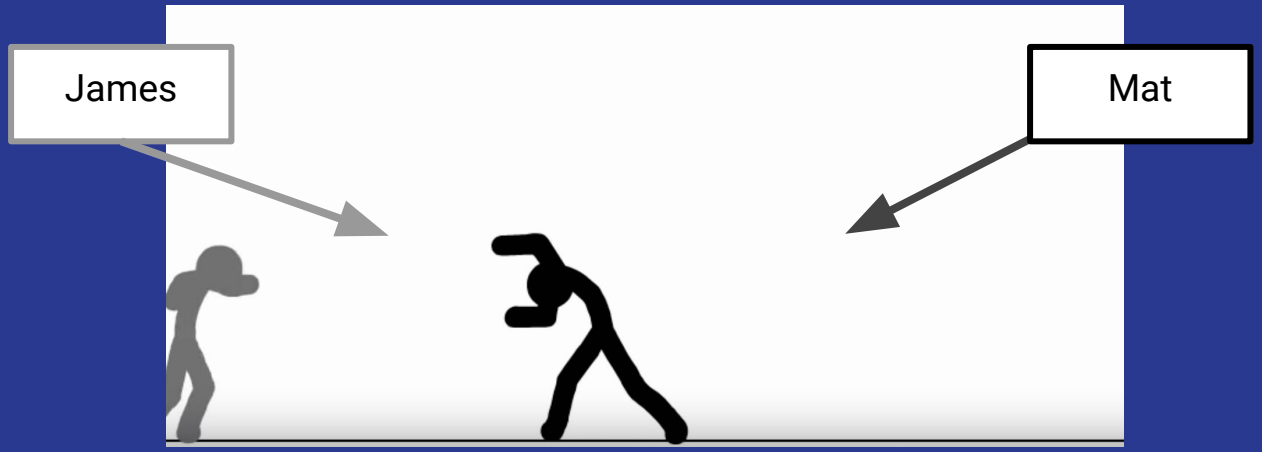
Stick Figure Dance Generation



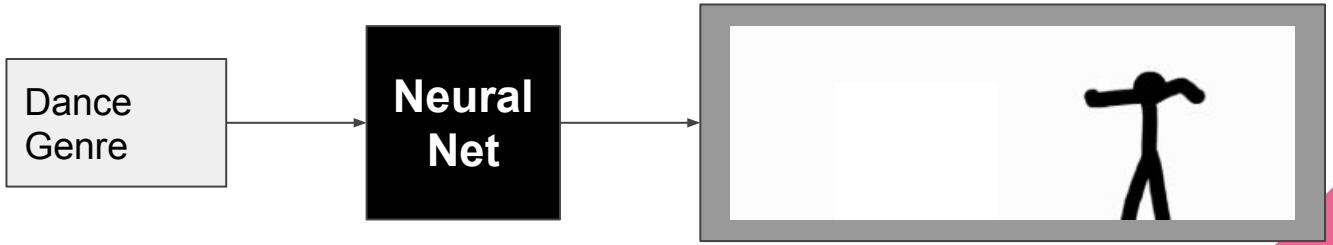
Mathew Varughese
James Wood

What?

What is this?

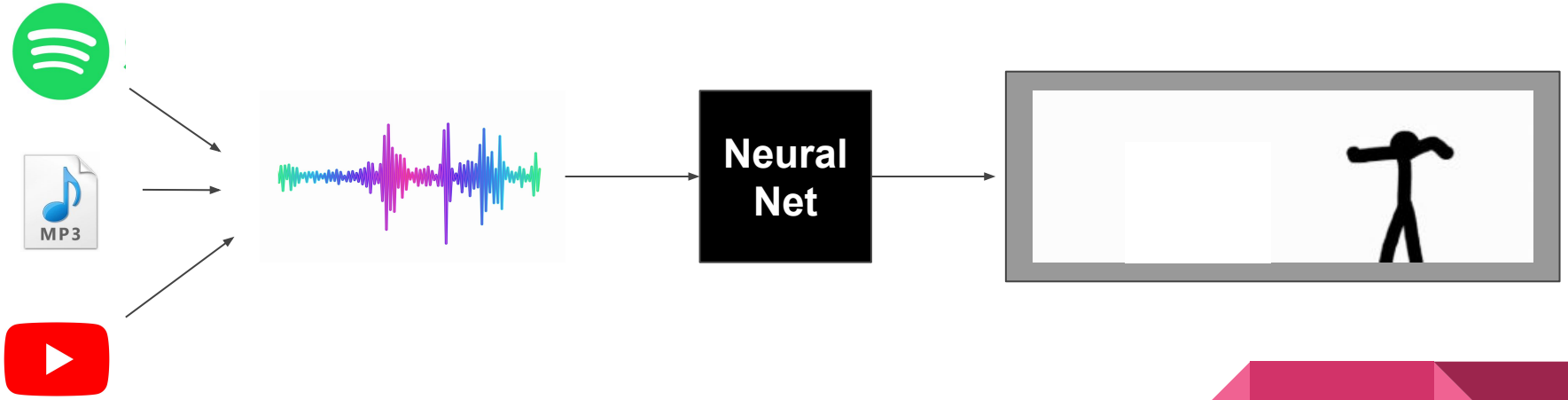
- Generate stick figures dancing using Neural Networks
- Give an input of a “dance” genre

