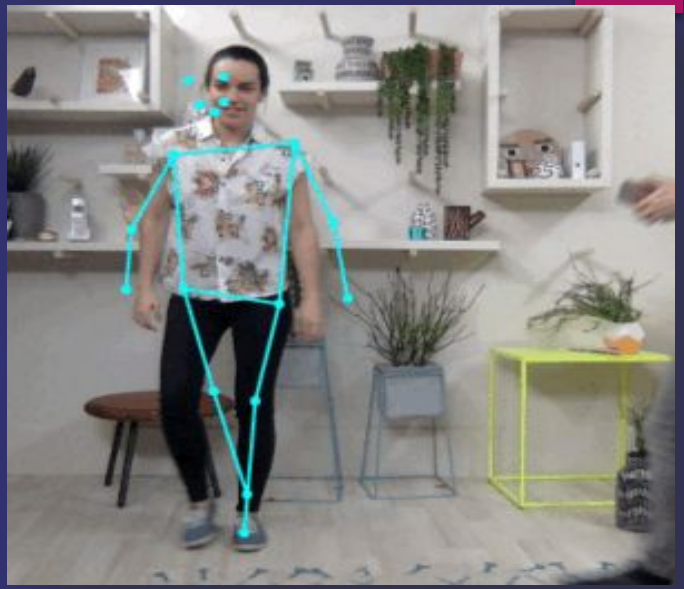
TYPES OF POSE ESTIMATION





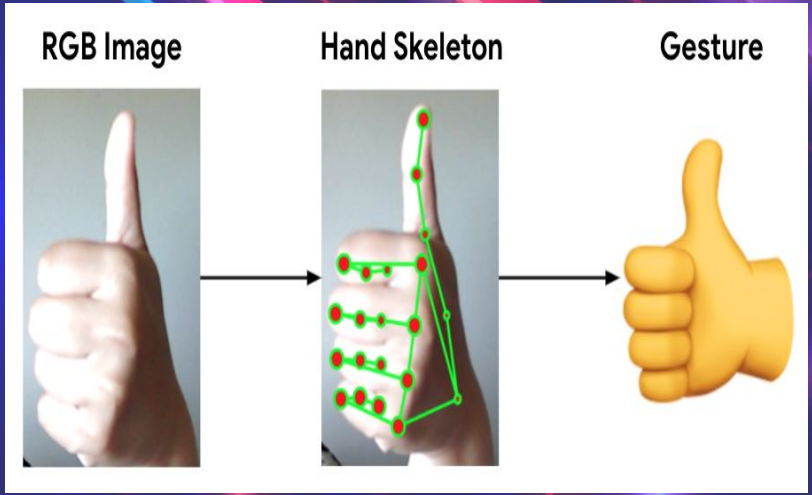
SINGLE POSE ESTIMATION

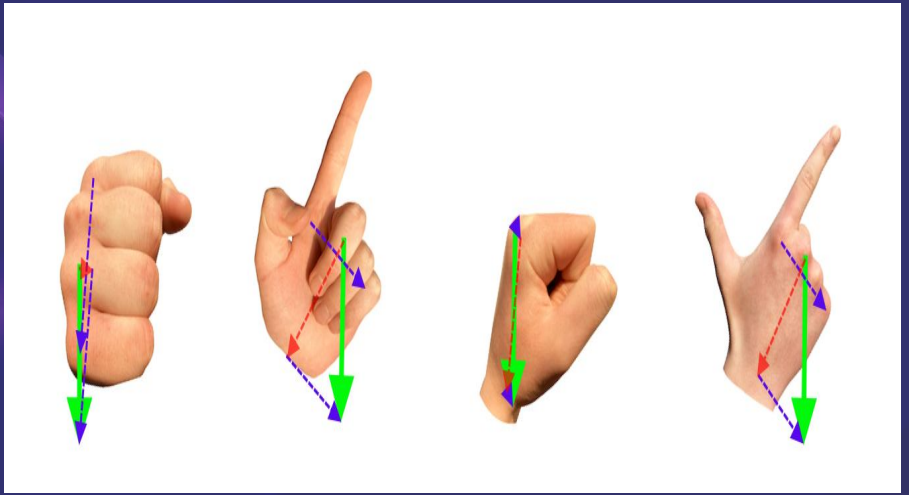
MULTI POSE ESTIMATION



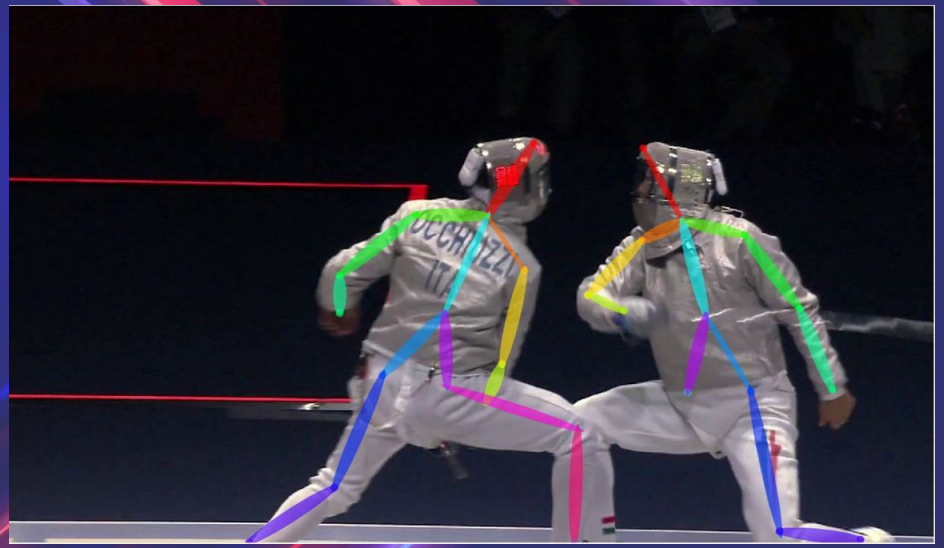
Pose Estimation can estimate either a single pose or multiple poses, meaning there is a version of the algorithm that can detect only one person in an image/video and one version that can detect multiple persons in an image/video.

