**A PRELIMINARY REPORT ON**

**CGAN – DRIVEN IMAGE MAPPING**

Submitted By:

Manali Jain (22110360)

Pushkar Patil (22110200)

Chetana Shinde (22110601)

Ayush Mundhe (22110630)

Under the Guidance of

**Prof. Mandar Mokashi**

In partial fulfilment of

**Bachelor of Technology**

**[**B.Tech. Information Technology**]**

[2024-25]

At



Department of Information Technology

Vishwakarma Institute of Information Technology, Pune 411048

*Affiliated To*

Savitribai Phule Pune University, Pune

**BANSILAL RAMNATH AGARWAL CHARITABLE TRUST,**

**(An Institute affiliated to Savitribai Phule Pune University)**

**VISHWAKARMA INSTITUTE OF INFORMATION TECHNOLOGY, PUNE**

**2024 – 2025**



**CERTIFICATE**

This is to certify that the project report entitles

**CGAN – DRIVEN IMAGE MAPPING**

Submitted by

**MANALI JAIN 22110360**

**PUSHKAR PATIL 22110200**

**CHETANA SHINDE 22110601**

**AYUSH MUNDHE 22110630**

This is to certify that the work “**CGAN-IMAGE DRIVEN MAPPING”** is a bonafide work carried out by Mr. Pushkar Patil, Ms. Chetana Shinde, Ms. Manali Jain, Mr. Ayush Mundhe in partial fulfilment of the award of Bachelor of Technology in Information Technology, Savitribai Phule Pune University, Pune, during the year 2024. The project report has been approved as it satisfies the academic requirements in respect of the project work prescribed for the Bachelor of Technology Degree.

**Project Guide Head of Department**

(Prof. Mandar Mokashi) **(**Prof. P. Futane)

Place: Date:

Remark:

**ACKNOWLEDGEMENT**

We take this opportunity to thank Head of the Department Prof. P. Futane and our project guide Prof. Mandar Mokashi for their valuable guidance and providing all the necessary facilities, which were indispensable in the completion of this project report. We are also thankful to all the staff members of the Department of Information Technology, VIIT, Pune for their valuable time, support, comments, suggestions and persuasion. We would also like to thank Prof. Mandar Mokashi, Mrs. Suruchi G. Dedgaonkar, Mrs. Shalini V. Wankhade and the institute for providing the required infrastructure, internet and library facility.

Manali Jain (22110360)

Pushkar Patil (22110200)

Chetana Shinde (22110601)

Ayush Mundhe (22110630)

**LIST OF FIGURES**

[Figure 1: Name of the figure **Error! Bookmark not defined.**](#_Toc179976571)

**CONTENTS**

[**INTRODUCTION** 6](#_Toc179982543)

# **ABSTRACT**

Conditional Generative Adversarial Networks (C-GANs) have demonstrated significant potential in the field of image-to-image translation, enabling the transformation of abstract representations into realistic images. This project leverages C-GANs to convert architectural sketches or blueprints into photorealistic images of buildings. By modifying the generator and discriminator components, the system can be adapted to various other applications, such as translating satellite images into real-world landscape visuals. The primary objective of this project is to develop a flexible and efficient model that can accurately map and visualize architectural designs in real-world contexts. The approach involves training the C-GAN on a dataset of paired sketches and corresponding real-world images, refining the model to optimize the quality and accuracy of the output. The results highlight the model's capability to produce highly detailed and realistic images, showcasing its applicability in fields such as architecture, urban planning, and geographic information systems (GIS). This work underscores the importance of generative AI in bridging the gap between abstract design concepts and tangible visualizations, offering a valuable tool for professionals in design and planning industries.

**CHAPTER 1**

**INTRODUCTION**

* 1. **Motivation**

The motivation behind **C-GAN Driven Image Mapping** stems from the need to transform abstract or incomplete visual representations, such as sketches or blueprints, into realistic images. This technology aims to assist architects, designers, and urban planners by automating the process of generating detailed visual interpretations from minimalistic inputs. By utilizing Conditional Generative Adversarial Networks (C-GAN), the system can streamline workflows, reduce design time, and enable a faster exploration of multiple design concepts. The ability to enhance preliminary sketches into more detailed and realistic images is invaluable for design communication, visualization, and rapid prototyping.

