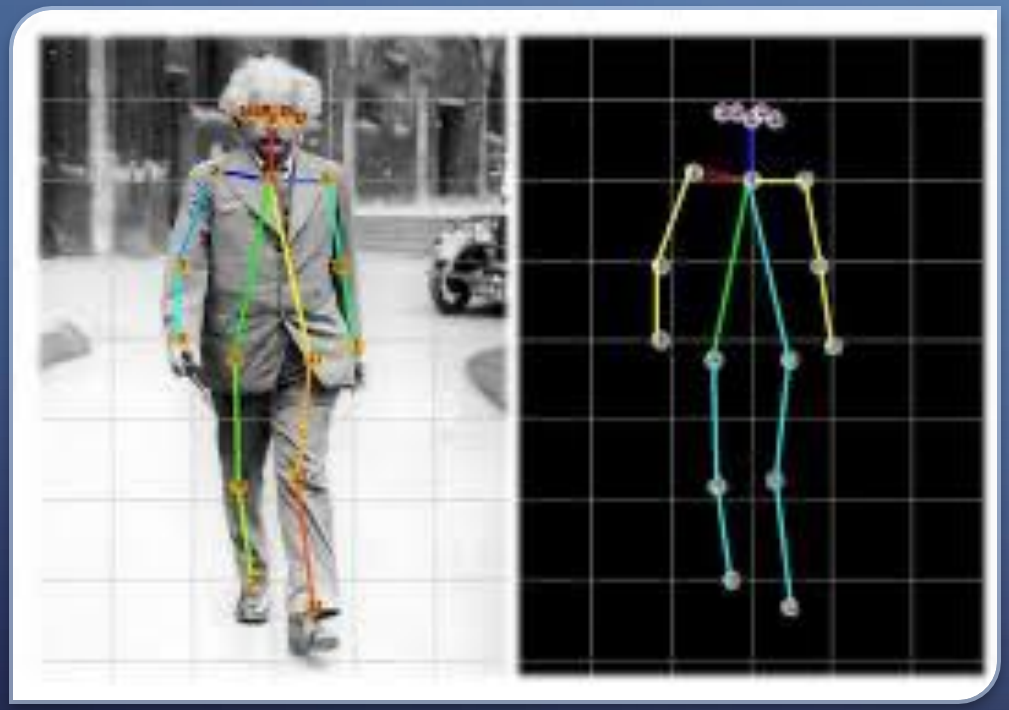
HUMAN POSE ESTIMATION: UNDERSTANDING AND MODELING HUMAN MOVEMENT

ELIZABETH THOMPSON

DECEMBER 4, 2025

WHAT IS HUMAN POSE ESTIMATION?



- Detect key body joints
- Connect joints to form a skeleton
- Analyze human movement in images and videos