Ballet	Swing	Break Dancing	Square
Flamenco	Foxtrot	Tango	Waltz
Latin	Quickstep		



Further Reach Goal

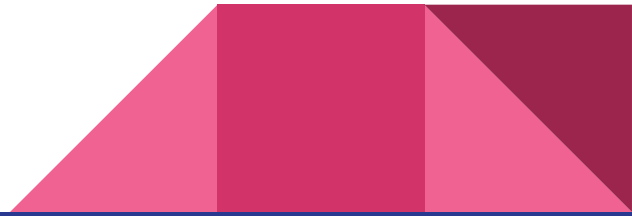
- Be able to listen to beat from any song and generate dance movements to follow that



Why?

Why?

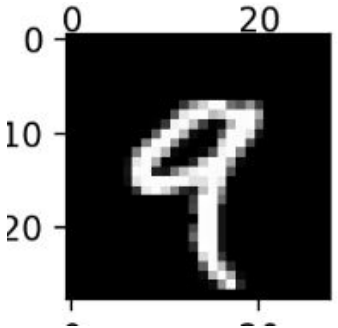
- Mimic human motion?
- Make AI more realistic
- Video + Audio

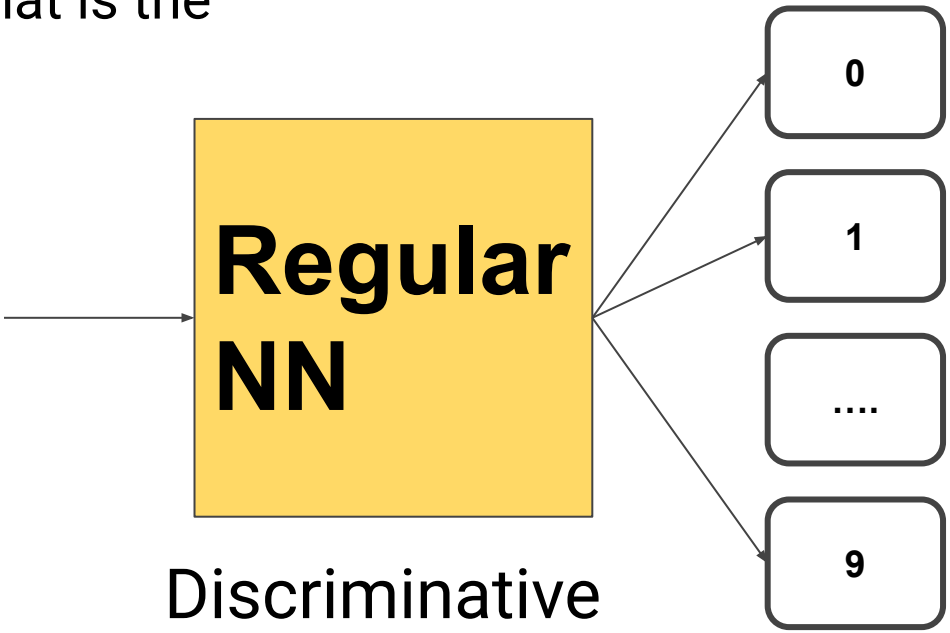


How can a Neural Nets *generate* data?

