Human Pose Estimation

COSC 444/544 Project

Team 3
Hamza Muhammad Anwar
Bridgette Hunt
Jeena Javahar
Aadil Shaji

Introduction

- Highlight the locations and shapes of the human body
- Pose = configuration of joints and body parts
- Approximate position of body parts and joints in relation to the rest of the body

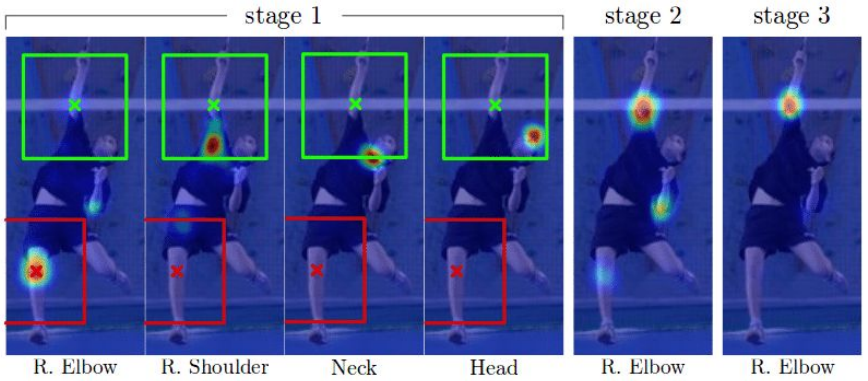
(Gupta, 2018)
Image credit: Oliver Sjöström

Application

- Rock climbing
- Complex poses
- Easier to teach techniques

Human Pose Estimation Techniques

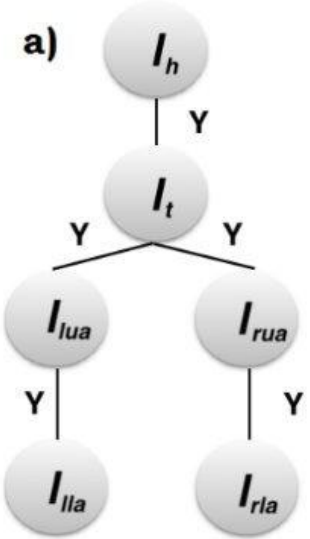
- Deep Learning
  - Heatmap-based
- Traditional
  - Flexible Mixtures of Parts
  - Pictorial Structures

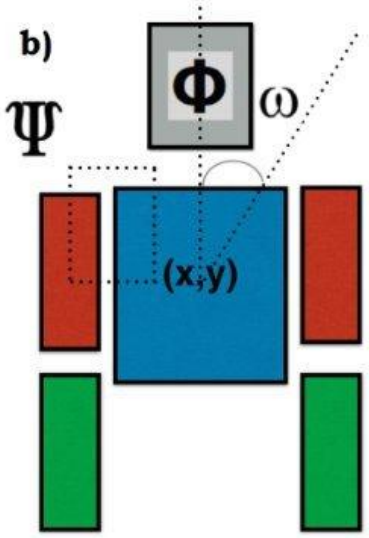


Hartmann, Michael & Massimiani, Gianluca. (2018). State of the Art: Safety for pedestrians and autonomous vehicles (Technical Report). 10.13140/RG.2.2.11639.27043.

Pictorial Structures

- Objects as a probabilistic graphical model of edges and vertices
- In HPE: represents body as a collection of body parts (vertices) and joints (edges)