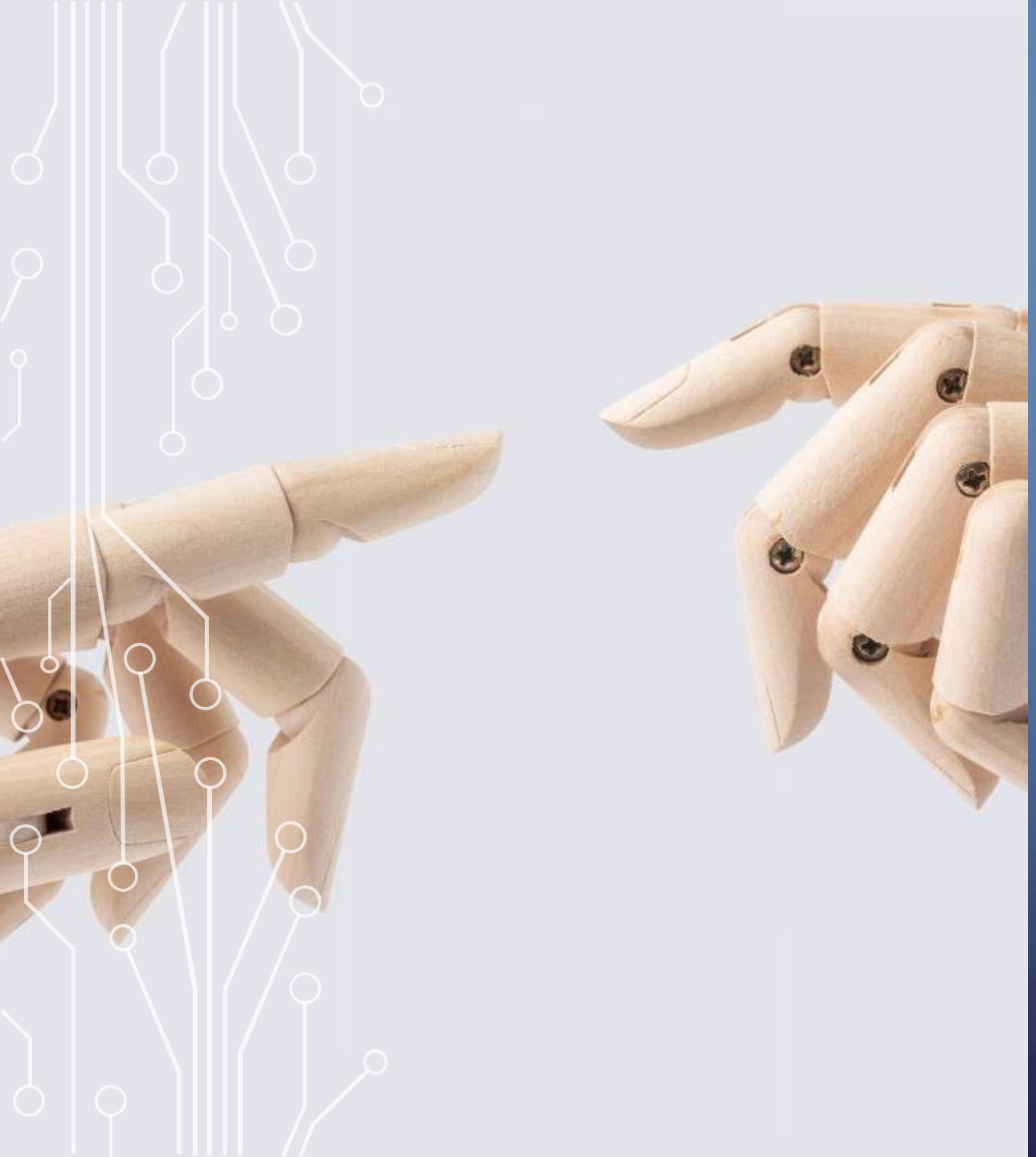
APPLICATIONS

Sports analytics: track player movement

Animation & gaming: motion capture

Healthcare: position / rehab monitoring

Human-computer interaction: gesture control



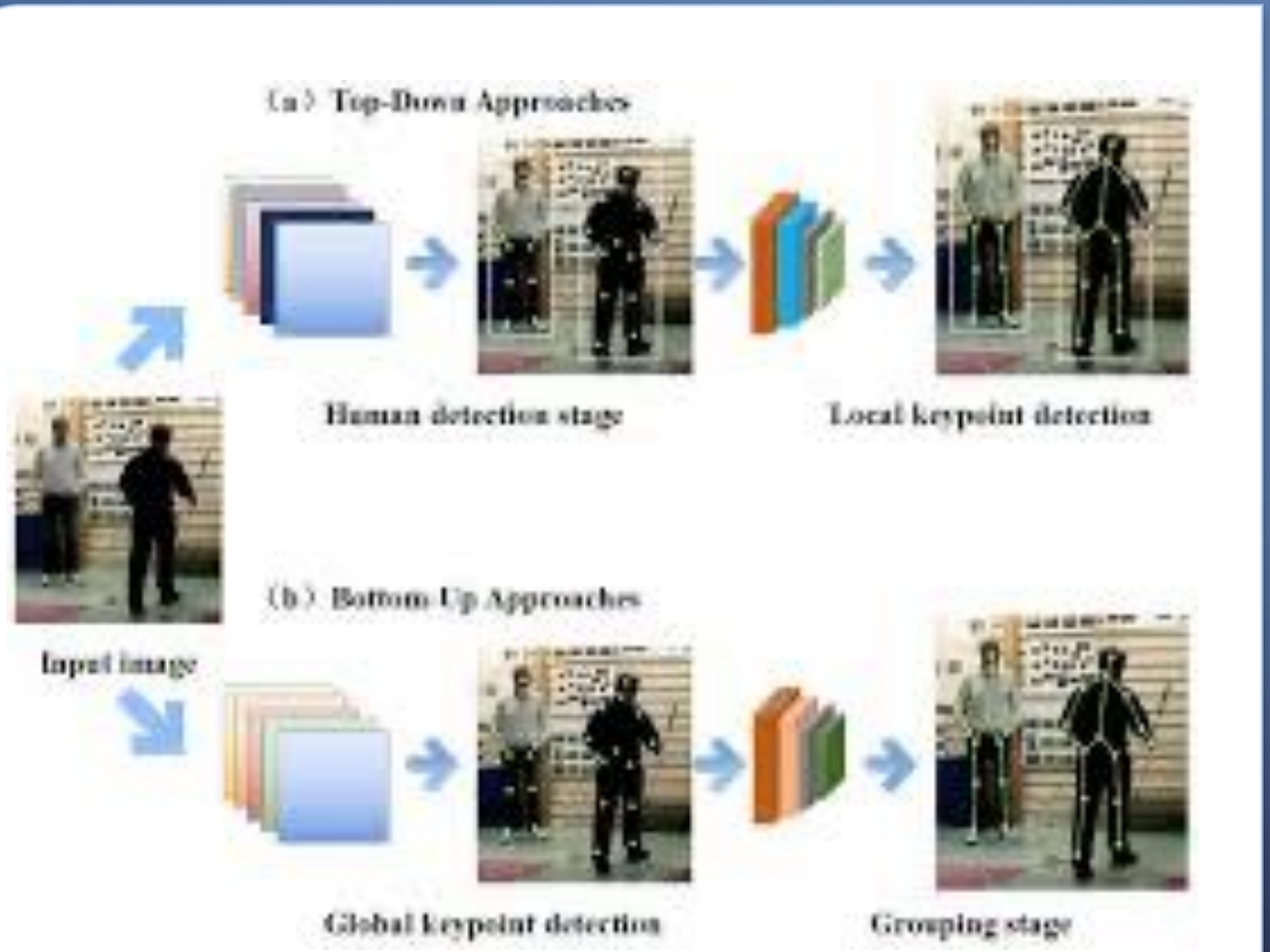
KEY CONCEPTS

- Keypoints: joints
- Skeletons: connect keypoints to visualize posture
- 2D vs 3D: 2D is simpler; 3D adds depth

(a) (b) (c) (d)

CHALLENGES

- Occlusion: body parts blocked
- Multiple people in the image
- Variation in clothing, lighting, camera angles
- Real-time video requires fast processing