Discriminative Network

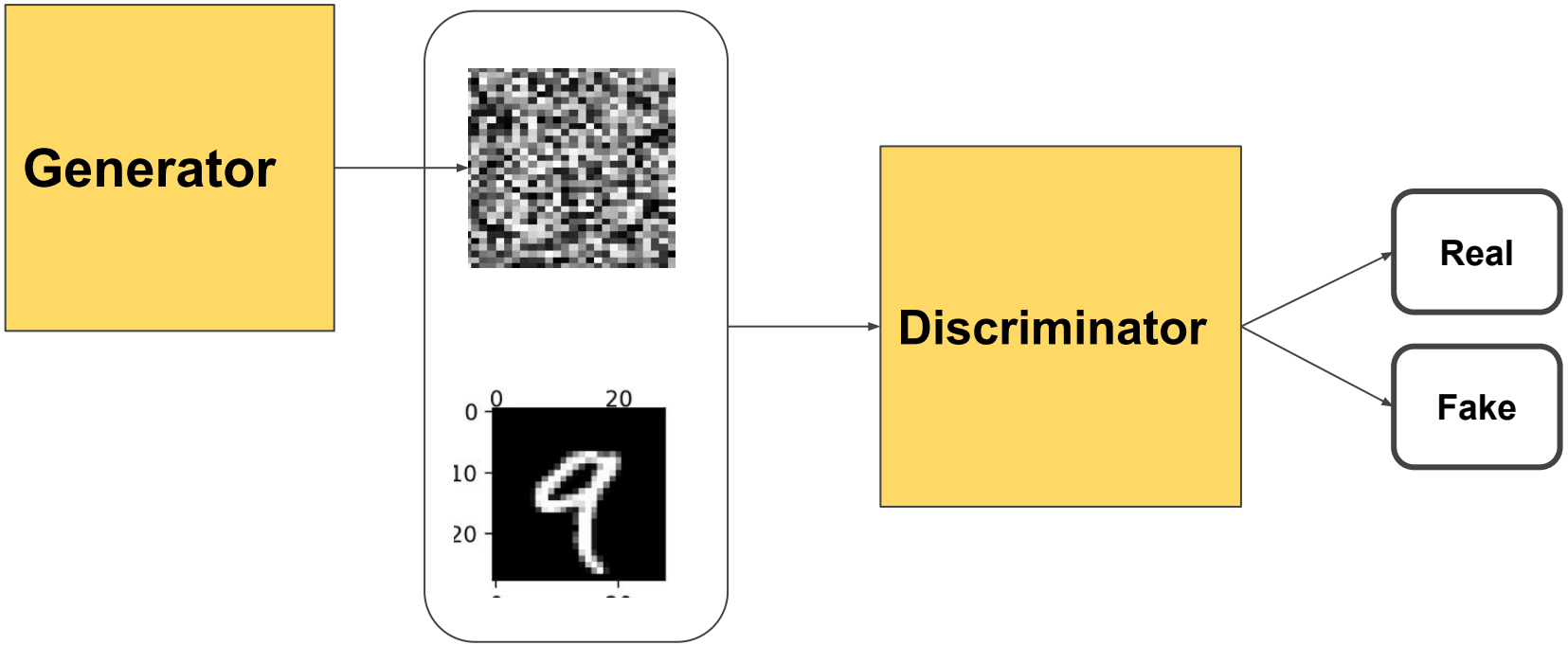
Given these pixels, what is the digit?



