

IIIT BHUBANESWAR



Department Of Computer Science & Engineering

MAJOR PROJECT PRESENTATION

Designing a Web Interface for Finding Human Body Pose Estimation

Project Members Name



```
graph TD; A[Project Members Name] --> B[Shrey Sahay  
B421047]; A --> C[Piyush Kumar  
Nayak  
B521044]; A --> D[Hemant Sah  
B421025];
```

The diagram is an organizational chart. At the top is a purple box labeled 'Project Members Name'. Three arrows point down from this box to three separate boxes below. Each box has a blue header and a light gray body. The first box contains 'Shrey Sahay' and 'B421047'. The second box contains 'Piyush Kumar', 'Nayak', and 'B521044'. The third box contains 'Hemant Sah' and 'B421025'.

Shrey Sahay
B421047

Piyush Kumar
Nayak
B521044

Hemant Sah
B421025

INTRODUCTION

Human Pose Estimation :

- Uses machine learning to estimate spatial locations of body joints (key points) from images or videos.
- It identifies key points like shoulders, elbows, knees, etc.
- Helps track the position and movement of body joints.

Project Aim : Demonstrate real-time pose estimation using Pose Net and webcam feed.

Technology Used : TensorFlow.js – A JavaScript library for training and deploying ML models in the browser.



CONSTRAINTS

Scope Boundaries

- Focus is limited to 2D Pose Estimation (X and Y coordinates of key points).
- 3D Pose Estimation is excluded for simplicity.
- We estimate both single and multi-person poses.
- Implementation runs in a web browser environment.
- No training of ML models is done manually.
- Uses JavaScript libraries with pre-trained ML capabilities.
- Access to models is done via simple import commands.



POSE ESTIMATION

Definition: Technique for detecting and tracking human body parts in real-time video or images

Explanation: Uses computer vision and machine learning to locate and determine orientation of body parts.

Use cases:

- Human action recognition
- Human-computer interaction
- Video surveillance
- Fitness & Health tracking
- Gesture recognition



[illegible][illegible]

Hand Pose Estimation

 All Pose Structures
  All Pose Structures
  Explore More



Left Hand

Id	Winger	Joint	Structure
0	Thumb	Root Joint	Extended Up
1	Index	Root Joint	Extended Up
2	Middle	Root Joint	Extended Up
3	Ring	Root Joint	Extended Up
4	Pinky	Root Joint	Extended Up

Right Hand

Id	Winger	Joint	Structure
0	Thumb	Root Joint	Thompson Up Right
1	Index	Root Joint	Thompson Up Left
2	Middle	Root Joint	Extended Up
3	Ring	Root Joint	Extended Up
4	Pinky	Root Joint	Extended Up

Pose Collections
[About Us](#)
[Terms Of Use](#)
[Privacy Policy](#)

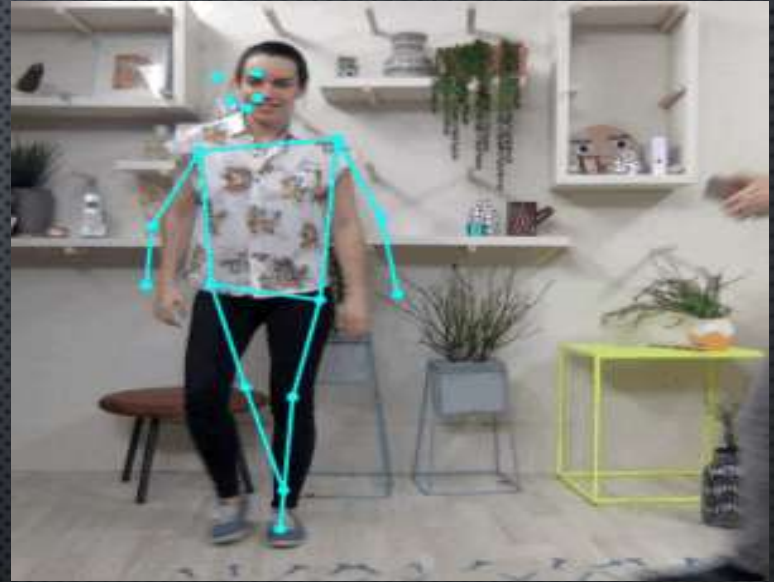
Get Help
[Feedback](#)
[Contact Us](#)
[Help & Support](#)

