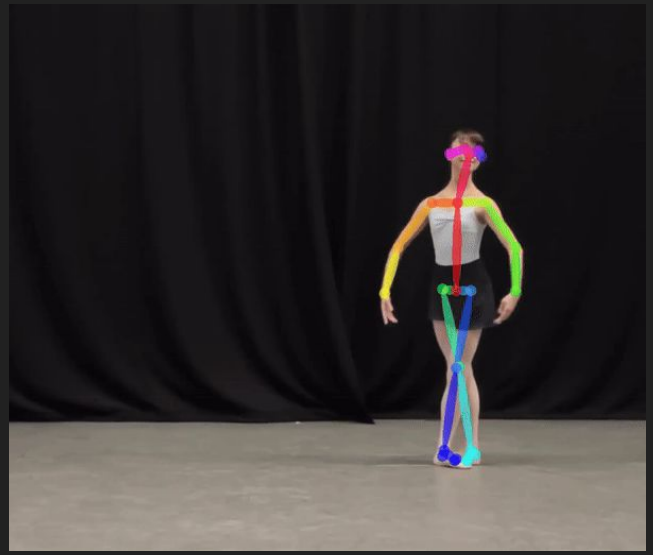
Pose Estimation & Dance Move Classification

Caroline Zeng, Elisse Chow

Problem/Motivation

Bottom-up pose estimation algorithm as a method to classify dance moves

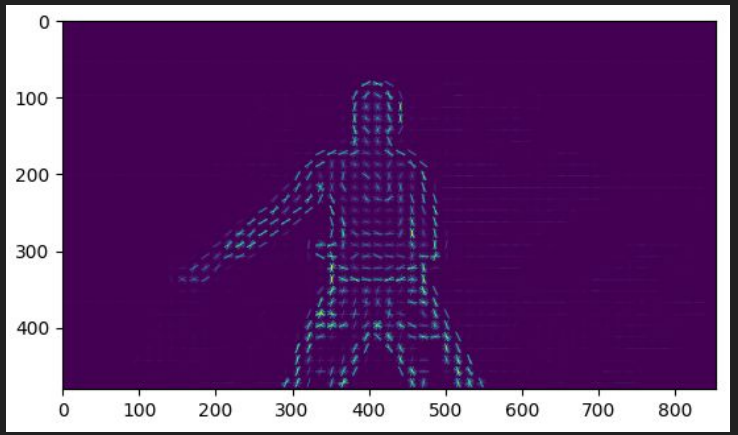


Methodology

- Pose Estimation Model
  - Trained on the MPII Human Pose dataset
  - Input: Feature vector from histogram of oriented gradients
  - Output: Estimated 16 joints that form the human pose skeleton
- Dance Move Classification Model
  - Trained on a dataset of images we collected and labeled with specific ballet movements (grand jeté, arabesque, and passé)
  - Input: Joint coordinates from our Pose Estimation Model (later OpenPose)
  - Output: Move classification dependent on the learned positions/relationships between the joints

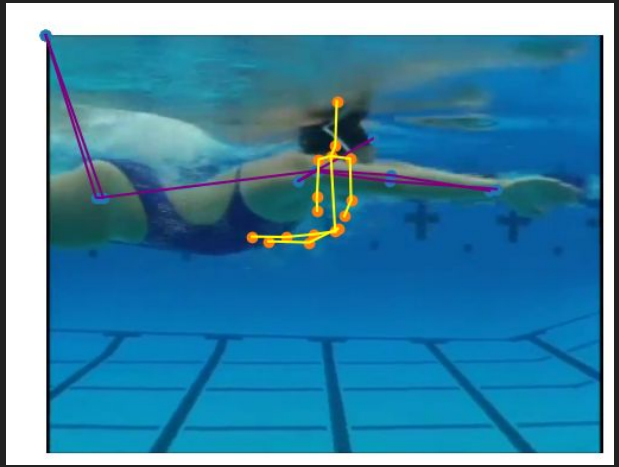
Pose Estimation: Histogram of Oriented Gradients

