



DS323: AI in Design (AIID)

Autumn 2023

# Week 04 Lecture 07

## AIID + Body

Wan Fang

Southern University of Science and Technology

---

# Agenda

- Body
  - Introduction to Human Pose Recognition
  - History of Human Pose Recognition
  - Re/defining Creativity
  - Proceed with Caution
- Generative Design

# Introduction to Human Pose Recognition

Human body language is an intrinsic component of our lived experience.

# From physical to digital

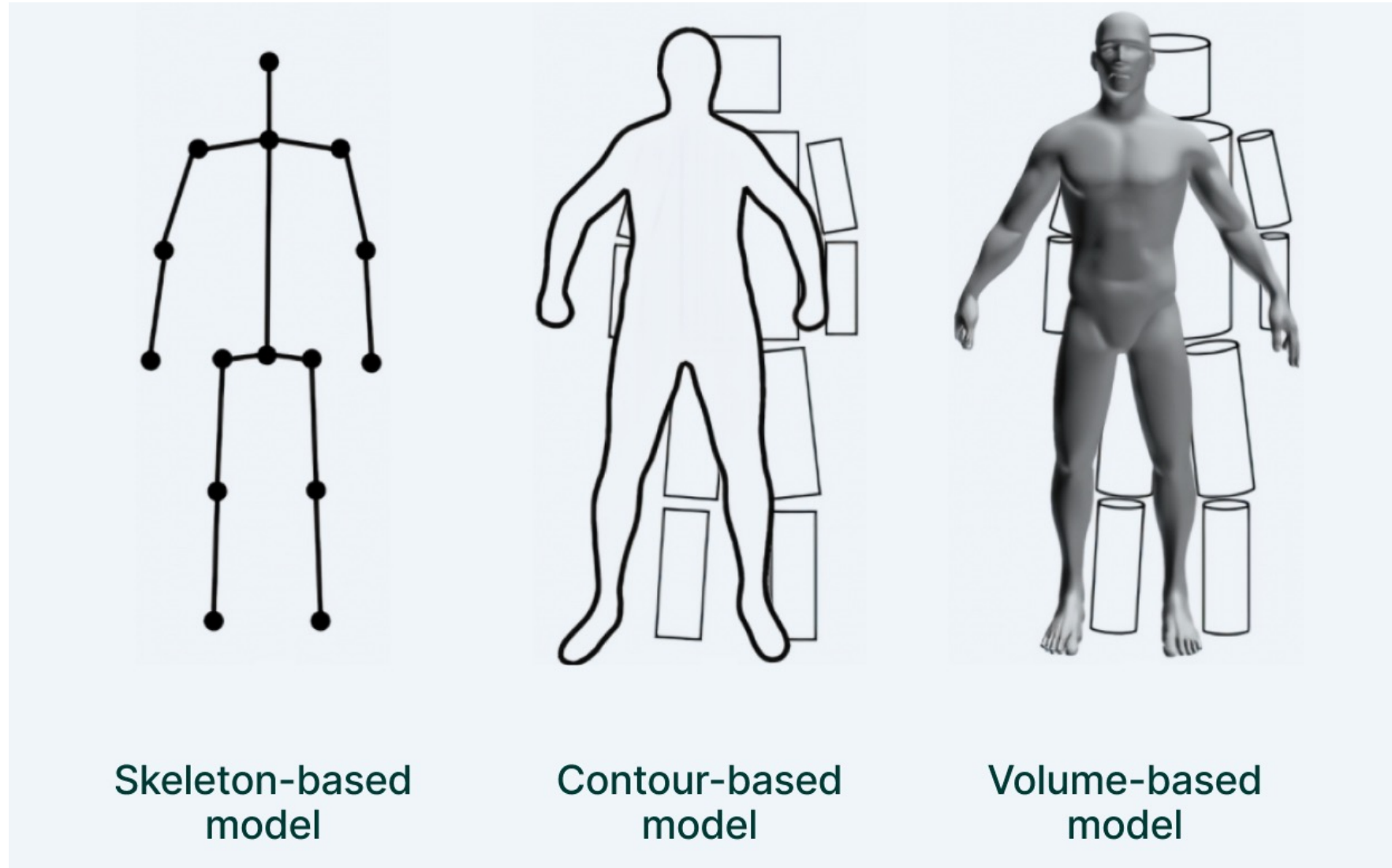
- In the digital age, we have unfortunately seen this complex language fade into the background.
- How can we encompass our body language as a tool for communication?