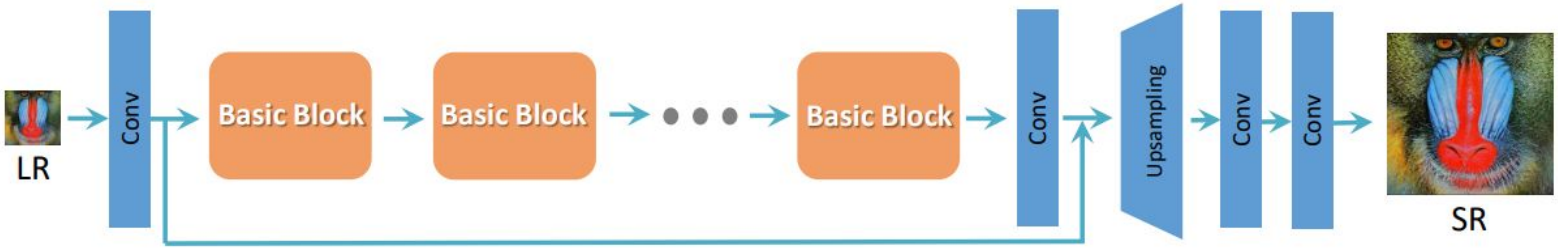


(Penmetsa et al., 2014)

Proposed Pipeline

ESRGAN (Enhanced Super-Resolution Generative Adversarial Networks)

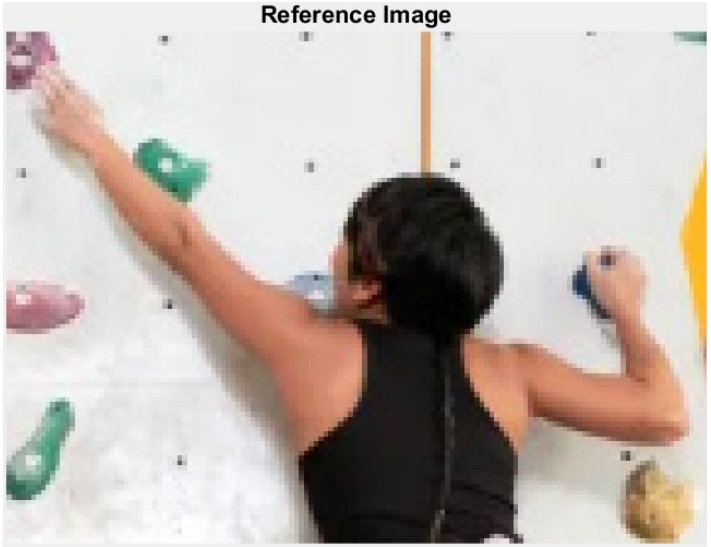
- Pre-processing
- Super Resolution
- Enhances resolution of low-quality images



(Wang et al., 2018)

ESRGAN

- Slight improvement



SIFT Clustering

- Finds regions of interest
- Creates clusters of SIFT features
- Selects largest cluster of features
- Bounding box limits region for pose estimation

Pose Estimator

- Pictorial Structures
- Pre-Defined Human Appearance Models
- Image Parsing for Human (Location & Orientation)

(Penmetsa et al., 2014)

Results (SIFT Bounding Box)

- Unsatisfactory
- Unable to properly recognize pose

Results (Manual Bounding Box)

- Performs better with manual bounding box

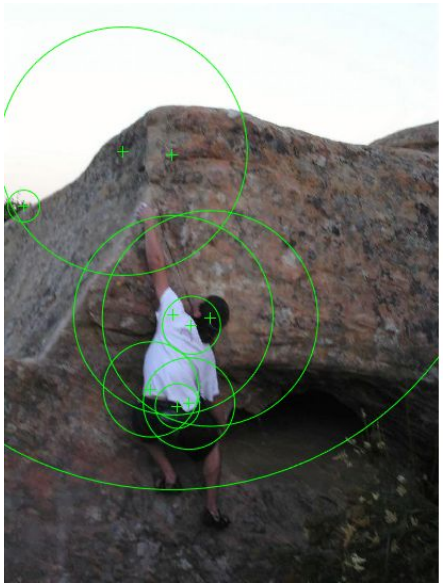
Results (Calvin Upper-body Detector)

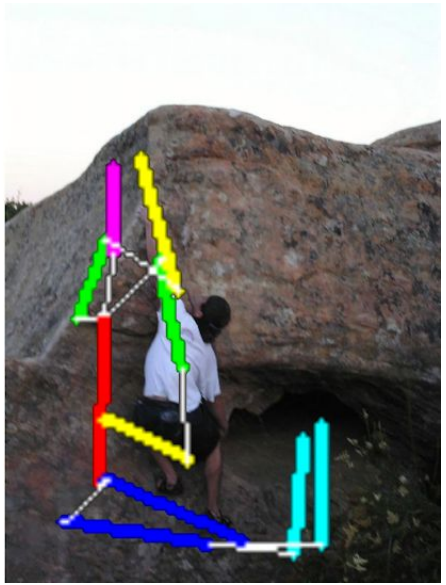
Unable to find a body at all!

