CS298 Proposal

Video Synthesis From a Single Given Frames

Lei Zhang (lei.zhang01@sjsu.edu)

Advisor: Dr. Chris Pollett

Committee Members: Philip Heller, Leonard Wesley

Abstract:

Generative Adversarial Networks (GANs) have shown impressive results in generating synthetic images. However, video generation is still hard to do even for these neural networks due to the additional time dimension. I am developing a framework to generate a video by searching through the latent space of GANs that have been trained on video frames. In contrast to other papers on the video generation, this framework does not use 3D convolution GAN or 2D+1D GAN to generate videos directly. Instead, it uses two stages. First, it generates a GAN latent space that trained from video frames. Then, it produces a video by searching the GAN latent space from the first step. A random image when specified provides a unique starting point in the latent space to generate videos.

CS297 Results

- Implemented a simple GAN with Python and use it to generate Chinese character digits
- Implemented a video GAN to create fake videos
- Created videos with pix2pix
- Created a GAN-based framework to generate fake videos with input from a single picture

Proposed Schedule

Week 1: Jan 28 - Feb 4	First Meeting
Week 2 - 6: Feb 5 - Mar 3	Use DCGAN and paper [1] to generate a 2D latent space of video frames. Write an algorithm to search for latent space to generate short videos.
Week 7 - 10: Mar 4 - Mar 31	Provide a random image, use paper [2] to search the closest mapping to the trained latent space. Use that latent space location to generate a short video.

Week 11 - 13: Apr 1 - Apr 21	Improve the quality of generated videos in terms of creating high resolution and more realistic pictures.
Week 14 - 17: Apr 22 - May 12	Complete the project report and slides for review.

Key Deliverables:

- Software
  - Deliverable 1: Write a program to extract videos to images. Implement a DCGAN and train it with video frames. Write an algorithm to search DCGAN latent space to generate a short video.
  - Deliverable 2: Write a program to map a random image to the latent space above.
  - Deliverable 3: Write a program to generate a video with a random start frame.
- Report
  - CS 298 Report
  - CS 298 Presentation

Innovations and Challenges

- Searching latent space is a relatively new area. I have to explore how to search for continuous frames from the latent space.
- Have to identify a latent vector that mapping to the input image as a starting point. Maybe hard to find the best latent vector to represent a random input image.
- It's not straightforward to find video frames from a GAN latent space. Have to explore the right GAN model to generate a latent space that is suitable for video generation.

References:

[1] T. Karras, et al., "Progressive growing of gans for improved quality, stability, and variation," *arXiv preprint arXiv:1710.10196*, 2017.

[2] Z. C Lipton, and S. Tripathi, "Precise recovery of latent vectors from generative adversarial networks." *arXiv preprint arXiv:1702.04782*, 2017.

[3] P. Isola, et al., "Image-to-image translation with conditional adversarial networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[4] S. Tulyakov, et al., "Mocogan: Decomposing motion and content for video generation," In *CVPR*, 2018.