* 1. **Need for CGAN – Driven Image Mapping**

C-GANs are crucial in **image-to-image translation tasks** due to their capability to condition the output based on an input image. Traditional GANs have been successfully used to generate images from random noise, but C-GANs take this a step further by incorporating conditions, such as sketches or satellite images, as input. This is particularly useful in scenarios where transforming input data (like architectural sketches or low-resolution images) into more visually detailed and contextually relevant outputs is required. C-GAN Driven Image Mapping serves several important use cases, such as:

* **Architectural Design**: Transforming architectural blueprints into realistic visualizations.
* **Urban Planning**: Converting low-resolution satellite imagery into detailed maps or street-level images.
* **Creative Design**: Enhancing artistic sketches into polished artwork for various industries.
* **Automation of Image Processing**: Reducing the manual effort required in creating detailed visual content from rudimentary inputs.
  1. **Brief Introduction to C-GAN Driven Image Mapping**

C-GAN Driven Image Mapping is an advanced technique based on Conditional Generative Adversarial Networks (C-GAN), which translates input images such as sketches, blueprints, or low-resolution satellite images into detailed, realistic images. The core idea behind this technology is to have a generator that creates realistic images conditioned on an input, and a discriminator that evaluates how closely the generated image matches the expected output. The interplay between these two networks enables the model to learn how to map abstract inputs into detailed, contextually accurate outputs.

C-GANs extend traditional GANs by incorporating an additional input (the condition), allowing for more control over the image generation process. In the case of image mapping, the condition could be a sketch or a satellite image, and the model is trained to generate a realistic representation of what that input would look like in a real-world context. This technique has found applications in architecture, urban planning, and other creative fields where transforming simplified visuals into fully detailed imagery is essential.

* 1. **Application**

C-GAN Driven Image Mapping has a wide range of applications across various industries, enabling efficient, automated, and creative transformations of abstract or incomplete images into more realistic and useful visual representations. Some of its notable applications include:

1. Architectural Design and Visualization:
   * Blueprint to Building Visualization: C-GANs can transform architectural blueprints and sketches into lifelike renderings, helping architects and clients visualize the final design before construction.
   * Interior Design: Designers can convert 2D layout sketches into realistic interior visuals, enhancing the design process for better decision-making.
2. Urban Planning and Geographic Mapping:
   * Satellite Image Enhancement: Low-resolution satellite images can be translated into high-resolution, street-level maps, aiding urban planners in assessing landscapes and infrastructure.
   * City Planning: Planners can use C-GAN models to generate realistic visualizations of city layouts, helping them analyze future development projects and make data-driven decisions.
3. Art and Creative Design:
   * Sketch-to-Image Translation: Artists and designers can convert hand-drawn sketches into polished artwork, saving time and effort while exploring various creative concepts.
   * Fashion Design: Rough fashion sketches can be converted into fully rendered images, helping designers and brands quickly visualize clothing designs.
4. Medical Imaging:
   * Medical Image Translation: C-GANs can assist in converting medical scans (e.g., low-resolution MRI or CT scans) into more detailed images, aiding in diagnosis and analysis.
   * Image-to-Image Translation for Diagnosis: C-GANs can help translate various imaging modalities, improving the accuracy and speed of diagnostic tools.
5. Automotive Design:
   * Concept Car Visualization: Designers can use sketches of concept cars to generate realistic 3D visualizations, assisting in the iterative design process.
   * Parts and Mechanisms Design: C-GANs can generate realistic images of car parts from engineering diagrams, helping streamline prototyping.
6. Gaming and Entertainment:
   * Game Character Design: Game developers can use rough sketches to generate fully rendered 3D models or character designs, accelerating the creation process.
   * Environment Rendering: Virtual environments for games or movies can be designed from rough landscape or city sketches, saving time and increasing creativity.
7. Agriculture and Remote Sensing:
   * Satellite-to-Terrain Mapping: C-GANs can be used to convert satellite imagery into highly detailed terrain maps, aiding agricultural monitoring, deforestation analysis, and environmental conservation.

These applications showcase how C-GAN-driven image mapping is transforming diverse fields, from architecture and urban planning to entertainment and healthcare, by enabling the automatic generation of high-quality, detailed images from minimal inputs.