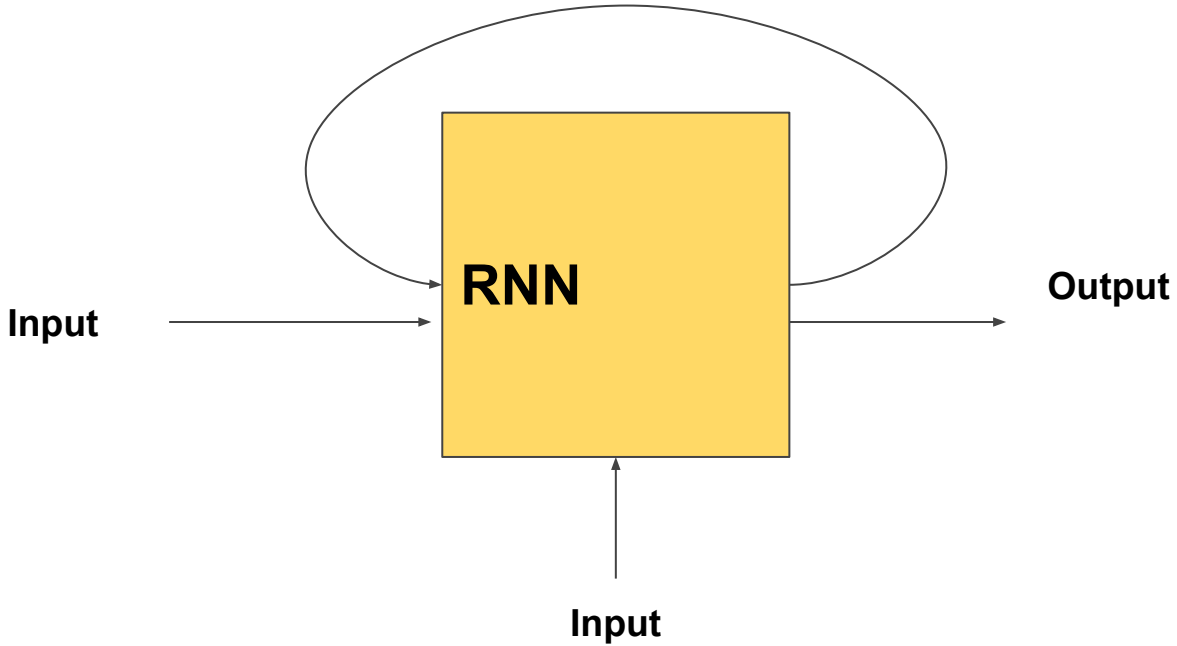


Discriminative

GAN - Generative Adversarial Network

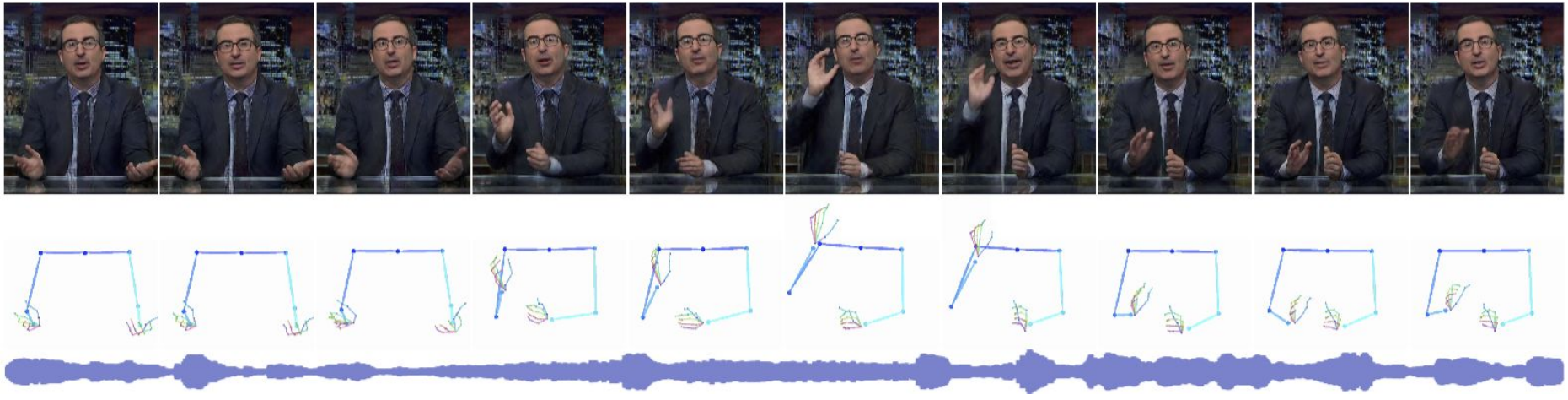


RNN - Recurrent Neural Network

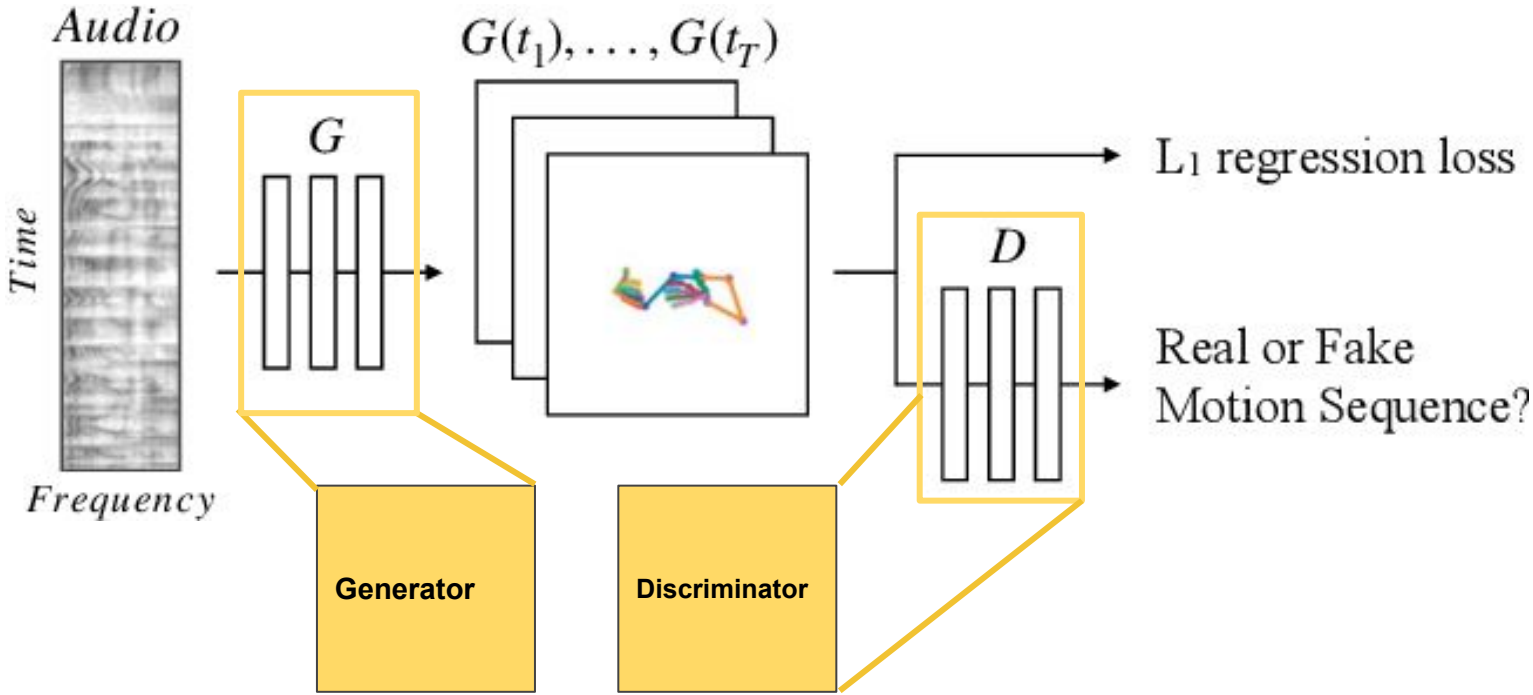


Prior Work