---

# What is Human Pose Estimation?

- A branch of *Computer Vision* research
- A technique that allows us to accurately detect and predict/estimate the pose of a person
- Achieved by identifying and classifying the *coordinates of the joints of a human body*, such as wrists, shoulders, knees, arms  
(..) commonly known as *landmarks*
- There are three types of approaches to model the human body

# Human Body Models



# History of Human Pose Recognition

Classical vs. Deep Learning-based approaches



# Challenges



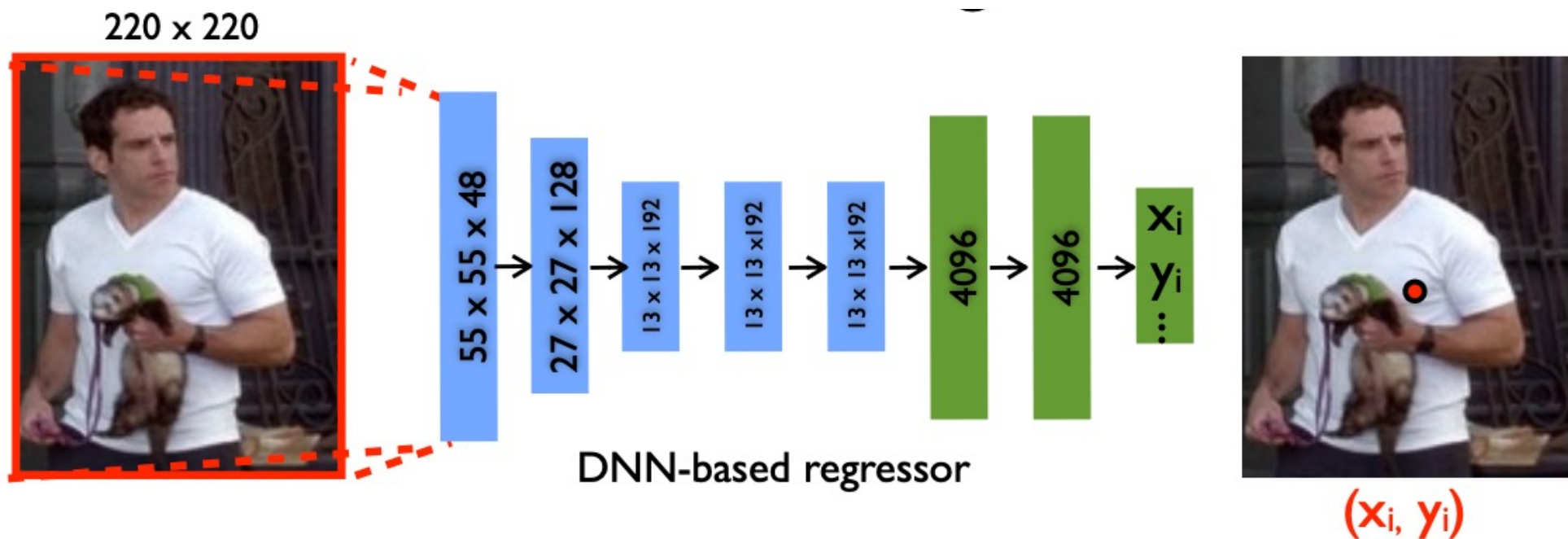


---

# Challenges

- Body's appearance joins changes dynamically due to
  - diverse forms of clothes,
  - arbitrary occlusion,
  - occlusions due to viewing angles,
  - background contexts
- Pose estimation needs to be robust to challenging real-world variations such as lighting and weather.

- **2014, Toshev et al, DeepPose: Human Pose Estimation via Deep Neural Networks**
  - Defined the whole problem as a CNN-based regression problem towards body joints



- DNNs are very proficient in estimating single human pose but when it comes to estimating multi-human they struggle

---

# Top-down vs. Bottom-up methods

- Top-down methods run a person detector first and estimate body joints within the detected bounding boxes.
- Bottom-up methods estimate each body joint first and then group them to form a unique pose.

- **2019**, Zhe Cao et. al. proposed **OpenPose**
  - One of the most popular *bottom-up* approaches for *multi-person* human pose estimation.