**CHAPTER 2**

**LITERATURE SURVEY**

**2.1 Literature Review**

**Table No 2.1: Literature Survey of Papers on PIX2PIX Model**

|  |  |  |
| --- | --- | --- |
| **Sr.No** | **Paper** | **Summary** |
| **1.** | Image-to-Image Translation with Conditional Adversarial Networks | Limited to paired image translation large data requirement |
| **2.** | CycleGAN: Unpaired Image-to-Image Translation | Complex training, hyperparameter tuning. |
| **3.** | Pix2PixHD: High-Resolution Image Synthesis | High computational requirements, complex architecture. |
| **4.** | DualGAN: Unsupervised Dual Learning | Limited domains, potential mode collapse. |
| 5. | GANimation: Anatomically aware Facial Animation | Limited to facial animation, preservation of details. |
| 6. | Progressive Growing of GANs | High computational resources. |
| 7. | Deep Photo Style Transfer | Limited to style transfer, content/style fidelity. |
| 8. | Towards Realistic Image Pairing and Transformation | Focuses on improving image pairing and transformation for better image generation results. |
| 9. | StarGAN – Muti Domain Image to Image Translation | Limited control over attributes, mode collapse. |
| 10. | Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks | Introduces Cycle GAN for unpaired  image translation, addressing the challenge of lack of paired data. |

**2.2 Review of Existing System**

The existing systems for **image-to-image translation** primarily leverage variations of **Conditional Generative Adversarial Networks (C-GANs)**, particularly the **Pix2Pix model**, which has set a foundational standard for paired image translation tasks. This review highlights key advancements and challenges in the field, as identified from the literature on C-GAN-driven models.

1. **Pix2Pix and Paired Image Translation:**

The Pix2Pix model, introduced by Isola et al., was a landmark in paired image translation tasks. It enables the conversion of an input image to a target image based on paired data, demonstrating successful applications in tasks like transforming sketches into images and satellite photos into maps.

**Strengths:** Pix2Pix's ability to generate high-quality results when the paired data is available, making it effective in supervised learning tasks such as semantic segmentation and image colorization.

**Challenges:** A major drawback is its reliance on large amounts of paired data, which is not always available or easy to collect. Additionally, the model is computationally intensive, limiting its scalability for larger datasets or higher-resolution images. Moreover, it requires careful parameter tuning to avoid issues like overfitting and poor generalization.

1. **Unpaired Image Translation with CycleGAN and DualGAN:**

To overcome the limitation of requiring paired data, models like CycleGAN (Zhu et al.) and DualGAN (Yi et al.) were developed to handle unpaired image translation tasks.

**Strengths:** These models eliminate the dependency on paired datasets by introducing a cycle-consistency loss, where the input image is translated into a target domain and then mapped back to the original domain. This method has shown success in applications such as translating photographs to artistic paintings and vice versa.

**Challenges:** Despite handling unpaired data, these models often suffer from mode collapse, where the generator fails to produce diverse outputs, leading to limited creativity in the generated images. They also face difficulties in achieving stable convergence during training, making it hard to ensure consistent results across different datasets.

1. **Multi-Domain Translation with StarGAN and BicycleGAN:**

StarGAN (Choi et al.) and BicycleGAN (Zhu et al.) extended the Pix2Pix framework to enable multi-domain image translation and attribute manipulation. StarGAN, for instance, allows translation between multiple domains (e.g., altering facial expressions or hair color in images) using a single model.

**Strengths:** These models introduced the ability to handle multiple image domains and attributes with a unified framework, significantly improving flexibility in image manipulation tasks.

**Challenges:** While these methods are powerful, they often struggle with controlling translation attributes accurately. For instance, it can be challenging to selectively manipulate one feature (like hair color) without unintentionally altering other features. Additionally, ensuring mode diversity (generating multiple distinct outputs for the same input) remains a significant obstacle in these systems.

1. **Datasets Used:**

The review of the literature also emphasizes the range of datasets that have been used for training these models, including Cityscapes (urban scenes), CelebA (celebrity faces), LSUN (large-scale scene understanding), and domain-specific datasets like Facades (architectural images) and Maps.

**Challenges with Data:** While these datasets provide robust training grounds, paired datasets are often time-consuming to create, and domain-specific datasets may not generalize well across different tasks or environments. Moreover, unpaired datasets, though easier to collect, introduce complexities in maintaining consistent translation quality.

* **Key Observations and Gaps:**

1. **Computational Complexity:** Most of these models, especially Pix2Pix and its variants, demand significant computational resources, particularly in terms of memory and training time, limiting their scalability for large datasets or high-resolution tasks.
2. **Training Stability:** Many models face difficulty achieving stable training, especially with unpaired image translation, leading to issues like mode collapse and degraded output quality.

