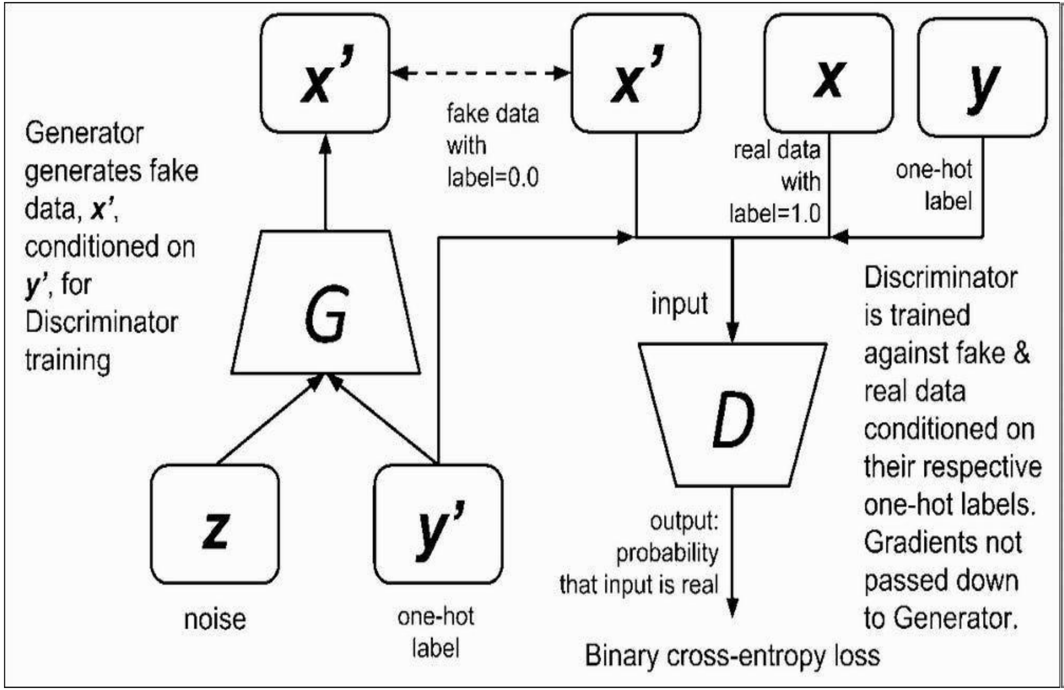
More on GANs

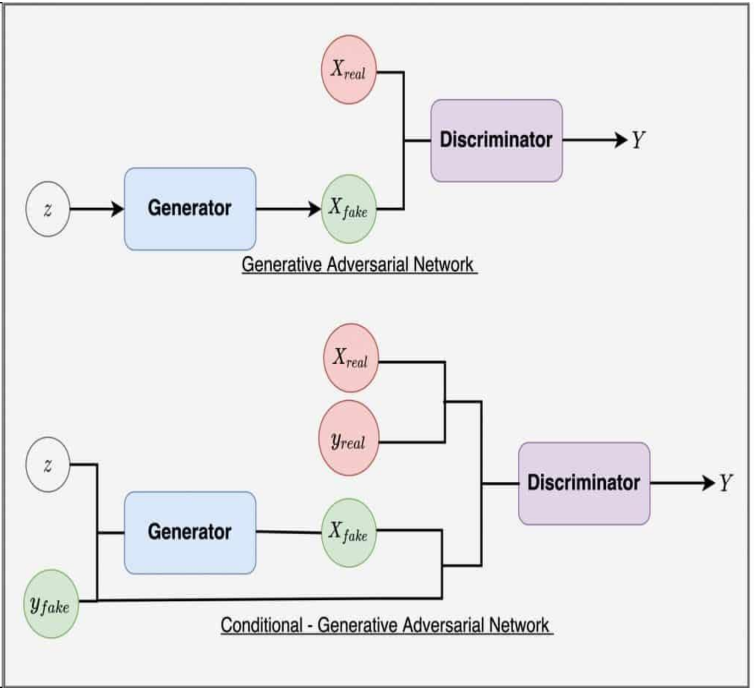
Conditional GAN

CycleGAN

StyleGAN

Conditional Generative Adversarial Networks (cGAN)





Conditional Generative Adversarial Networks (cGAN)

- *Controlled Generation:*
  - Additional conditional input guides the generative process.
  - Enables targeted and controlled data synthesis.
- *Task-Specific Synthesis:*
  - Designed to address task-specific generative challenges.
  - Focuses on producing outputs aligned with given conditions.
- *Adaptive Learning:*
  - Both generator and discriminator adapt to conditional input.
  - Enhances adaptability and generative capabilities.

cGAN Continued

- **Duality of Input:**
  - Blend of random noise and specific conditional information.
  - Provides a dual-input mechanism for the generator.
- **Diverse Applications:**
  - Enables cGANs to tackle a spectrum of tasks.
  - From image-to-image translation to attribute-specific synthesis.
- **Strategic Data Synthesis:**
  - Harnesses the conditional input for strategic and targeted data synthesis.
  - Defines a new paradigm in generative capabilities.

