**Conditional GAN (cGAN) for Image-to-Image Translation**

**Author:** Sunny Kumar  
**Course:** B.Tech CSE, 2nd Year  
**Internship:** Prodigy Infotech – Task 04

**Abstract**

This report presents the implementation of a Conditional Generative Adversarial Network (cGAN) for image-to-image translation using the Pix2Pix architecture. The CMP Facades dataset was used, which contains paired images of building facades and their edge maps. The goal of this project was to generate realistic facade images from structural outlines.

The project demonstrates how cGANs can successfully learn mappings between input and output image domains. The results show progressive improvement in generated images across training epochs, validating the effectiveness of the Pix2Pix framework. This work has practical applications in architectural design automation, satellite imagery analysis, and computer vision research.

**1. Introduction & Literature Review**

Generative Adversarial Networks (GANs), first introduced by Goodfellow et al. in 2014, are a class of deep learning models where two neural networks (a generator and a discriminator) compete with each other in a minimax game. The generator attempts to produce realistic samples, while the discriminator tries to distinguish between real and generated data.

Conditional GANs (cGANs) extend this framework by introducing conditioning information, such as class labels or images, into both networks. This allows for controlled generation of data rather than purely random outputs.

**Pix2Pix** (Isola et al., 2017) is a widely used implementation of cGANs designed specifically for image-to-image translation tasks. Unlike traditional GANs, Pix2Pix learns mappings between paired domains, such as:

* Sketches → Photos
* Aerial images → Maps
* Edge maps → Buildings

In this project, Pix2Pix was applied to the **CMP Facades dataset**, which consists of building facades paired with their edge maps.

**2. Methodology**

**2.1 Dataset**

The **CMP Facades dataset** contains around 400 images of real-world building facades, each paired with an annotated edge map. These paired images enable supervised training for image-to-image translation tasks.

* Input: Edge maps (structural outlines of buildings)
* Output: Realistic building facade images

**2.2 Preprocessing**

* The dataset was loaded into Google Colab and extracted.
* Images were resized and normalized to the range [-1, 1].
* Data loaders were created for batching and shuffling during training.

**2.3 Model Architecture**

**Generator (U-Net)**

* Encoder-decoder structure with skip connections.
* Converts input edge maps into realistic facade images.

**Discriminator (PatchGAN)**

* Classifies each image patch as real or fake.
* Provides local realism feedback rather than global.

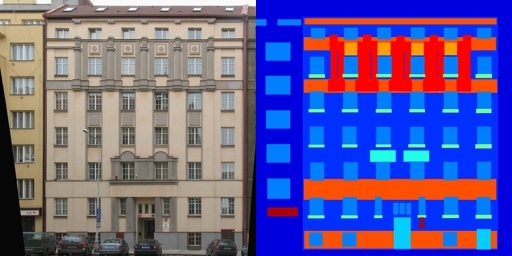
**2.4 Training Setup**

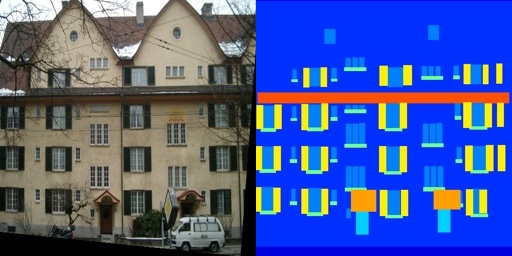
* Framework: PyTorch
* Optimizer: Adam (learning rate 2e-4, β1 = 0.5)
* Loss Functions:
  + Adversarial Loss (to fool the discriminator)
  + L1 Loss (to ensure closeness to target image)
* Training: 20 epochs on GPU (Google Colab CUDA support)

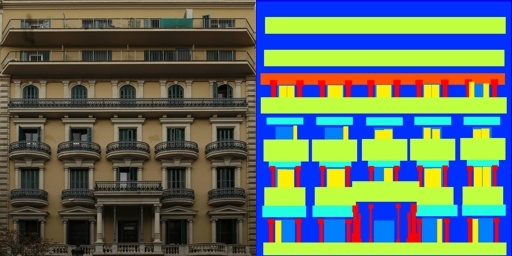
**3. Implementation & Results**

Training produced progressively better results across epochs. The discriminator loss (Loss\_D) and generator loss (Loss\_G) were monitored along with L1 reconstruction loss.

**Sample Results**

**Epoch 1 Output**  
*Figure 1: Early training output shows blurry and inconsistent structures.*  


**Epoch 5 Output**  
*Figure 2: Facade details begin to emerge, though artifacts remain visible.*  


**Epoch 20 Output**  
*Figure 3: Final output with improved clarity and realistic building structures.*  


The results clearly show how the model gradually improves at mapping edge structures into detailed facades.

**4. Discussion & Applications**

This project demonstrates that cGANs are capable of effective image-to-image translation. In particular, Pix2Pix successfully mapped edge drawings of buildings into realistic facade images.

**Applications**

* **Architectural Design Automation** – Generating facade proposals from sketches.
* **Urban Planning** – Converting structural outlines into building visualizations.
* **Satellite Imagery** – Enhancing or reconstructing low-quality map data.
* **Artistic Style Transfer** – Translating artistic sketches into photorealistic scenes.

**Limitations**

* Requires paired training data (not always available).
* Sensitive to dataset size; small datasets limit generalization.
* Training instability due to adversarial learning.

**5. Conclusion & Future Work**

In this internship task, a Pix2Pix-based Conditional GAN was implemented on the CMP Facades dataset. The generator learned to transform edge maps into realistic facades with increasing accuracy across epochs.

**Future directions include:**

* Using **CycleGAN** for unpaired datasets.
* Exploring **improved architectures** for more stable training.
* Scaling to **larger, more diverse datasets**.

This project illustrates the potential of cGANs in real-world computer vision tasks and serves as a foundation for further research in generative AI.

**References**

1. Goodfellow, I., et al. (2014). *Generative Adversarial Networks.*
2. Isola, P., Zhu, J., Zhou, T., & Efros, A. A. (2017). *Image-to-Image Translation with Conditional Adversarial Networks (Pix2Pix).*
3. CMP Facades Dataset – Czech Technical University.