Learning Individual Styles of Conversational Gestures

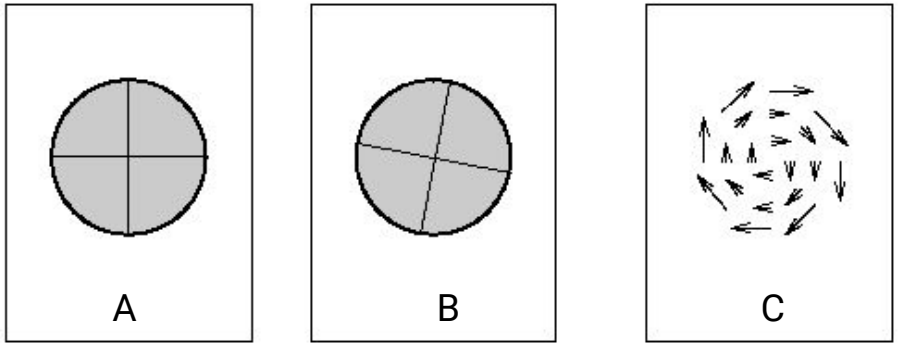


Learning Individual Styles of Conversational Gestures



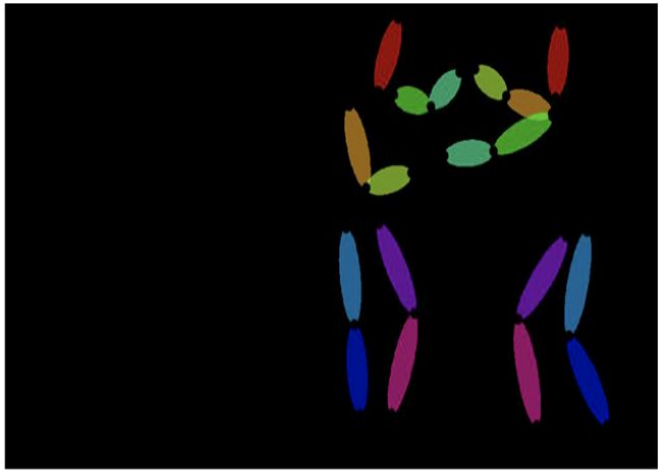
Let’s Dance: Learning from Online Dance Videos