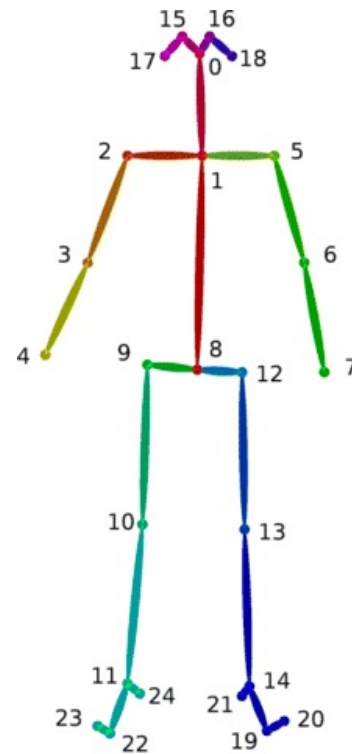


- 2019, Zhe Cao et. al. proposed OpenPose
  - For the first time shown the capability to jointly detect the human body, foot, hand, and facial keypoints on single images
  - Total 135 keypoints: body 25, hand 21x2, facial 70

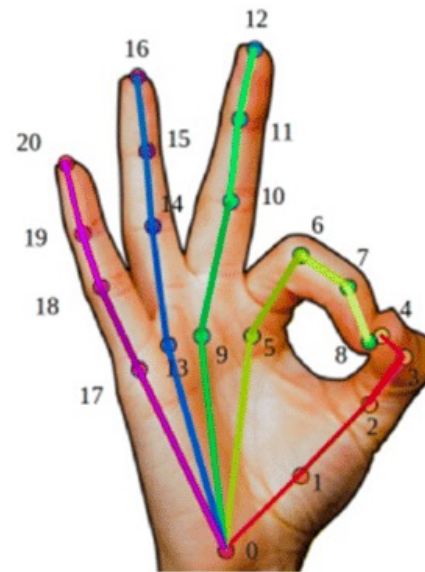
```

{0, "Nose"},
{1, "Neck"},
{2, "RShoulder"},
{3, "RElbow"},
{4, "RWrist"},
{5, "LShoulder"},
{6, "LElbow"},
{7, "LWrist"},
{8, "MidHip"},
{9, "RHip"},
{10, "RKnee"},
{11, "RAnkle"},
{12, "LHip"},
{13, "LKnee"},
{14, "LAnkle"},
{15, "REye"},
{16, "LEye"},
{17, "REar"},
{18, "LEar"},
{19, "LBigToe"},
{20, "LSmallToe"},
{21, "LHeel"},
{22, "RBigToe"},
{23, "RSmallToe"},
{24, "RHeel"},
{25, "Background"}

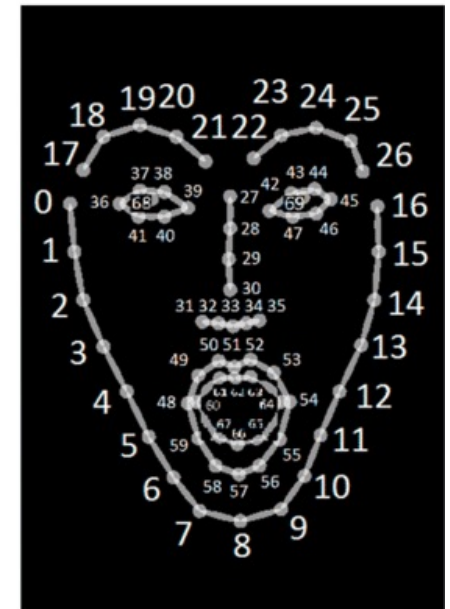
```