Address
 617 Shikharwadi,
 710005 Gandhinagar,
 Gandhinagar, India

Follow Us



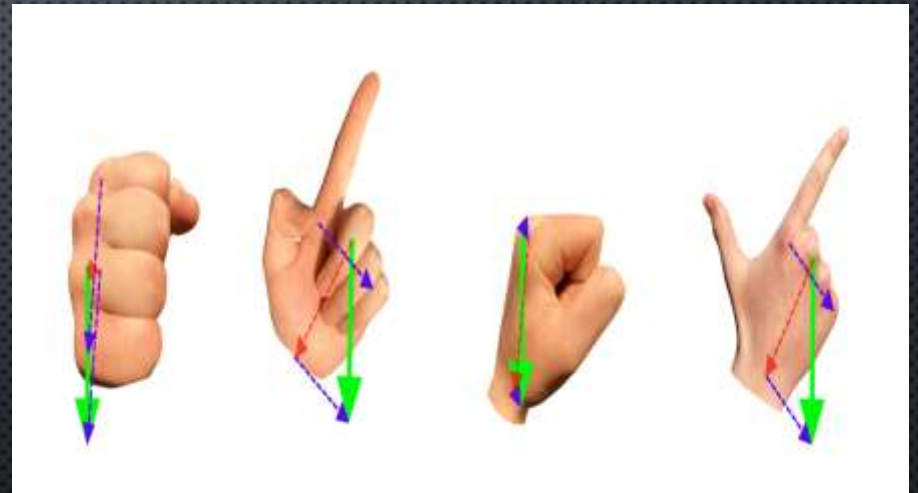
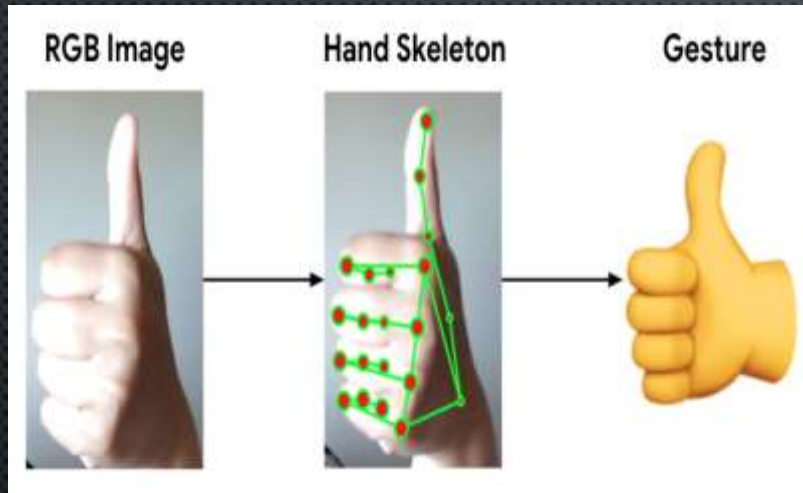
SINGLE POSE & MULTI POSE ESTIMATION