Optical Flow

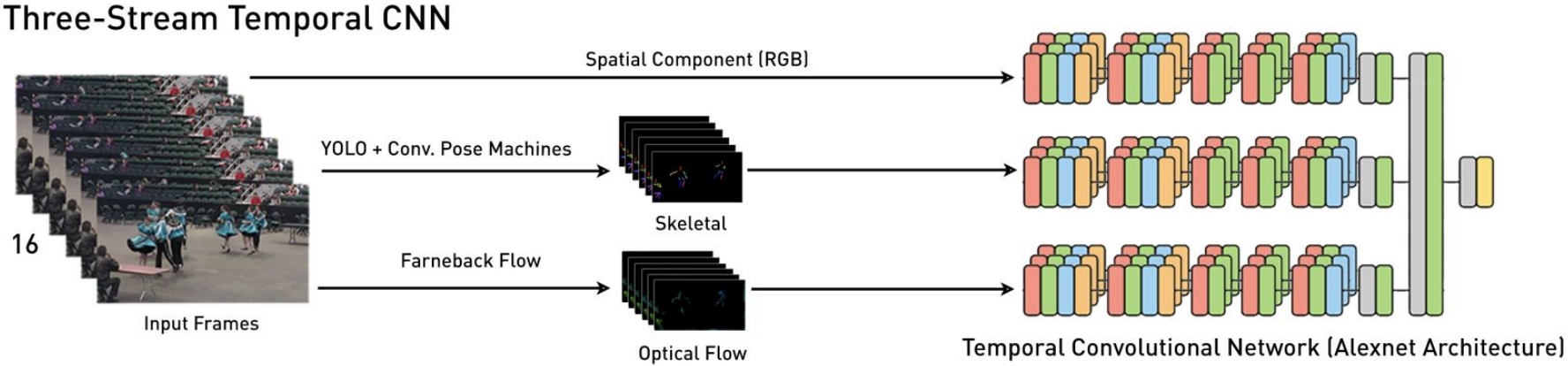


Let’s Dance: Learning from Online Dance Videos

Skeletal Pose Detection

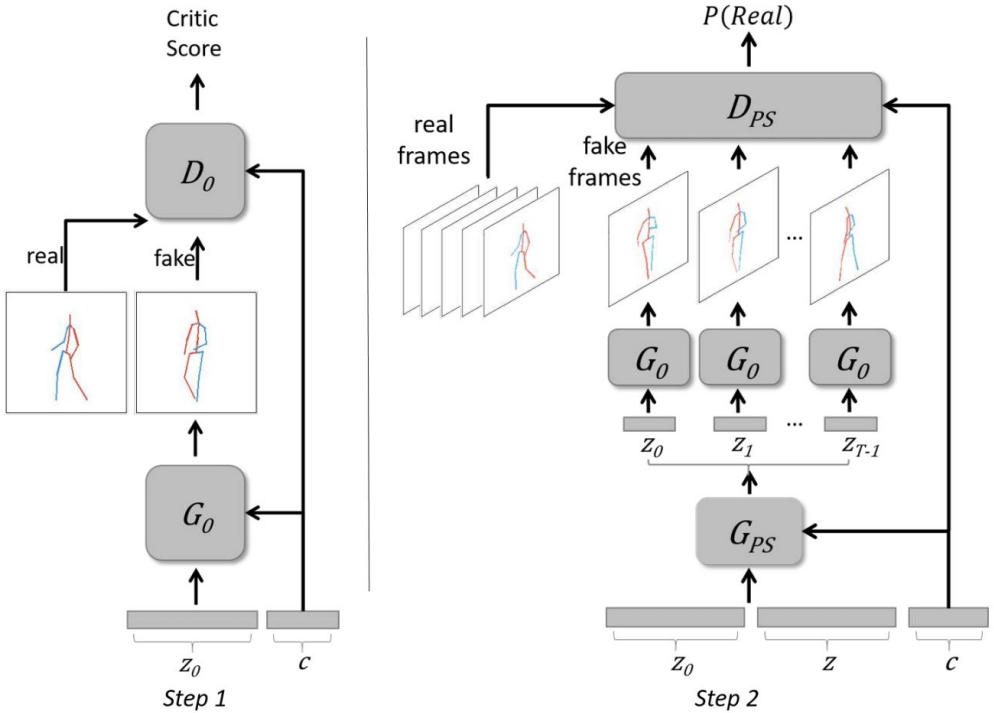


Let's Dance: Learning from Online Dance Videos



Approach	Testing Accuracy
Frame-by Frame CNN	56.4%
Two-Stream CNN	68.89%
Temporal 3D CNN (RGB)	70.11%
Temporal 3D CNN (Skeletal)	57.14%
Three-Stream CNN	69.20%
Temporal Three-Stream CNN	**71.60%**