a) Body

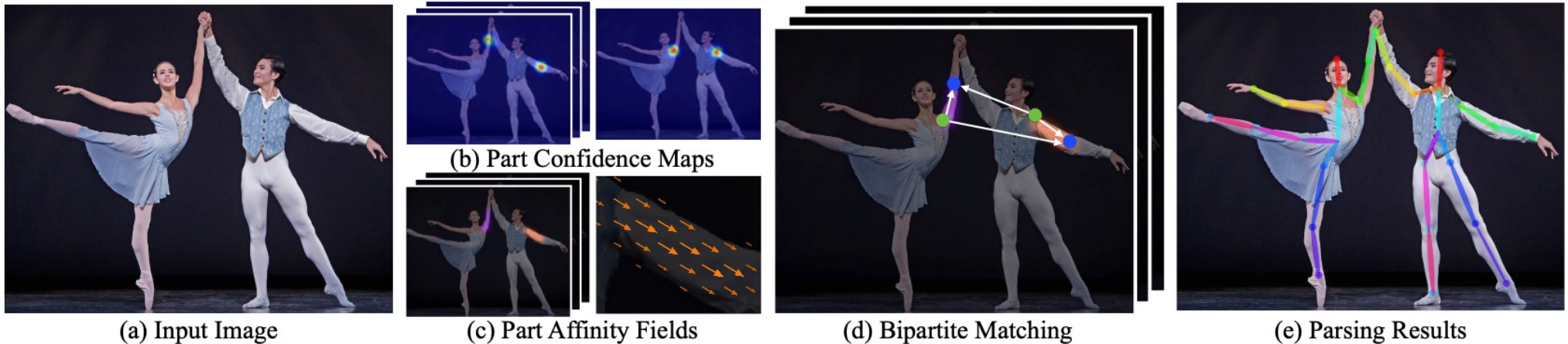


b) Hand



c) Face

# Pipeline of OpenPose, a multi-stage CNN



The OpenPose Pipeline consists of multiple sequential tasks:

1. Acquisition of the entire image as input (image or video frame)
2. Two-branch CNNs jointly predict confidence maps for body part detection
3. Estimate the Part Affinity Fields (PAF) for parts association
4. Set of bipartite matchings to associate body parts candidates
5. Assemble them into full-body poses for all people in the image

# Becoming more accessible



Authors [Ginés Hidalgo](#) (left) and [Hanbyul Joo](#) (right) in front of the [CMU Panoptic Studio](#)



# Becoming more accessible



---

# Opportunities

- Human pose recognition is becoming more ubiquitous across numerous ecologies
- How can we use Human Pose Recognition to translate our intimate bodily movements in a digital environment?
- What elements do we lose during that process?

# Re/defining Creativity

Translate bodily movements into something meaningful

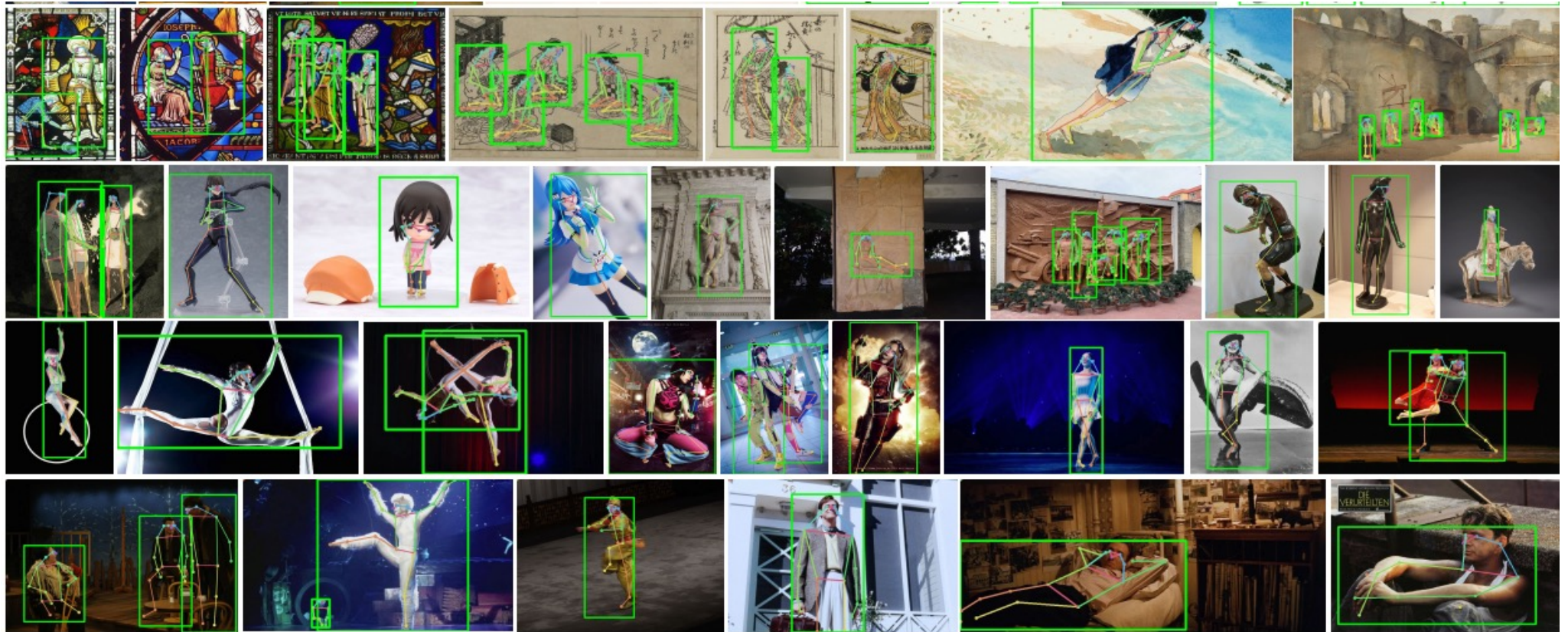
# Human-Art: A Versatile Human-Centric Dataset