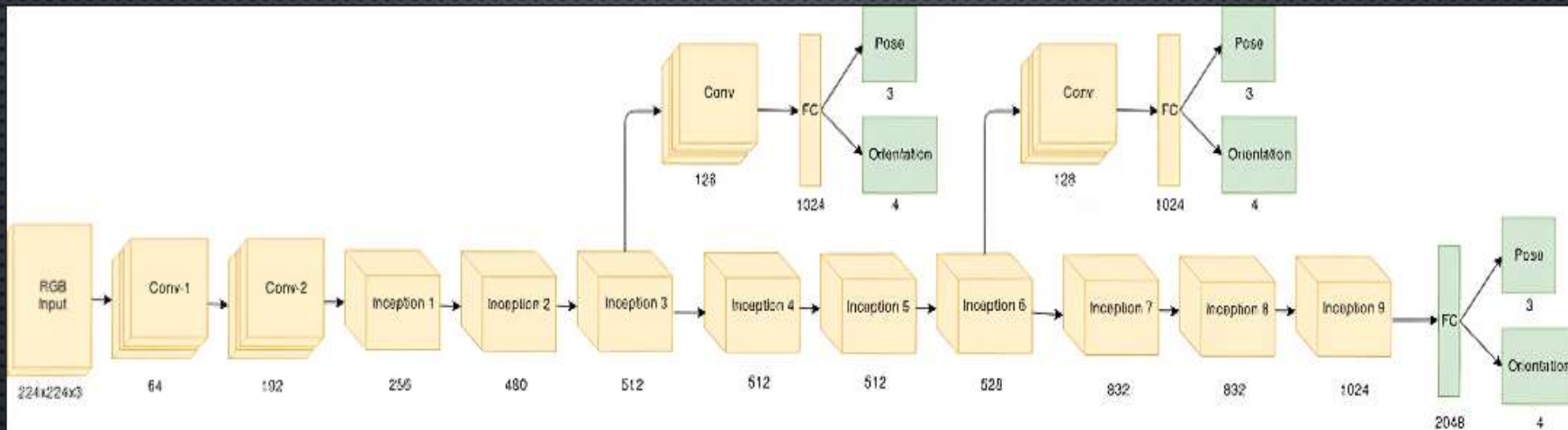
Pose estimation can operate in two modes :

- Single Pose Estimation – Detects the pose of only one people in an image/ video.
- Multi Pose Estimation – Detects the poses of multiple people in an image/ video.

HAND POSE ESTIMATIONS

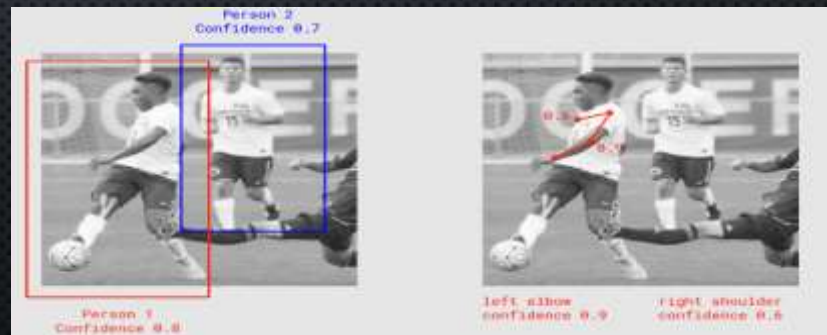


POSE NET ARCHITECTURE



Key Terminologies :

- Pose
- Pose Confidence Score
- Keypoint
- Keypoint Confidence Score
- Keypoint Position



ALGORITHMS



Fig - 17 Pose Key points Returned By PoseNet

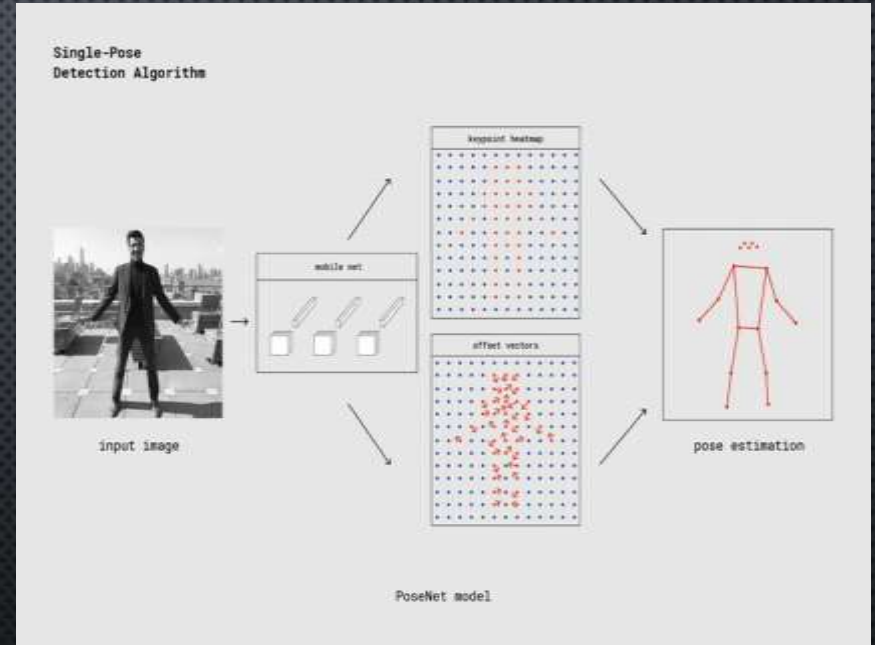


Fig - Single person pose detector pipeline using PoseNet

ESTIMATING POSES FROM THE OUTPUTS OF THE MODEL

Key points Detection Steps :