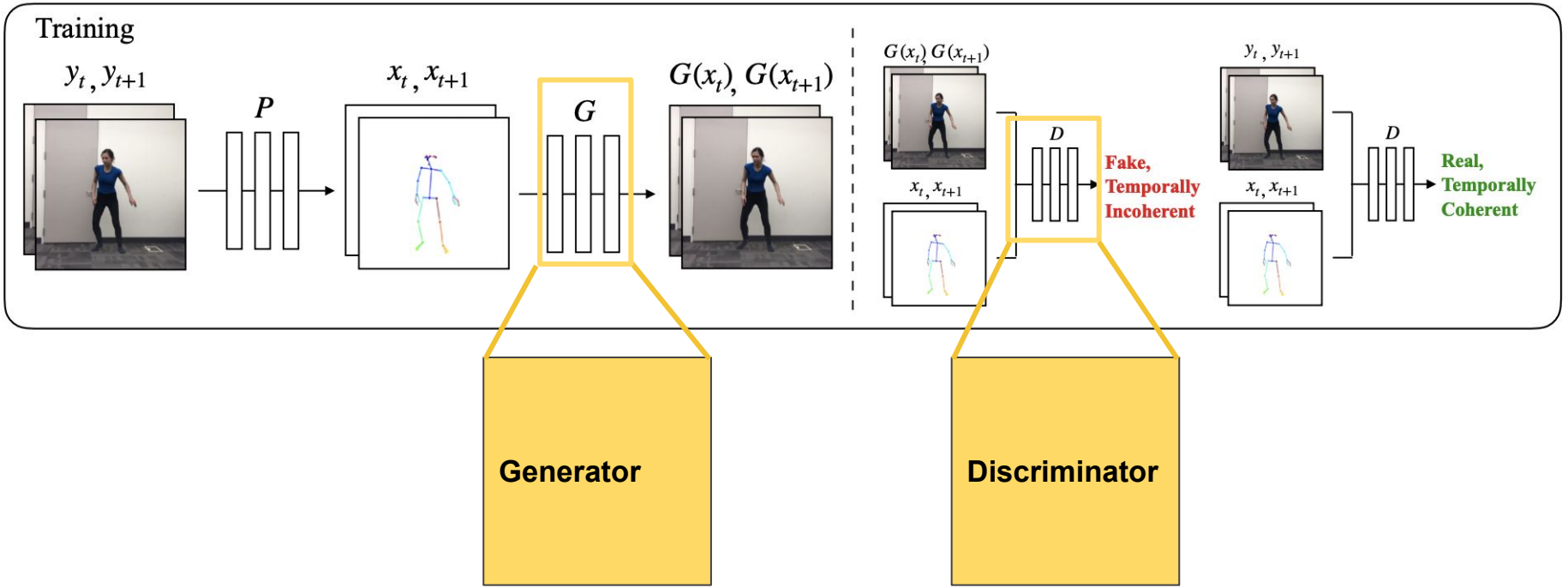
Deep Video Generation, Prediction and Completion of Human Action Sequences



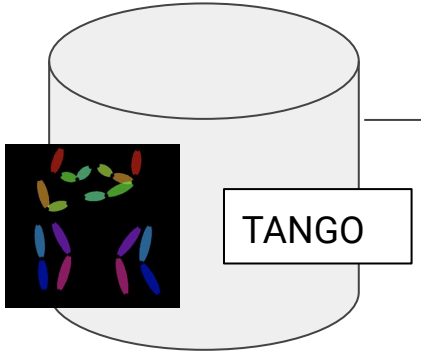
Everybody Dance Now

“Deepfakes for Dancing”

Everybody Dance Now



Approach



Let's Dance Skeleton's Dataset

Multiple Frames of Joint Data

C-RNN-GAN

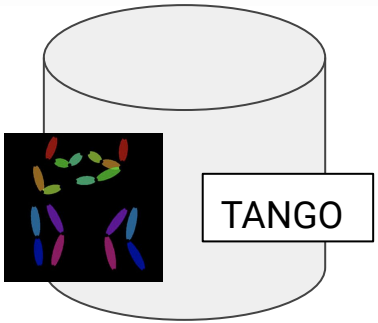
Other Video Datasets (YouTube, Everybody Dance Now)

Pose Detection Algorithms

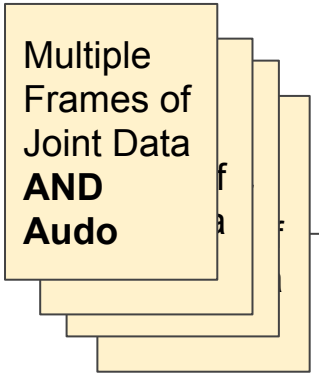
(If needed)