## Bridging Natural and Artificial Scenes



**50,000** images including human figures in **20** scenarios  
 (5 natural scenarios, 3 2D artificial scenarios, and 12 2D artificial scenarios)

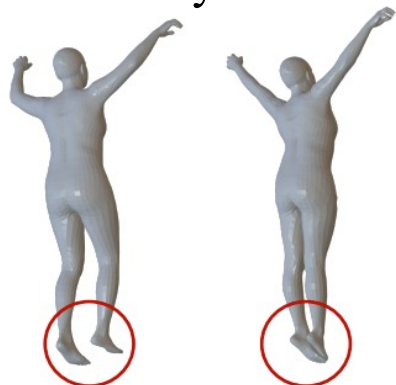




## Human Mesh Recovery

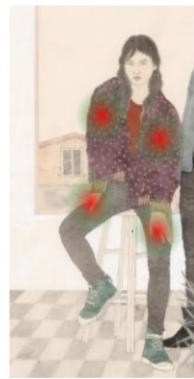


Oil Painting

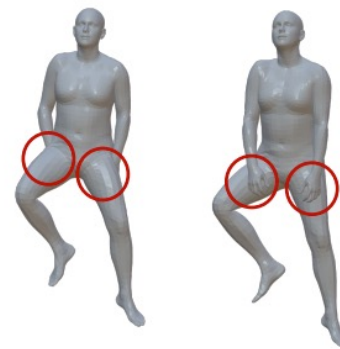


(a)

(b)



Watercolor

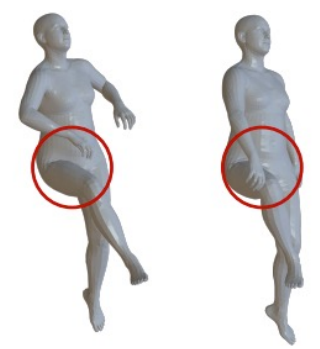


(c)

(d)



Digital Art



(e)

(f)

# MMPose for AIGC







# Animal Kingdom

A Large and Diverse Dataset for Animal Behavior Understanding



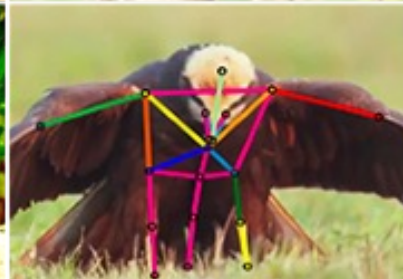
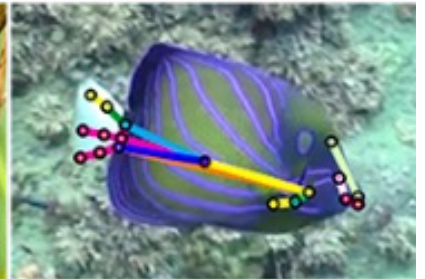
**Mammals**