- Sigmoid Activation applied to heatmap to obtain key point scores
- Argmax2D used to find (y, x) index with highest score for each key point
- Offset Vectors retrieved for each key point using corresponding heatmap indices
- Final Key point Positions computed using output stride and offset vectors
- Key point Confidence = score at heatmap position
- Pose Confidence = average of all key point confidence scores



ALGORITHMS

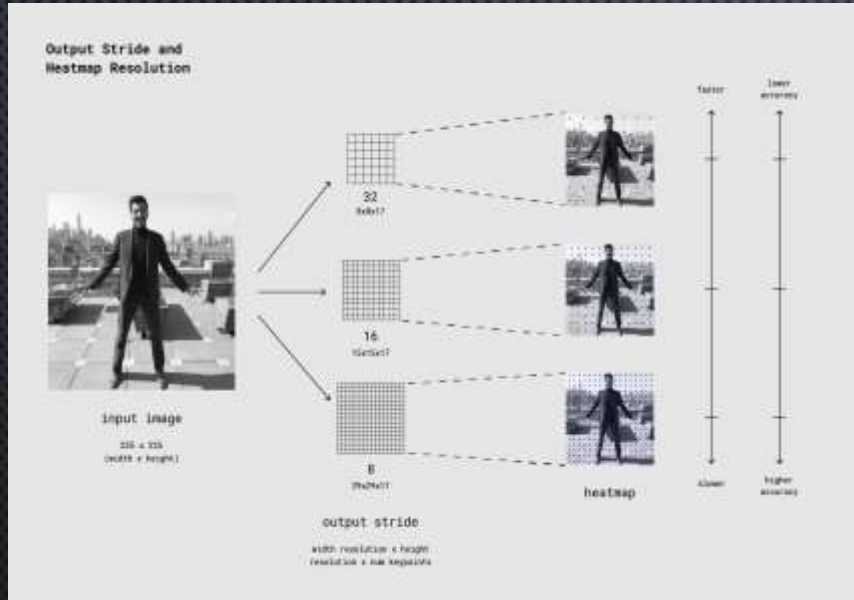


Fig - Processing Model Inputs: an Explanation of Output Strides

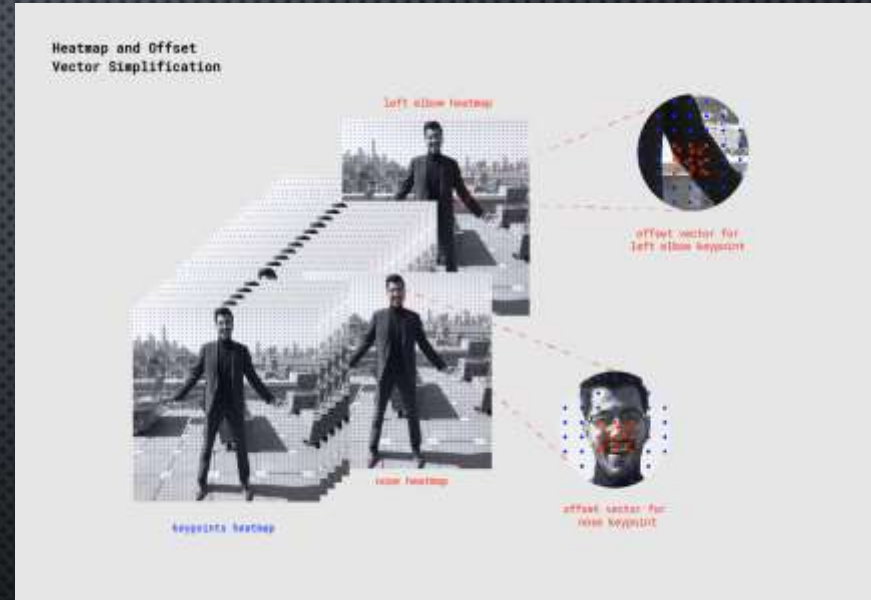


Fig - Model Outputs: Heatmaps and Offset Vectors

LIBRARIES/ FRAMEWORKS/ MODELS

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

The TensorFlow logo is represented by a stylized, rounded rectangular box with a magenta border and a light pink fill. The text "TensorFlow" is centered within the box in a black serif font.