Incorporating Conditional Information

- **Versatile Applications:**
  - cGANs become versatile with the ability to consider conditions.
  - Ideal for tasks where explicit control and guidance are required.
- **Image-to-Image Translation:**
  - Seamless transformation from input to output images with specified conditions.
  - Enables style transfer and content preservation.generative capabilities.
- **Super-Resolution Enhancement:**
  - Generates high-resolution images based on conditional input.
  - Enhances image details and quality.
- **Attribute-Specific Synthesis:**
  - Synthesizes images with specified attributes or features.
  - Tailors outputs based on conditional cues.

Training Procedure of Conditional GANs

- **Dual Input Setup:**
  - Generator takes both random noise and conditional information.
  - Conditional information guides the generation process.
- **Discriminator Training:**
  - Discriminator evaluates real and generated samples.
  - Aims to maximize likelihood for real samples and minimize for fake samples.
- **Generator Training:**
  - Generator strives to minimize adversarial loss and fool the discriminator.
  - Conditional loss ensures alignment with specified conditions.
- **Back-and-Forth Iterations:**
  - Iterative process of training G and D in a competitive manner.
  - Each iteration refines both the generator's creativity and the discriminator discernment.
- **Convergence Challenges:**
  - Ensuring a balance where the generator produces realistic and conditionally aligned outputs.
  - Hyperparameter tuning crucial for stable convergence.

Loss Function Modification

This modification changes the loss function:

$$\min_G \max_D \mathbb{E}_{x \sim p_{\text{data}}(x)}[\log D(x|y)] + \mathbb{E}_{z \sim p_z(z)}[\log(1 - D(G(z|y)|y))]$$

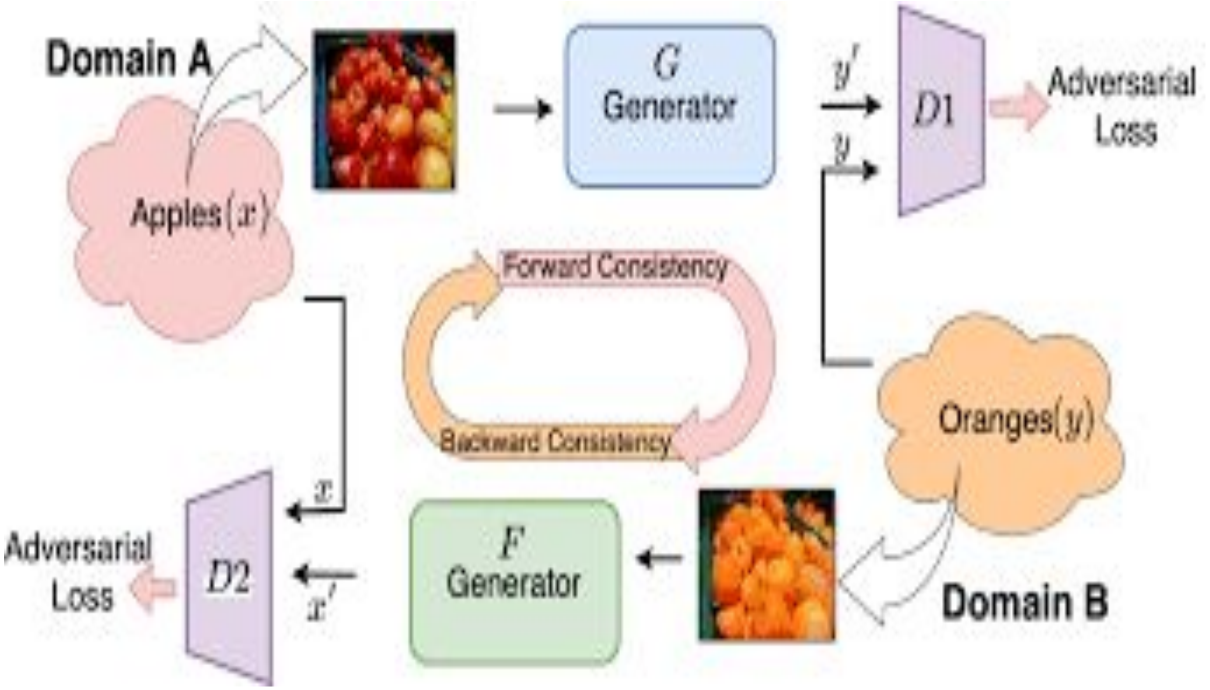
Where:

- $x \sim p_{\text{data}}(x)$ is real data conditioned on y.
- $z \sim p_z(z)$ is random noise.
- $G(z|y)$ generates data conditioned on y.
- $D(x|y)$ predicts whether x is real given y.

Success in Generative Challenges

- **Task-Specific Triumphs:**
  - cGANs excel in tasks requiring specific generative outcomes.
  - Tailored solutions for diverse challenges.
- **Precision Synthesis:**
  - Delivers controlled and targeted data synthesis.
  - Ideal for scenarios demanding fine-tuned outputs.
- **Adaptive Learning:**
  - Demonstrates adaptability in learning from diverse conditional inputs.
  - Thrives in complex generative landscapes.
- **Beyond Random Generation:**
  - Moves beyond random generation to intentional and strategic synthesis.
  - Defines a paradigm shift in generative capabilities.
- **Multifaceted Applications:**
  - From image-to-image translation to attribute-specific synthesis.
  - Proven efficacy across a spectrum of domains.

CycleGAN



Domain A
Apples(x)
G
Generator
y'
y
$D1$
Adversarial
Loss
Forward Consistency
Backward Consistency
Oranges(y)
Domain B
F
Generator
x
x'
$D2$
Adversarial
Loss

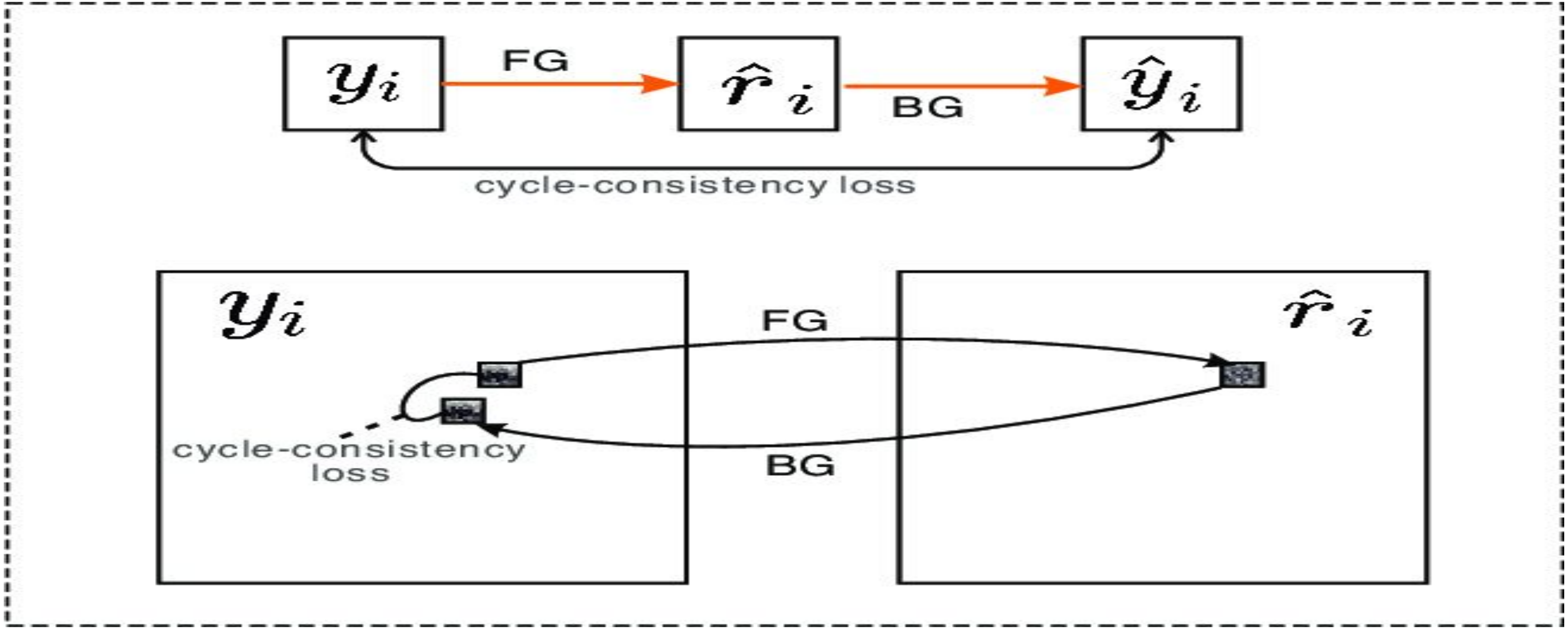
CycleGAN

- **Unpaired Image-to-Image Translation (Unsupervised):**
  - Translates images from one domain to another without paired examples.
- **Cycle-Consistency Loss:**
  - Ensures translated images can be reconstructed back to the original domain.