HAND POSE ESTIMATION



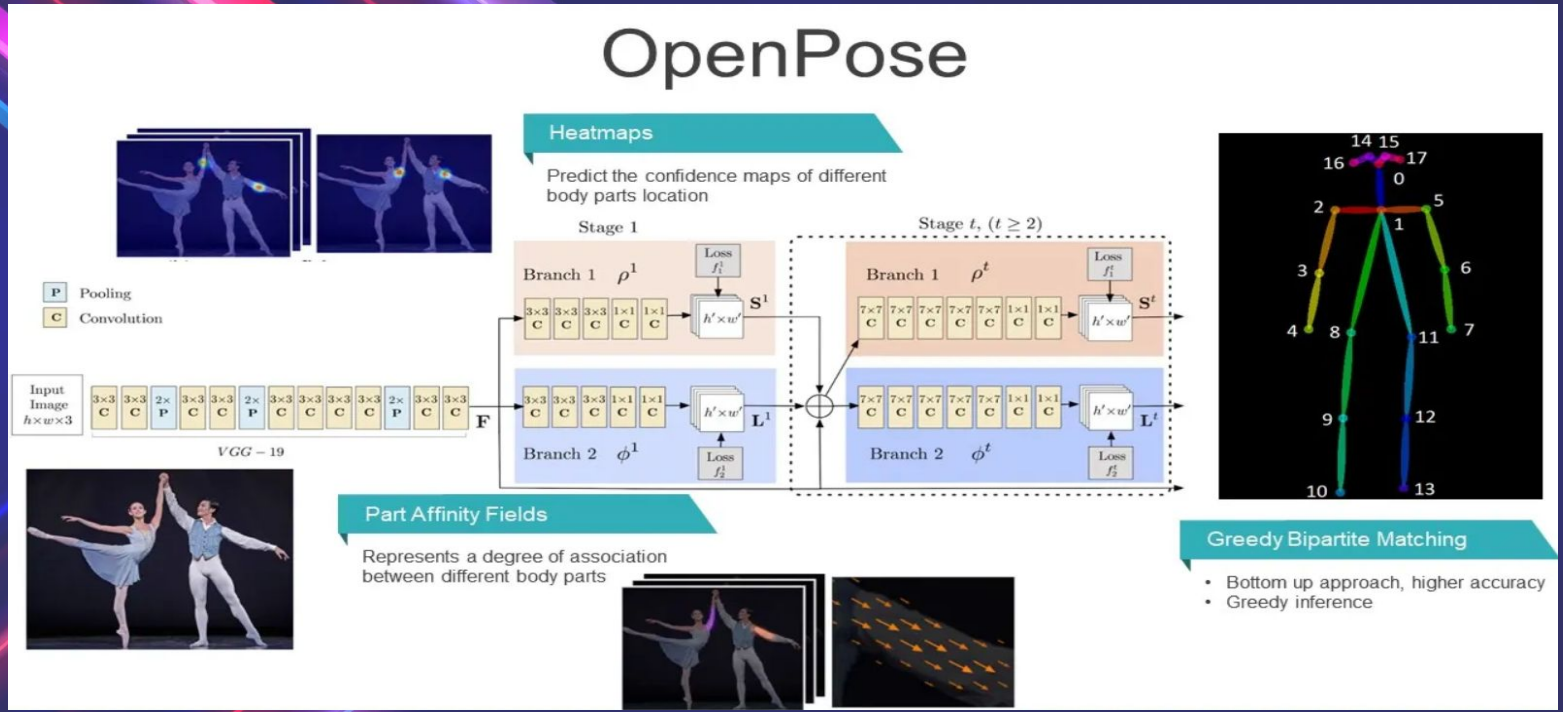


Bottom Up approach

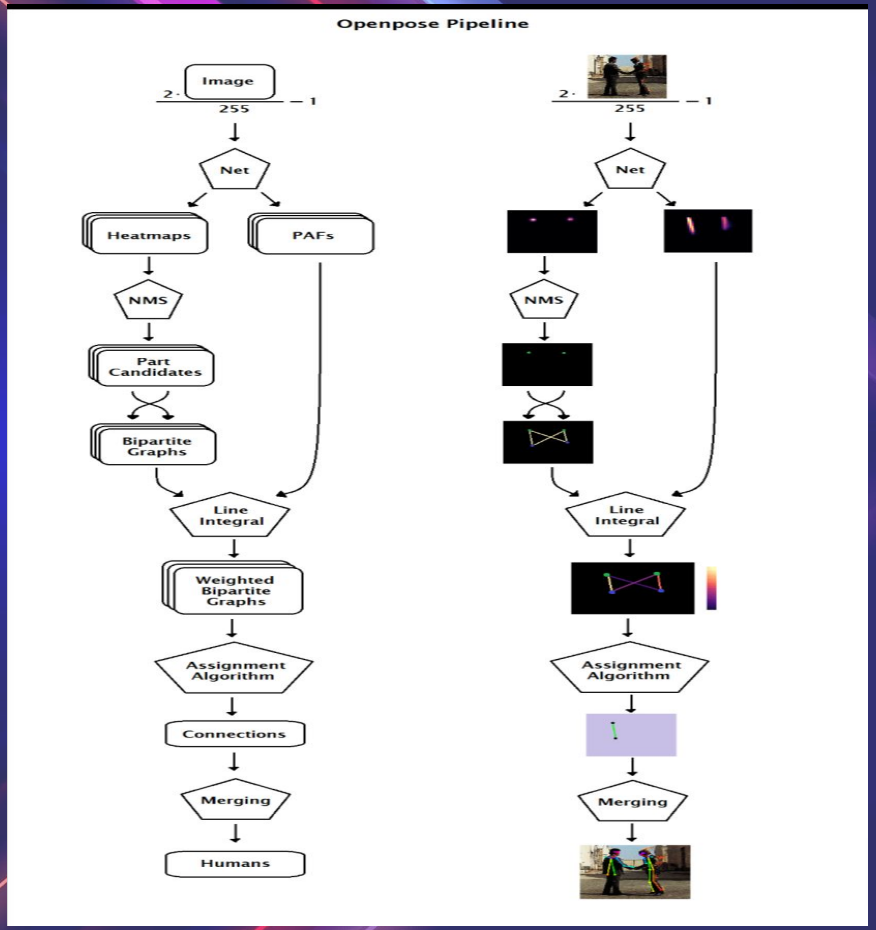


- Instead of detecting human and then going on for pose estimation for every individual, we are going for a bottom up approach.
- In Bottom up approach, we will find keypoints like elbow, knee, hand etc and then go on for association of them.

