

# Pose-Tracking Model Decision Worksheet

Five questions, three topologies, four pitfalls — pick the right pose model in ninety seconds.

## Five-question decision tree

### 1. Is this a commercial product?

- Yes → do not ship OpenPose. License is non-commercial.
- No (research / internal) → any family is fine.

### 2. Where does the model run?

- Browser / user device → MediaPipe Pose.
- Edge box near camera → RTMPose-m / -s.
- Cloud server → continue to question 3.

### 3. How many concurrent streams?

- > 10 / server → RTMPose-m on CPU (90+ FPS).
- ≤ 10 / server, max accuracy → ViTPose-L on GPU.

### 4. Need whole-body keypoints?

- Hands + face landmarks → RTMW (133 keypoints).
- Just COCO 17 / BlazePose 33 → smaller model.

### 5. Real 3D measurements?

- Clinical-grade → multi-camera triangulation rig.
- UX feedback only → MediaPipe 3D / RTMW3D.

## Deployment topology — default model

Topology	Model	Cost owner	Latency
Browser	MediaPipe Pose	User device	~16 ms
Edge	RTMPose-m / -s	Hardware	~50 ms
Server CPU	RTMPose-m	Cloud / stream	~30 ms
Server GPU	ViTPose-L	Cloud / stream	~20 ms

## Four pitfalls — pre-launch checklist

- License: OpenPose not shipped in a paid product.
- Smoothing: One-Euro filter applied per keypoint.
- Frame rate: pose runs ≤ 10 FPS for slow-motion features.
- 3D: single-camera 3D is UX feedback, not measurement.

## Pipeline stages

- 1 · Detection — RTMDet-tiny or YOLOv8-person.
- 2 · Pose — one of the four families.
- 3 · Tracking — ByteTrack on bounding boxes.
- 4 · Smoothing — One-Euro filter per keypoint.