**Reptiles**

**Amphibians**

**Birds**

**Fishes**

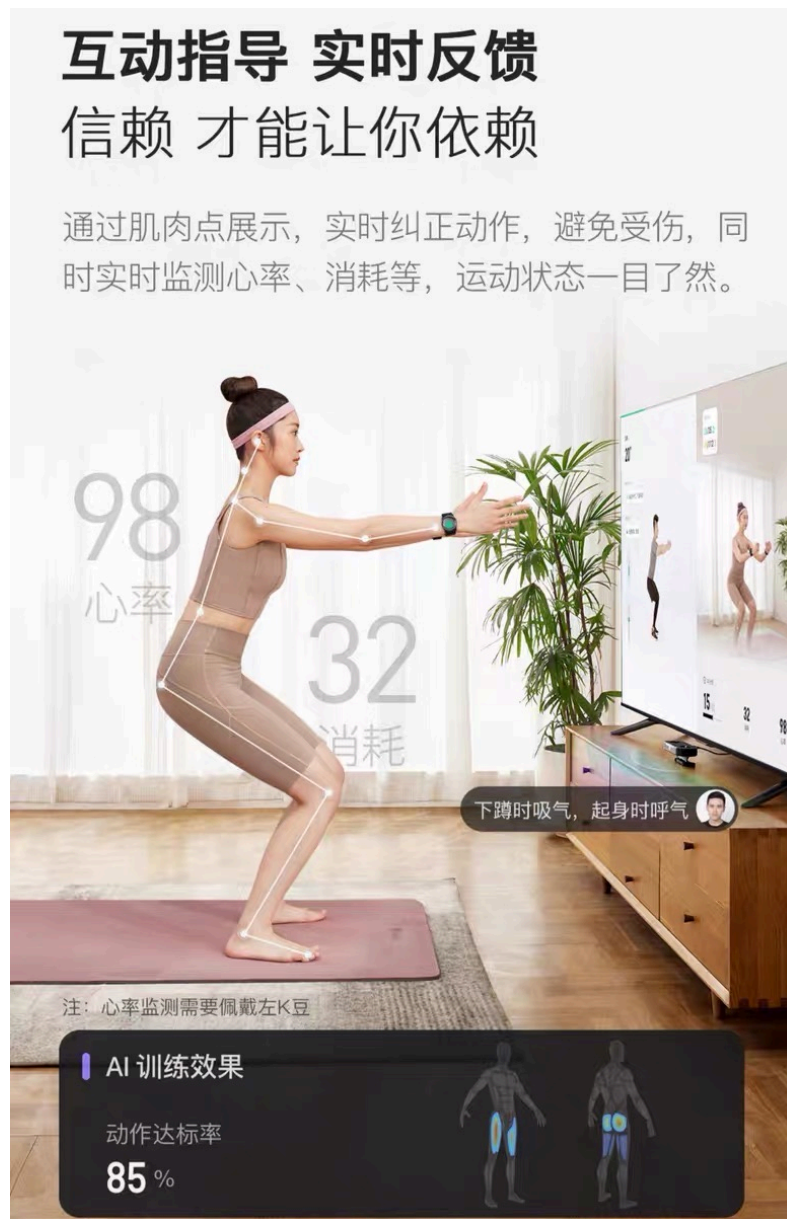






- 50 hours of annotated videos
- 30K video sequences for the fine-grained multi-label action recognition task
- 33K frames for the pose estimation task
- 850 species across 6 major animal classes

