# Application of Human Pose Estimation: *A Case Study using Salsa Dance Form*

Bhanu Yerra
Metis Data Science Bootcamp, Project #5

## Introduction

Human pose estimation is a process of estimating anatomical "key points" of people in images or videos. In the last few years, a number of researchers have made incredible progress and demonstrated real-time performance[1, 2]. Some applications of human pose estimation include gesture recognition, augmented reality, animation, gaming, surveillance, sports and human assistance.

For my project #5, I would like to study the application of one such estimation method's ability to learn, detect and/or generate dance movements from watching Salsa videos on YouTube. What challenges would an exercise like this involves? Potential applications might be focused in areas where anticipating human movements and a system's readiness to react to those potential movements improves performance and speed.

## Approach

Zhe Cao et al, from The Robotics Institute of Carnegie Mellon University have won inaugural COCO 2016 keypoint challenge using a "bottom-up" approach. They have achieved real-time speeds using GPU installed laptops. My approach involved using their algorithm to collect human poses at 8 frames per second speed from YouTube salsa videos.
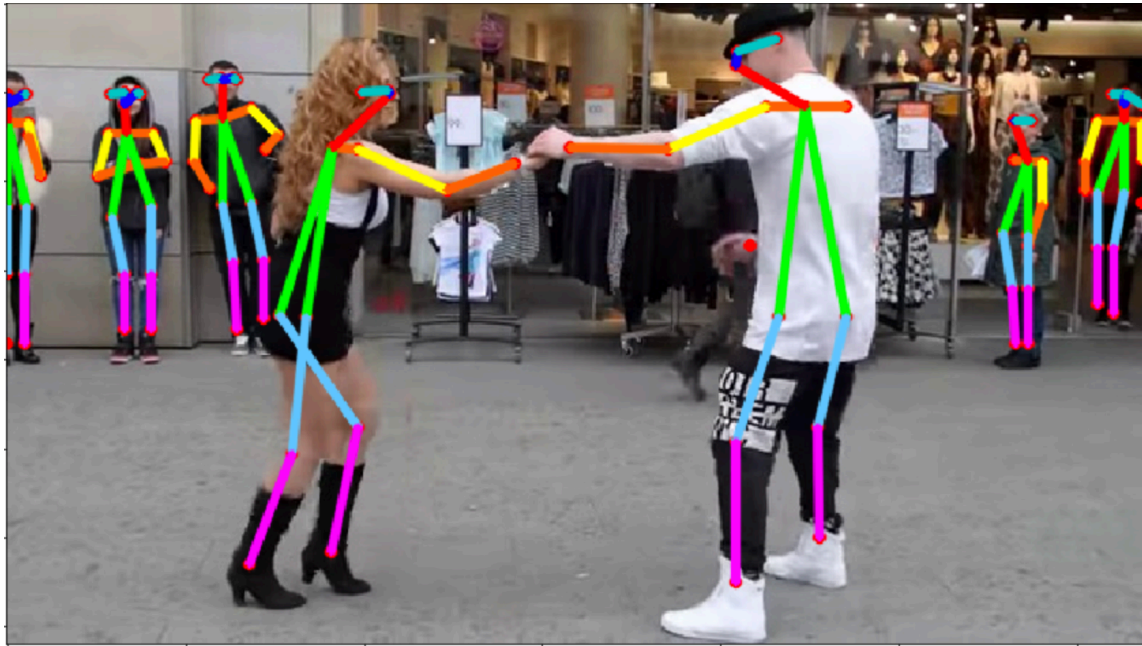
Why Salsa?

Besides being a challenging dance form that involves subtle and quick movements, a number of videos on YouTube provides easy access to videos with solo dancers. The Zhe Cao et al's pose estimation is quick and accurate with multi-person video frames, but those algorithms do not track a person from frame to frame. This poses a problem when persons in a video switch places. To avoid this and simplify my problem, I've selected solo dance videos.

The other dance form that I have considered is ballet. However, the female dancers clothing in ballet results in false positives for certain dance positions and will pose a challenge for my task.

Data collected is about 90 minutes from World Salsa Summit's videos on YoutTube and includes 58 performances by about 30 artists:

- https://www.youtube.com/watch?v=4nHElVbT3HY
- https://www.youtube.com/watch?v=ITNiqNcl6Mw
- https://www.youtube.com/watch?v=L5mqL7ADEsY
- https://www.youtube.com/watch?v=M_rPhEjym1o



## Additional Considerations

Given this is a risky project that might take longer than time allotted for project #5, an alternative is to classify Salsa dance form based on male and female dance sequences. Additionally, Salsa vs. another dance form can be evaluated.

## References:

1. Zhe Cao et al, Realtime Multi-person 2D Pose Estimation using Part Affinity Fields. https://arxiv.org/abs/1611.08050
   Github: https://github.com/ZheC/Realtime_Multi-Person_Pose_Estimation
2. A 2019 Guide to Human Pose Estimation, Medium Article, https://heartbeat.fritz.ai/a-2019-guide-to-human-pose-estimation-c10b79b64b73