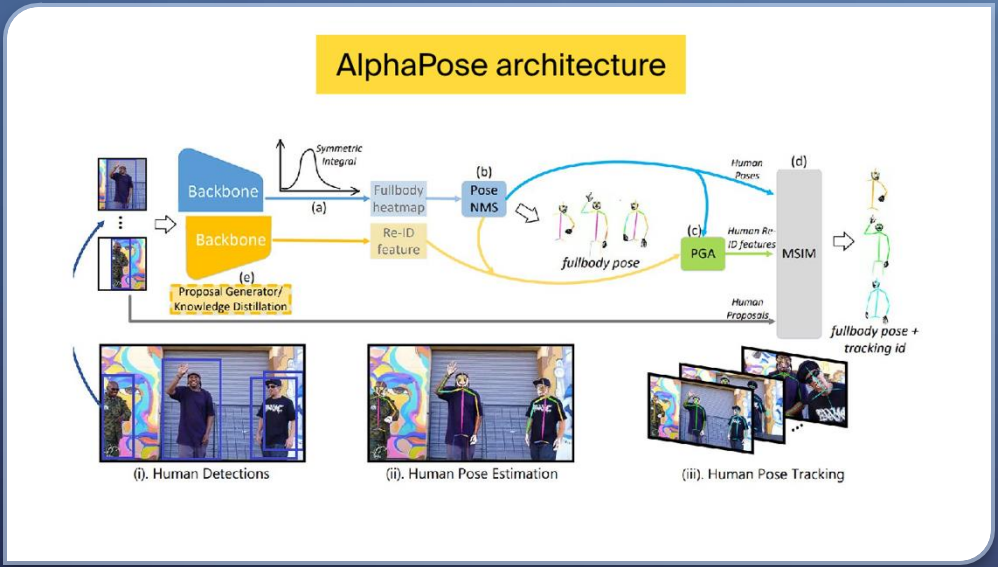
APPROACHES

- Top-down: detect people → then keypoints
- Bottom-up: detect all keypoints → group into people

POPULAR MODELS AND TECHNIQUES

- OpenPose: real-time multi-person 2D pose estimation
- HRNet: maintains high-resolution features
- PoseNet: lightweight, runs in-browser or on mobile