# AI-powered personal trainers

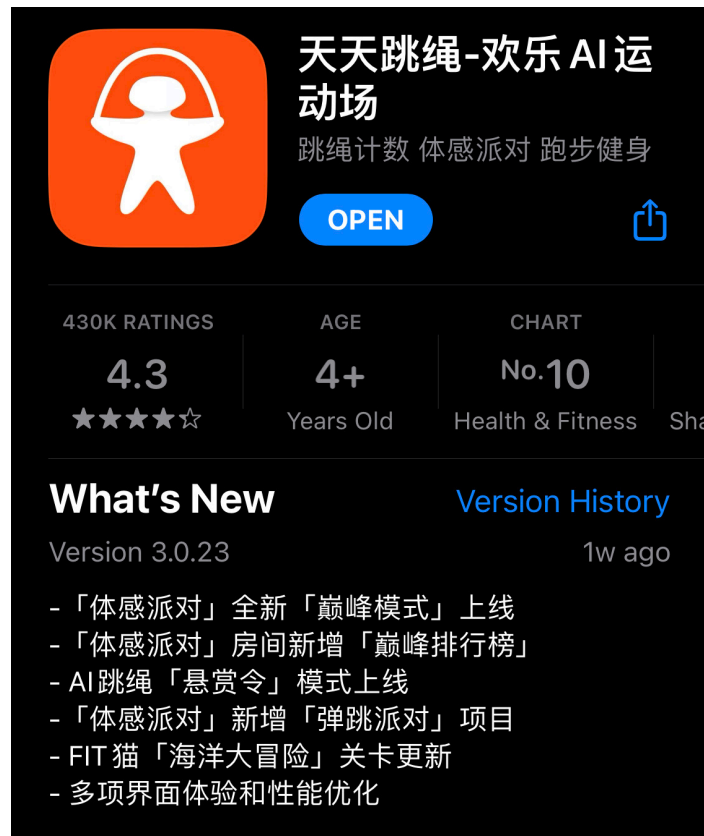


# AI-powered personal trainers





# AI-powered personal trainers



**天天跳绳-欢乐AI运动场**  
跳绳计数 体感派对 跑步健身

OPEN

430K RATINGS  
**4.3**  
★★★★☆

AGE  
**4+**  
Years Old

CHART  
**No.10**  
Health & Fitness

**What's New** [Version History](#)

Version 3.0.23 1w ago

- 「体感派对」全新「巅峰模式」上线
- 「体感派对」房间新增「巅峰排行榜」
- AI跳绳「悬赏令」模式上线
- 「体感派对」新增「弹跳派对」项目
- FIT猫「海洋大冒险」关卡更新
- 多项界面体验和性能优化



# Augmented Reality and Virtual Reality



<https://pose-animator-demo.firebaseio.com/>

# Fall Detection using Pose Estimation

A person is standing on a ledge or roof edge, looking down. The background is a blurred cityscape. The text "50 WAYS TO FALL" is overlaid in large, bold, black letters.

**50 WAYS  
TO  
FALL**

# Fall Detection using Pose Estimation

FPS: 4.426411753097155  
Fall Count: 6

©kevinbparry



Drunk



# Fall Detection using Pose Estimation

FPS: 4.286598757855384  
Fall Count: 4

@kevinbparry



Stubbed Toe

# Fall Detection using Pose Estimation

FPS: 4.378928378285264  
Fall Count: 3

@kevinbparry



Walking Trip

# AI无距 帕友无惧

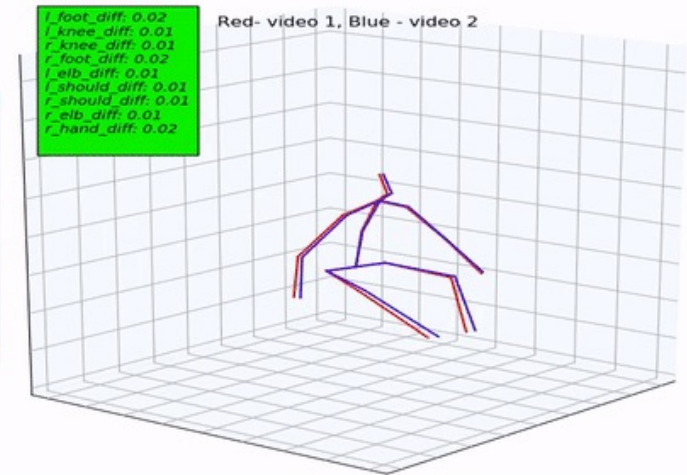
- <https://v.qq.com/x/page/c085944k2yq.html>

# Proceed with Caution

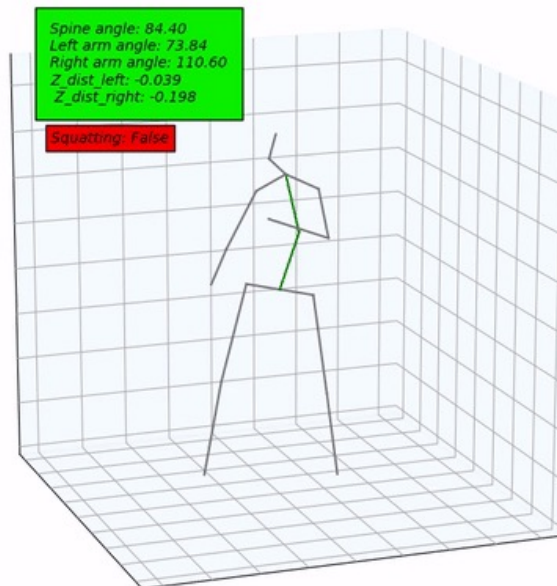
When talking about anything AI x ML related, we must be very careful in how we plan and develop such systems and tools. We should also consider their wider influence and impact on various ecologies that might extend outside the more typical use cases for such technologies.