- **Dual Generators and Discriminators:**
  - Utilizes two generators and discriminators for bidirectional translation.
  - Each generator learns to map images from one domain to the other, enabling bidirectional transformation.
  - Corresponding discriminators assess the authenticity of translated images in both domains.

CycleGAN Continued

- **Adversarial Framework:**
  - Facilitates realistic image generation by generators while discriminators distinguish real from fake images.
- **Stability and Quality:**
  - Dual pairs contribute to stable learning and enhanced image translation quality.
- **Applications:**
  - Enables diverse tasks like style transfer, object transfiguration, and domain adaptation.
  - Applicable across various domains including art, photography, and medical imaging.

Cycle-Consistency Loss



Cycle-Consistency Loss

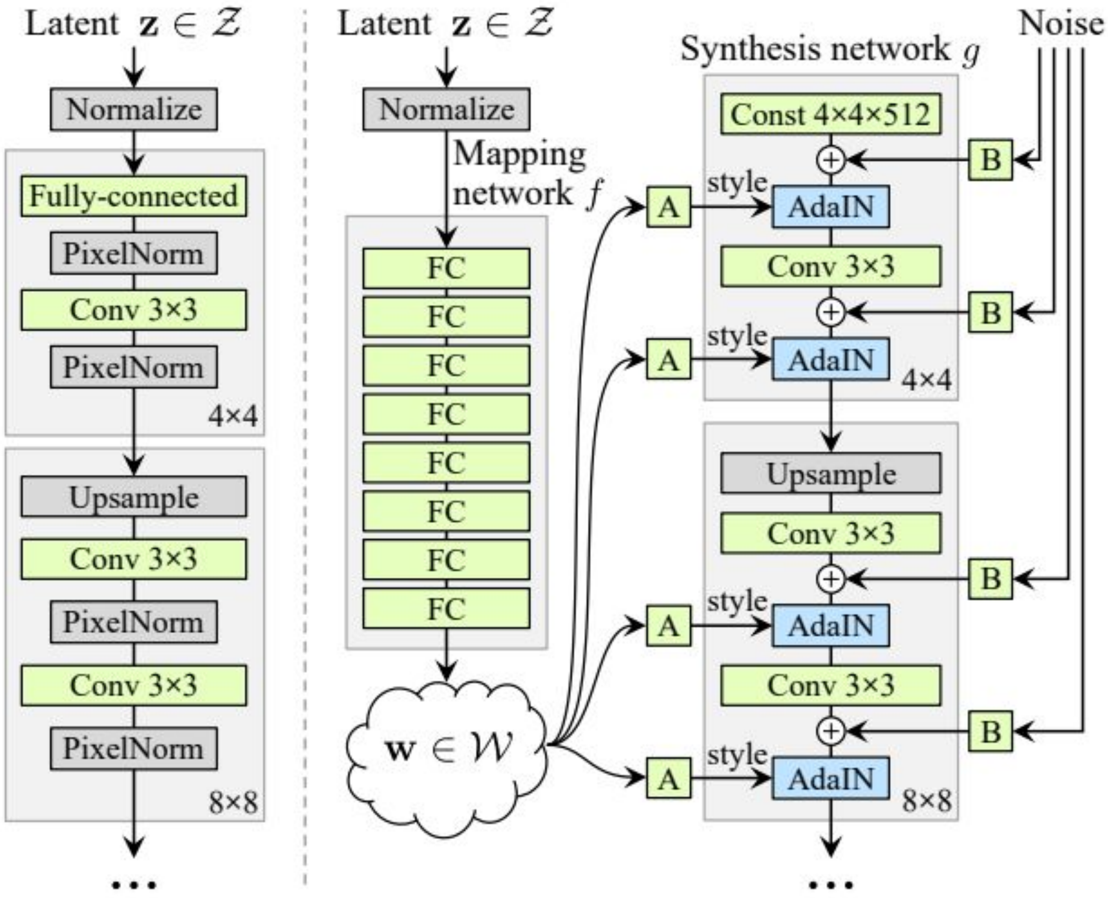
- **Principle:**
  - Ensures consistency between original and translated images.
- **Bidirectional Mapping:**
  - Encourages generators to produce images that can be accurately reconstructed.
- **Loss Calculation:**
  - Measures the discrepancy between the original and reconstructed images using L1 or L2 norms.
- **Enforces Realism:**
  - Contributes to the realism of translated images by maintaining image structure and content.
- **Stabilizing Training:**
  - Helps stabilize the training process and prevents mode collapse.