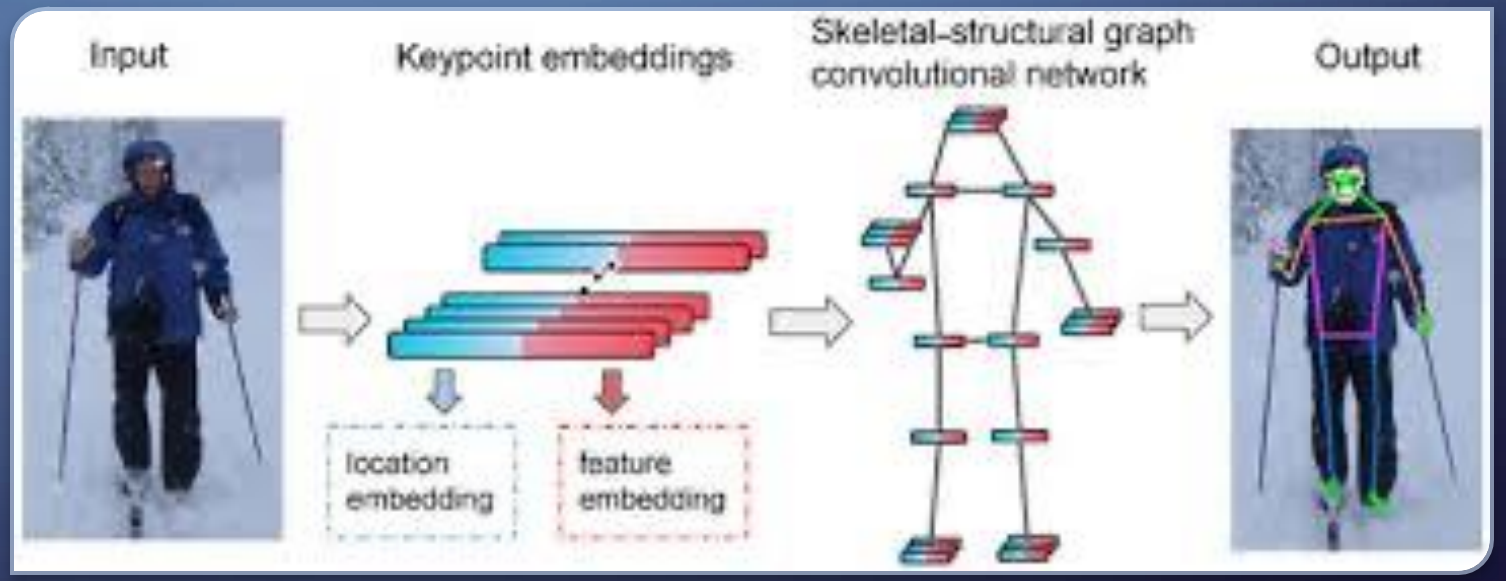
HOW POSE ESTIMATION IS MODELED



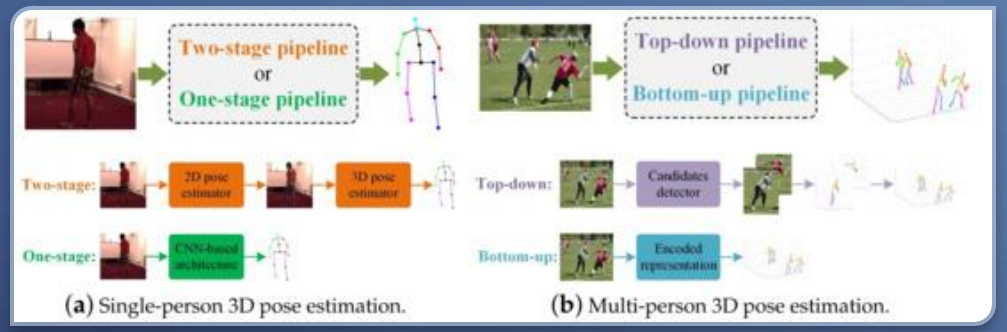
- Uses convolutional neural networks (CNNs) to extract image features
- Outputs heatmaps for each keypoint
- Loss function compares predicted vs. true keypoints
- Training on images with labeled keypoints

MODEL ARCHITECTURE AND TRAINING WORKFLOW

- Backbone CNN extracts multi-scale features
- Feature maps passed through refinement layers
- Heatmaps decoded to keypoint coordinates



ADVANCED MODELING AND OPTIMIZATION TECHNIQUES



- Multi-person pose estimation: top-down vs. bottom-up pipelines
- Data augmentation: rotation, scaling, flipping for robustness
- Loss functions: MSE for heatmaps, PCK-based penalties
- Optimization: Adam or SGD, learning rate scedules

MODERN ENHANCEMENTS AND DEPLOYMENT

- Lightweight architectures for mobile / embedded devices
- Real-time optimizations: model pruning, quantization
- Integration with video pipelines for smooth temporal tracking
- Multi-task learning: pose + action recognition
- Pretrained models and transfer learning

REFERENCES

Foundational Sources:

Cao, Z., Hidalgo, G., Simon, T., Wei, S.-E., & Sheikh, Y. (2018). *OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields*. arXiv preprint arXiv:1812.08008.

Pishchulin, L., Insafutdinov, E., Tang, S., Andres, B., Gehler, P., & Schiele, B. (2015). *DeepCut: Joint Subset Partition and Labeling for Multi-Person Pose Estimation*. arXiv preprint arXiv:1511.06645.

Bogo, F., Kanazawa, A., Lassner, C., Gehler, P., Romero, J., & Black, M. J. (2016). *Keep it SMPL: Automatic Estimation of 3D Human Pose and Shape from a Single Image*. arXiv preprint arXiv:1607.08128.

Image Sources:

MDPI. (2021). "Human Pose Estimation: …" (source URL: https://www.mdpi.com/2078-2489/13/3/109

ResearchGate. Illustration of top-down vs bottom-up pose estimation approaches. (source URL: https://www.researchgate.net/figure/Approaches-to-2D-multi-person-pose-estimation-a-top-down-and-b-bottom-up_fig4_393274050)

MobiDev blog. "Human Pose Estimation Technology Guide." (source URL: https://mobidev.biz/blog/human-pose-estimation-technology-guide)

Nature article. (source URL: https://www.nature.com/articles/s41598-024-58175-8)

V7 Labs blog. "Human Pose Estimation Guide." (source URL: https://www.v7labs.com/blog/human-pose-estimation-guide) ITRex Group blog. "Human Pose Estimation: Use Cases & Implementation Tips." (source URL: https://itrexgroup.com/blog/human-pose-estimation-use-cases-implementation-tips/)