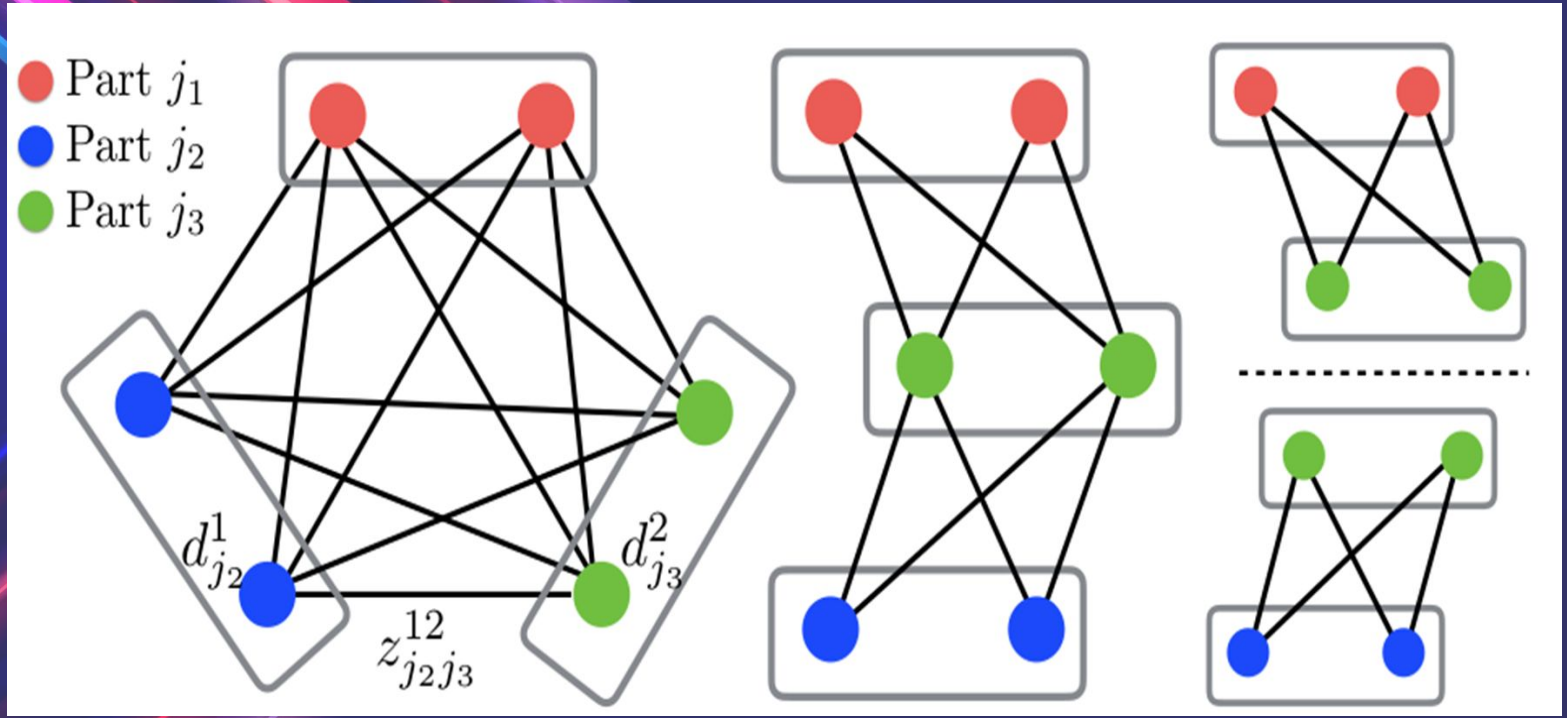
OpenPose Architecture



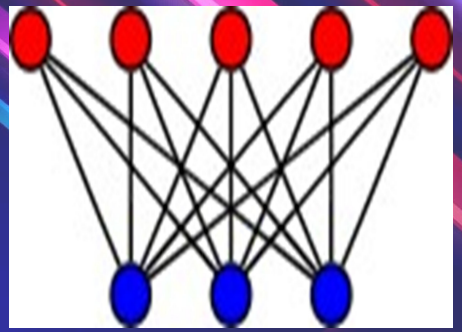
Path affinity fields

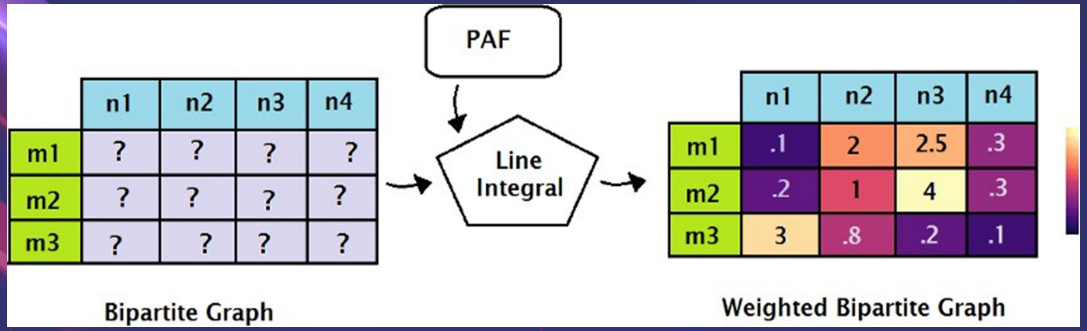


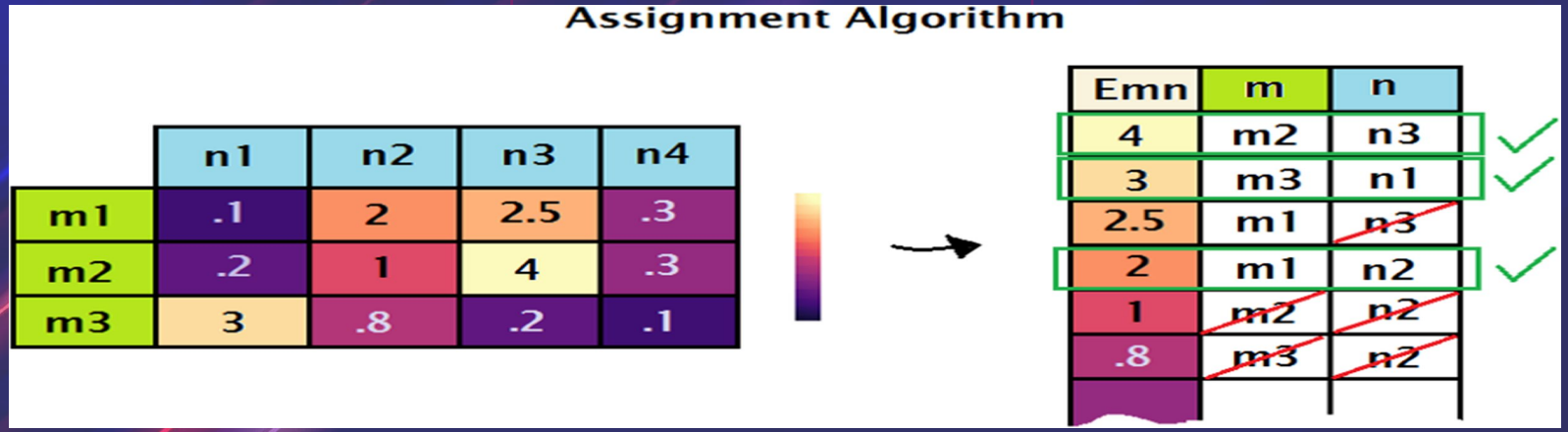
Bottom Up Parsing



Bipartite Graph







Merging Algo

if $H_1 \cap H_2 \neq \emptyset$
then
$H_1 = H_1 \cup H_2$
delete(H_2)

Pose Estimation Output

Libraries/Frameworks and Models

Following are the technologies that we had used to implement the real time human pose estimation project in web browser .

