CS 512 F24 PROJECT PROPOSAL

Image Translation utilizing CycleGAN

Nehal Dhanraj Patil (A20539654)

Main Paper: Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. arXiv:1703.10593v7 [cs.CV], August 24, 2020.

Authors: Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A. Efros.

Problem Statement:

Image to image translation requires paired training data, which can be difficult or impossible to obtain for many applications. This project aims to address this challenge by implementing the CycleGAN model, which eliminates the need for paired examples by using unpaired datasets. The goal is to develop a system that can translate images between two domains (ex., day to night, sunny to cloudy) while ensuring the translations are realistic & cycle-consistent.

Approach:

We will implement the CycleGAN model, which consists of two generator networks (G: X → Y & F: Y → X) & two discriminator networks (DY & DX). These networks are trained simultaneously using adversarial losses to ensure that the generated images are indistinguishable from real images in the target domains. To prevent arbitrary mappings & mode collapse, the model introduces a cycle consistency loss, ensuring that translating an image to the other domain & back results in the original image.

Key components:

Adversarial Loss: Encourages the generators to produce realistic images that the discriminators cannot distinguish from real images.

Cycle Consistency Loss: Ensures that translations between domains are meaningful & revertible.

Main Components:

- *Two Generators:* G: X → Y & F: Y → X that learn the mappings between domains.
- *Two Discriminators:* DY & DX that distinguish between real & generated images in each domain.
- *Loss Functions:*
  - Adversarial loss for generating realistic images.
  - Cycle consistency loss to preserve structure during domain translation.

Data:

We will use unpaired datasets, such as:

- Horses - Zebras for object transfiguration.
- Summer - Winter for season transfer.
- Photo - Monet for artistic style transfer.

These datasets will be pre-processed to reduce the image resolution, ensuring feasible training times. We will evaluate the model's performance using both quantitative metrics like Frechet Inception Distance (FID) & qualitative perceptual studies to assess the realism & coherence of the generated images.

Evaluation:

The effectiveness of the model will be assessed based on the ability to generate realistic images that are indistinguishable from real images in the target domain. We will measure the model's ability to retain key characteristics of the original domain (ex., object structure or scene composition). Also, we will perform qualitative visual comparisons to assess quality of the translations, with focus on applications such as artistic style transfer and object transfiguration.

References:

- Johnson, J., Alahi, A., & Fei-Fei, L. (2016). Perceptual losses for real-time style transfer and super-resolution. *arXiv*. https://arxiv.org/abs/1603.08155
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. *arXiv*. https://arxiv.org/abs/1406.2661
- Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. *arXiv*. https://arxiv.org/abs/1611.07004
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *arXiv*. https://arxiv.org/abs/1512.03385
- Shrivastava, A., Pfister, T., Tuzel, O., Susskind, J., Wang, W., & Webb, R. (2017). Learning from simulated and unsupervised images through adversarial training. *arXiv*. https://arxiv.org/abs/1612.07828
- Gatys, L. A., Ecker, A. S., & Bethge, M. (2016). Image style transfer using convolutional neural networks. *arXiv*. https://arxiv.org/abs/1508.06576
- *CycleGAN Software*: Official implementation of CycleGAN. Retrieved from: https://github.com/junyanz/CycleGAN
- *ImageNet Dataset*: A large-scale hierarchical image database used for training. Retrieved from: http://www.image-net.org
- UC Berkeley CycleGAN datasets: http://efrosgans.eecs.berkeley.edu/cyclegan/datasets/
- *Set5, Set14, and BSD100 Datasets*: Standard benchmark datasets for evaluating super-resolution models.

Responsibilities(Nehal Dhanraj Patil):

- Implementation of generator networks, training the model on the selected dataset, & conducting experiments.
- Implementation of the discriminator networks, performance evaluation using metrics like FID, & preparation of the presentation.