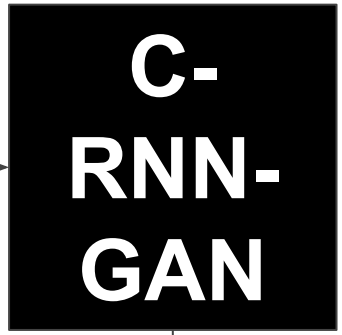
Stretch Goal

Data Augmentation - overlay music onto dataset



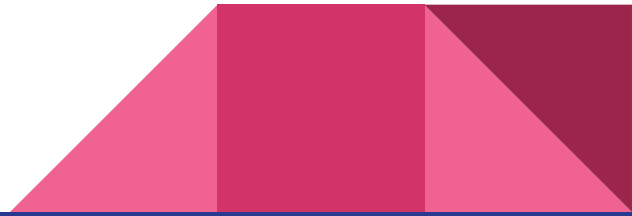
Let's Dance Skeleton's Dataset





How is this different?

- Exploring the use of a C-RNN-GAN
- Music “genre” as input is another dimension of complexity
- By incorporate audio we are doing something much more difficult than the prior work



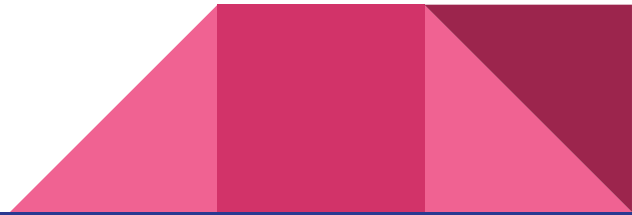
Challenges

1. Learning more about RNNs, GANs, LTSMs

- As new students of Deep Learning, we are excited (scared) to learn new neural networks
- Movement-based physics concepts (optical flow)

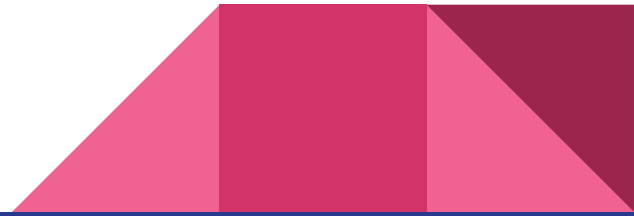
2. Evaluation Metrics for GAN

- How do we assess if a generated dance is “good”? Is manually the best option?
  - Average Log-likelihood
  - Coverage Metric
  - Inception Score (IS)
  - Modified Inception Score (m-IS)
  - Mode Score
  - AM Score
  - Frechet Inception Distance (FID)
  - Maximum Mean Discrepancy (MMD)
  - The Wasserstein Critic
  - Birthday Paradox Test



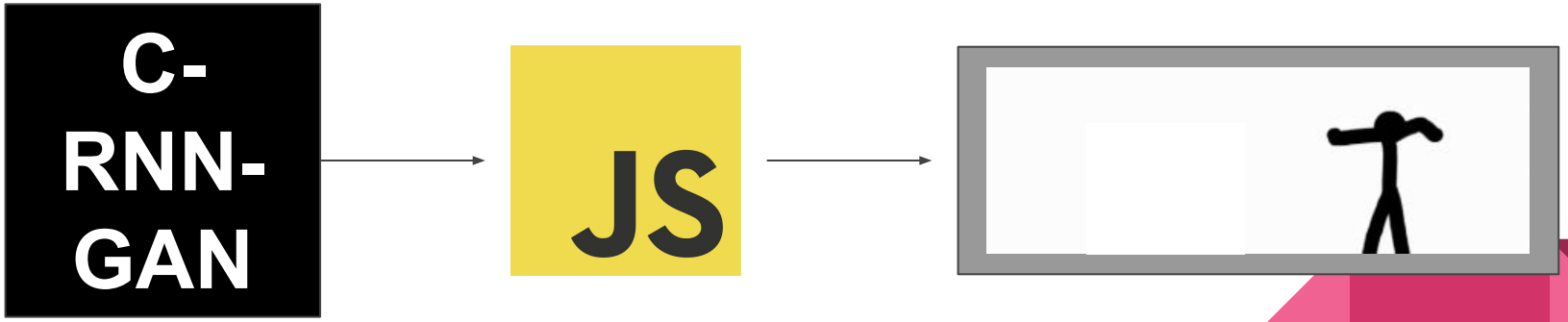
3. Combining Audio and Video

- What is the best approach to use both audio and video into our network to make it generate dance moves based on beats?



4. Visualizing Dances

- Creating an application that can turn output data of our model into visual stick figures



Thank you! Questions? Suggestions?