(Due to reliance on facial detection to find the upper body)

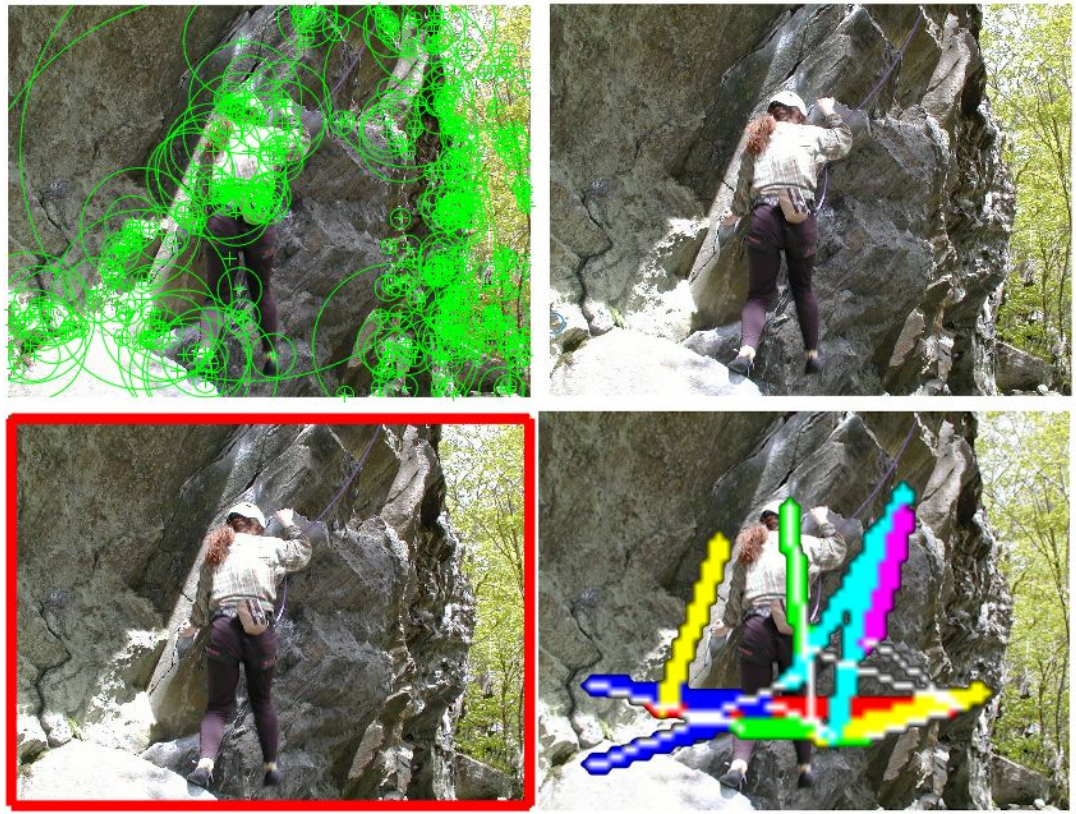
Results

Way too small!





Results



Way too big!

Live Demo

Analysis

- **The performance of pose estimation is highly contingent on the input image and the appearance model matching each other**

- Available models were meant for images with forward-facing and upright people
- This didn’t match our image domain well

Analysis

- **The SIFT feature detection created bounding boxes that did not translate to the expected bounding boxes for the pose estimator**

- The pose estimator expects a bounding box surrounding the head and torso
- Our bounding boxes, even in an ideal situation, didn't capture the right parts