# Technology limitations



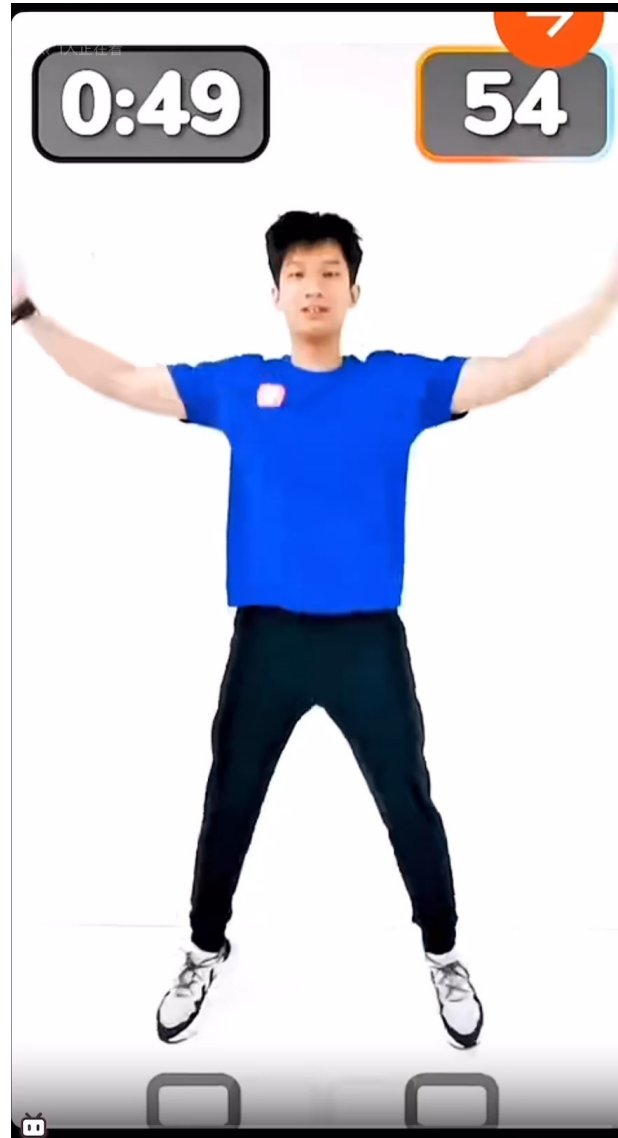
Frontal View Error



- Body diversity, Age Diversity
- Occlusion
- Don't explicitly model depth
- Fail to consider contact

Quick Movements of the Lower Body Part

# 天天跳繩作弊

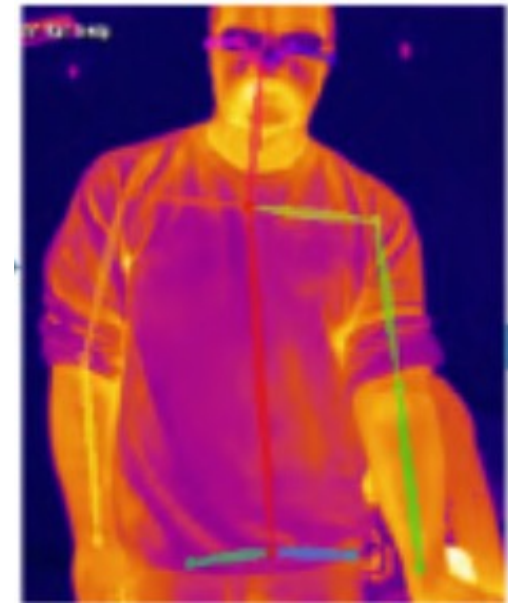


# Privacy concerns

- With every corner of urban landscapes infested with cameras that collect, mine and process information about individuals and their interactions with their environment, for opaque agendas.
- Could human pose estimation be used as a method to preserve privacy, preventing ethnic bias and reducing complexity?
  - reconstructed skeletal structure VS facial features
  - thermal and depth camera



Cormier, et al, 2022



Baselizadeh, et al, 2023

---

# Practice

- PoseNet:
  - <https://mmla.gse.harvard.edu/tools/posenet/>
- OpenPifPaf:
  - <https://github.com/openpifpaf/openpifpafwebdemo>
  - [https://github.com/cwlroda/falldetection\\_openpifpaf](https://github.com/cwlroda/falldetection_openpifpaf)
- Openpose
  - <https://github.com/CMU-Perceptual-Computing-Lab/openpose>
- MMPose:
  - <https://github.com/open-mmlab/mmpose/tree/main>
- ml5.js - Friendly Machine Learning for the Web:
  - <https://examples.ml5js.org/>
- <https://www.tensorflow.org/js/demos>
- <https://pose-animator-demo.firebaseio.com/>





# DS323: AI in Design (AIID)

<https://ds323.ancorasir.com/>

Autumn 2023

**Thank you~**

Wan Fang  
Southern University of Science and Technology