**CHAPTER 3**

**PROJECT STATEMENT**

**3.1 Purpose behind the Project**

The C-GAN Driven Image Mapping project aims to explore and extend the capabilities of Conditional Generative Adversarial Networks (C-GANs) to transform input images, such as sketches, blueprints, or low-resolution satellite imagery, into high-quality, realistic visuals. By automating the traditionally manual process of converting rough or incomplete inputs into fully detailed images, this project seeks to reduce human intervention, streamline workflows, and enhance creativity across various industries, including architecture and urban planning. Additionally, the project focuses on improving visualization and decision-making by enabling stakeholders to visualize complex designs or layouts with greater clarity. To push the boundaries of AI in image synthesis, the project will address existing challenges such as training stability, multi-domain image translation, and ensuring diversity in the generated outputs, while experimenting with innovative network structures to enhance image-to-image translation across diverse applications.

**3.2 Decision Scope**

The scope of the **C-GAN Driven Image Mapping** project focuses on transforming architectural blueprints, satellite imagery, and hand-drawn sketches into realistic images, with key applications in architecture, urban planning, and creative design. The project will utilize and compare C-GAN variants like Pix2Pix for paired image translation, CycleGAN for unpaired tasks, and StarGAN for multi-domain image translation. It will prioritize datasets such as Cityscapes, CelebA, and Facades, along with domain-specific data. The technical focus includes improving training stability, ensuring scalability, and enhancing image quality for more efficient and reliable results.

**Chapter 4**

**PROJECT ANALYSIS AND DESIGN**

**4.1 Use Case Diagram**

A diagram of a diagram

Description automatically generated

**4.2 Sequence Diagram**

A diagram of a process

Description automatically generated

**4.3 Class Diagram**

A diagram of a model

Description automatically generated

**4.4 Activity Diagram**

A diagram of a process

Description automatically generated

**CHAPTER 5**

**CONCLUSION AND FUTURE WORK**

**5.1 Conclusion**

The C-GAN Driven Image Mapping project successfully demonstrates the potential of Conditional Generative Adversarial Networks (C-GANs) in transforming rough sketches, blueprints, and low-resolution images into high-quality, realistic visuals. By automating image translation, the project reduces manual effort and accelerates workflows across industries like architecture, urban planning, and design. Additionally, it addresses key technical challenges such as training stability, scalability, and image quality enhancement. The integration of multiple C-GAN variants (Pix2Pix, CycleGAN, StarGAN) allows for greater flexibility in both paired and unpaired data translation, further showcasing the wide applicability of C-GANs in real-world scenarios.

**5.2 Future Scope**

The C-GAN Driven Image Mapping project successfully demonstrates the potential of Conditional Generative Adversarial Networks (C-GANs) in transforming rough sketches, blueprints, and low-resolution images into high-quality, realistic visuals. By automating image translation, the project reduces manual effort and accelerates workflows across industries like architecture, urban planning, and design. Additionally, it addresses key technical challenges such as training stability, scalability, and image quality enhancement. The integration of multiple C-GAN variants (Pix2Pix, CycleGAN, StarGAN) allows for greater flexibility in both paired and unpaired data translation, further showcasing the wide applicability of C-GANs in real-world scenarios.

**CHAPTER 6**

**REFERENCES**

**[1]** Christopher hesse. https://affinelayer.com/ pixsrv/.

**[2]** L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille. Semantic image segmentation with deep convolutional nets and fully connected crfs. In ICLR, 2015.

**[3]** E. Denton, S. Chintala, A. Szlam, and R. Fergus. Deep generative image models using a laplacian pyramid of adversarial networks. In NIPS, 2015

**[4]** D. Eigen and R. Fergus. Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture. In ICCV, 2015.

**[5]** L. A. Gatys, A. S. Ecker, and M. Bethge. Image style transfer using convolutional neural networks. CVPR, 2016.

**[6]** I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. WardeFarley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In NIPS, 2014.

**[7]** L. Karacan, Z. Akata, A. Erdem, and E. Erdem. Learning to generate images of outdoor scenes from attributes and semantic layouts. arXiv preprint arXiv:1612.00215, 2016.

**[8]** C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi. Photo-realistic single image super-resolution using a generative adversarial network.

**[9]** C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi. Photo-realistic single image super-resolution using a generative adversarial network.

**[10]** M. Mathieu, C. Couprie, and Y. LeCun. Deep multi-scale video prediction beyond mean square error. ICLR, 2016.