- TensorFlow
- Mediapipe
- Posenet

1. Tensor Flow

Definition: TensorFlow.js is a JavaScript library that enables the training and deployment of machine learning models on the browser and Node.js.

Overview of features: It allows developers to use high-level APIs and pre-trained models, run models in the browser, and create interactive applications with machine learning capabilities.

Links: https://www.tensorflow.org/

Datasets Used : COCO , MPII for pose estimation
300W-LP for face detection
Hand Tracking 21-Keypoint

2. MediaPipe

Definition: MediaPipe is an open-source machine learning framework developed by Google for building real-time applications that process video, images, and audio.

Overview of features: It supports cross-platform deployment (Web, Android, iOS, Desktop) with low-latency inference optimized for CPU & GPU acceleration. Its modular pipeline enables efficient multi-task learning, making it ideal for AI-powered applications in AR, gaming, and healthcare .

Link

3. Posenet

Definition: PoseNet is a deep learning-based real-time human pose estimation model developed by Google. It detects keypoints of the human body (nose, eyes, shoulders, elbows, etc.) from images or
Videos .

Overview of features: Optimized for mobile and web applications, it runs seamlessly in browsers using TensorFlow.js and supports CPU, GPU, and WebGL acceleration.

Link

FLOW OF THE PROJECT

1. Data Collection And Data Preprocessing
2. Model Selection
3. Integration with Web Browser
4. User Input
5. Inference
6. Visualization

FLOW OF THE PROJECT

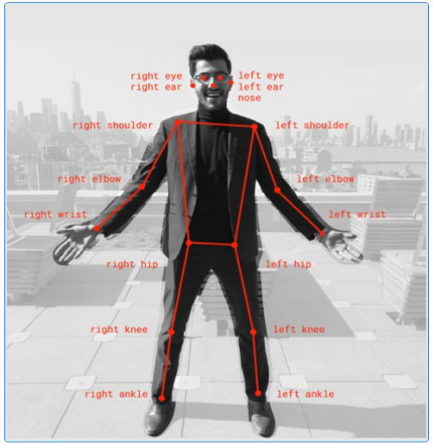
7. Optimization
8. Testing
9. Deployment

WEBSITE HOME PAGE

Welcome To Human Pose Estimation

The Human Pose Estimation is the task of using a machine learning model to estimate the approximate pose of a person from an image or a video by estimating the spatial locations of key body joints that is called keypoints.

- There are total 17 keypoints that are used by algorithm to estimate the pose of human body.
- This step is a crucial prerequisite to multiple tasks of computer vision which include human action recognition, human tracking, human-computer interaction and video surveillance.
- It can be used to estimate either a single pose or multiple poses, meaning there is a version of the algorithm that can detect only one person in an image/video and one version that can detect multiple persons in an image/video.
- The aim is to deliver the basic use cases of the Pose Net model for real-time human pose estimation using a webcam feed as the data. Now, the challenge is to create an advanced webcam filter that has detection functionalities like the Snapchat camera.



Download Document

- Project Report
- Project Presentation

|Human Pose Estimation © 2025 All Right ®|
|View Tearms Of Use & Privacy Blueprint|

USE CASES

Action prediction

Surveillance

Cloth Parsing

Online Coaching

Movie and Game

AR and VR

Healthcare

THE FUTURE SCOPE

1. Real-Time Pose Estimation on Edge Devices
2. Privacy and Security Considerations
3. Real-Time Feedback and Coaching
4. Collaborative Pose Estimation
5. Self-Supervised Learning
6. Virtual Try-On and Fashion
7. 3D Pose Estimation

CONCLUSION

1. Demonstrated PoseNet for real-time pose estimation via webcam.
2. Next goal: build a Snapchat-style filter with pose detection.
3. Pose estimation has broad uses in VR, gaming, sports, and health.

RELATED LINKS

1. **Figma Design**

   https://www.figma.com/proto/IwBTcGW9T4qGUhQMAbepp8/Human-Pose-Estimation?node-id=210%3A67&scaling=contain&page-id=0%3A1

2. **Source Code Is Pushed To GitHub**
   https://github.com/sahemant12/HumanBody-Pose-Estimation

1. **Our Pose Estimation Website**
   https://humanbody-pose-estimation.netlify.app/