TensorFlow

The Mediapipe logo is represented by a stylized, rounded rectangular box with a magenta border and a light pink fill. The text "Mediapipe" is centered within the box in a black serif font.

Mediapipe

The Posenet logo is represented by a stylized, rounded rectangular box with a magenta border and a light pink fill. The text "Posenet" is centered within the box in a black serif font.

Posenet



TENSORFLOW

Definition: A JavaScript library for training and deploying ML models in the browser and Node.js.

Overview of features:

- Supports high-level APIs and pre-trained models
- Enables browser-based model execution
- Facilitates interactive ML-powered applications

Official Links: <https://www.tensorflow.org/>

Datasets Used :

- COCO , MPII for pose estimation
- 300W-LP for face detection
- Hand Tracking 21-Keypoint for hand pose tracking



MEDIA-PIPE

Definition: Open-source ML framework by Google for real-time processing of video, images, and audio.

Key features:

- Supports cross-platform deployment: Web, Android, iOS, Desktop
- Optimized for low-latency inference on CPU & GPU
- Modular pipeline for efficient multi-task learning
- Ideal for AR, gaming, healthcare, and other AI applications

Official Link: [Media-Pipe](#)



POSENET

Definition:

- A real-time human pose estimation model by Google.
- Detects key points like nose, eyes, shoulders, elbows, etc., from images or videos.

Key features:

- Optimized for mobile and web apps
- Runs directly in browsers via TensorFlow.js
- Supports CPU, GPU, and WebGL acceleration
- Suitable for interactive and low-latency applications

Official Link: [Posenet](https://github.com/google/posenet)



FLOW OF THE PROJECT

S.No.	Flow
1	Data Collection And Data Preprocessing
2	Model Selection
3	Integration with Web Browser
4	User Input
5	Inference
6	Visualization
7	Optimization
8	Testing
9	Deployment



WEBSITE HOMEPAGE

Welcome To Human Pose Estimation

[M-Pose Detection](#)[H-Pose Detection](#)[Explore More](#)

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](#)



Pose Estimation

[About Us](#)[Terms Of Use](#)[Privacy Blueprint](#)

Get Help

[Feedback](#)[Contact Us](#)[Help & Support](#)

Address

HIT Bhubaneswar,
751003 Gothapatna,
Odisha, India

Follow Us



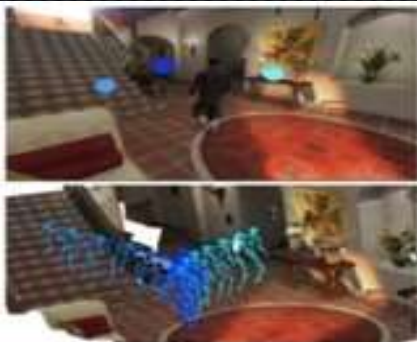
[Human Pose Estimation © 2025 All Right @]
[View Terms Of Use & Privacy Blueprint]

Links :

- [Repository](#)
- [Website](#)



USE CASES



Action prediction



Surveillance



Cloth Parsing



Online Coaching



Movie and Game



AR and VR



Healthcare

FUTURE SCOPES

S.No.	Future Scopes
1	Real – Time Pose Estimation on Edge Devices
2	Privacy and Security Considerations
3	Real – Time Feedback and Coaching
4	Collaborative Pose Estimation
5	Supervised Learning
6	Virtual Try – On and Fashion
7	3D Pose Estimation



CONCLUSION

Demonstrated basic PoseNet use cases with real-time webcam input.

Real-Time human pose estimation uses CV + ML to track joint positions and orientation.

Widely applicable in:

- Virtual Reality
- Gaming
- Sports Analytics
- Health Monitoring