Drawbacks of CycleGAN

- **Domain Specificity:**
  - Performance may vary across different domains, requiring domain-specific tuning.
- **Limited Control:**
  - Lack of control over specific features or attributes in the translated images.
- **Training Complexity:**
  - Training process can be complex, requiring careful parameter tuning and computational resources.
- **Mode Collapse Risks:**
  - Potential for mode collapse, where the generator focuses on a limited range of outputs.
- **Quality Trade-offs:**
  - Balance between image quality and cycle-consistency may result in compromises.

Exploring StyleGAN: The Art of Image Synthesis

- What is StyleGAN?
  - A cutting-edge GAN architecture by Nvidia for high-quality image synthesis
  - separates content (structure) from style (appearance)
- Evolution from Earlier GAN Models:
  - Addresses limitations of previous architectures (e.g., mode collapse)
  - Introduces a more stable and scalable framework
- Components of StyleGAN:
  - Generator: Transforms random noise into high-resolution images
  - Discriminator: Distinguishes between real and generated images
  - Latent Space: Controls image features such as age, gender, and more
- Adaptive Instance Normalization (AdaIN):
  - Allows for flexible style transfer and image manipulation



(a) Traditional

(b) Style-based generator

StyleGAN Architecture

- Mapping Network
  - Instead of directly using a noise vector z, StyleGAN first processes it through an 8-layer fully connected neural network.
  - This network maps z to an intermediate latent space w, which allows for better disentanglement of image features.
- Adaptive Instance Normalization
  - The mapped latent vector w is transformed into style vectors at different layers.
  - This lets different layers control different levels of detail, from coarse structure (e.g., face shape) to fine details (e.g., skin texture).

StyleGAN Architecture

- Stochastic Noise Injection
  - Random noise is added at each layer to introduce stochastic variations, ensuring natural-looking images with details like wrinkles or freckles.
- Progressive Growing
  - Images are generated progressively, starting from a low resolution (4×4) and growing to high resolution (1024×1024).
  - This stabilizes training and improves image quality.

How AdaIN Separates Structure from Style?

- The early layers (coarse resolution, e.g., 4×4, 8×8) control structure (e.g., face shape, pose).
- The later layers (fine resolution, e.g., 64×64, 1024×1024) control style (e.g., colors, texture, lighting).
- This means we can change the style while keeping the structure fixed, or vice versa.

Advantage of StyleGAN

- Style Mixing & Interpolation
  - One major advantage of StyleGAN is style mixing, where two different latent codes w1 and w2 are used at different layers. This enables:
    - Hybrid images that combine features from two different sources.
    - Smooth interpolation between different styles, allowing for morphing effects.

StyleGAN's Key Features

- High-Resolution Image Synthesis:
  - Ability to generate images of up to 1024x1024 resolution
  - Unprecedented level of detail and realism
- Controllable Image Attributes:
  - Fine-grained control over features like age, pose, and expression
  - Enables creation of diverse and customizable images
- Diversity in Generated Images:
  - Showcase of the wide range of outputs, from faces to landscapes
  - Highlighting the versatility and richness of StyleGAN's creations

Applications of StyleGAN

- Art and Creative Expression:
  - Artists using StyleGAN for surreal and imaginative artworks
- Face Synthesis and Editing:
  - Applications in entertainment, virtual avatars, and digital personas
- Fashion Design and Visualization:
  - Creating virtual fashion lines, exploring new styles and trends
- Data Augmentation in Machine Learning:
  - Generating synthetic datasets for training robust ML models

Example