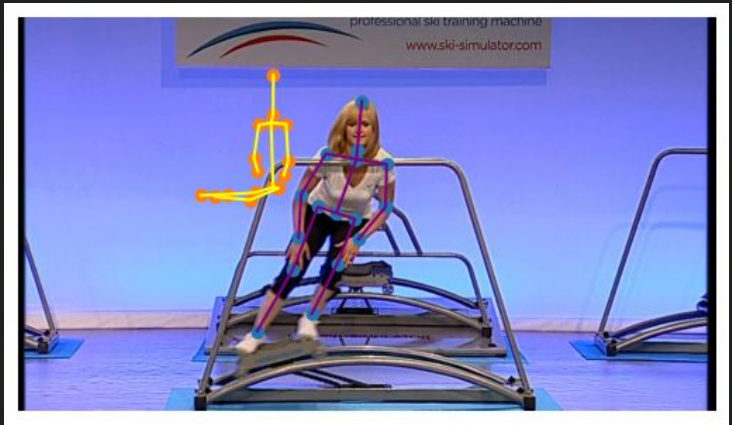
- Sci-kit image hog - captures contour, silhouette and some texture information
- Ravel to feature vector



Pose Estimation: Challenges

- Sitting person for all samples
- Responsive to location of person but not pose



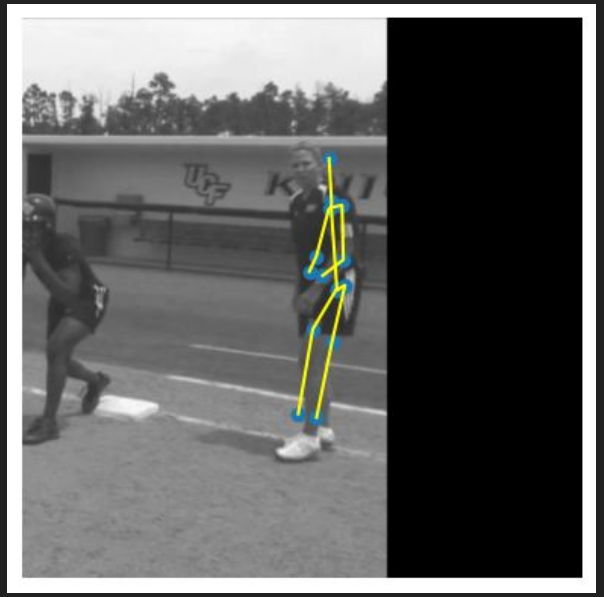


Pose Estimation: Small Success

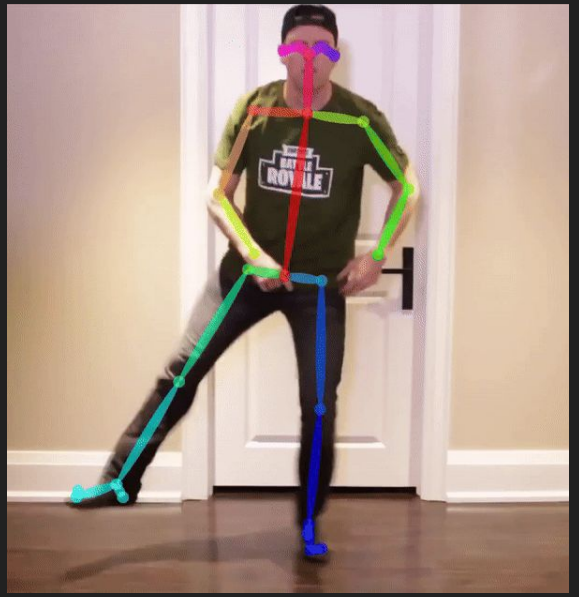
- Slightly improved, but overall still performing poorly, especially loss
  - Loss: 45k training, 80k validation
- General stance of body identified often, imprecise joint locations

Change of Plans: OpenPose Model

- Using a real-time multi-person pose detection library moving forward that gives us far better results



vs.



Dance Move Classification: Dataset/Model

Grand jeté

Arabesque

Passé

- 2397 frames, pulled from ballet videos on Youtube (10 fps)
- Annotated with label of move being performed
  - 2079 none
    - Downsized to 20% (~200)
  - 103 grand jeté
  - 109 arabesque
  - 106 passé
- Pose skeleton obtained for every image
- Model trained on joint locations to classify action

Dance Move Classification: Results

- Classifying moves with reasonable accuracy
  - Accuracy: 0.6 validation, 0.6 test
  - Loss: 25 validation, 25 test
- Ran predictions on sequenced videos of us and our friends dancing to see if model could correctly identify moves

Dance Move Classification: Analysis

- Passé often represented in dataset during turns - skeleton features common with turns (arms up, vertical body alignment) potentially falsely associated to passé
- Higher representation of “None” in dataset - accuracy boosted by identification of no move happening
- Obscured joints, motion blur - incorrect skeleton

True: None, Guess: Passe

Guess: None, True: Arabesque

Discussion/Future Work

- Pose estimation model accuracy
- Smoothing
- Incorporating more motion
  - Temporal-based moves - grooves
  - Sliding window
- Incorporating more moves
- Differentiating between different dance styles