IMAGE MANIPULATION USING GENERATIVE ADVERSARIAL NETWORKS

VISHAL JC	(2127200501164)
VISHAL M	(2127200501165)
VIVEK MUNNAA D	(2127200501169)

CS18711 – MINIPROJECT

FINAL REVIEW

Batch No	: 62
Name of the Supervisor	: Ms. P. UMA
Date of Review	: 10/11/2023
Domain	: DEEP LEARNING

ABSTRACT

Employing Generative Adversarial Networks (GANs), this project transforms grayscale images into vibrant color representations. Its significance lies in historical restoration, visual media enhancement, medical diagnostics support, and artistic expression. The GAN-based generator learns to add color, while the discriminator ensures realism, harnessing deep learning and image processing for potential applications across domains, from preserving history to improving healthcare and multimedia content. This project revolutionizes the use of GANs (Generative Adversarial Networks) by exploiting the potential of deep learning and image processing.

PROBLEM STATEMENT

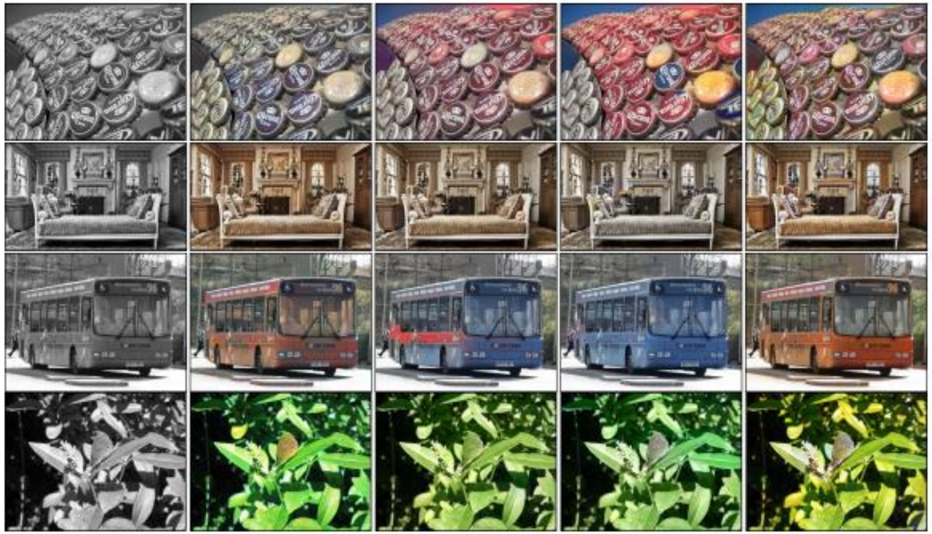
The task of image processing is both complex and crucial role in many fields including medicine, forensics, art and architecture. These fields have the ability to edit and enhance images essential for diagnosis, analysis, creative expression and visual perspective communication. This emphasizes the importance of a strong and solid finish intelligent image processing software solutions. Here, we apply Conditional GAN(CGAN) to specify labels during GAN training and enable the network to quickly absorb essential details, thus discovering the deep representations of steerable dimensions facilitates promising applications of real image editing and interactive image manipulation.

INTRODUCTION TO GENERATIVE ADVERSARIAL NETWORKS

- Generative Adversarial Networks, or GANs are an approach for generative modeling using deep learning methods, such as convolutional neural networks.
- Generative modeling is an unsupervised learning task in machine learning that involves automatically discovering and learning the regularities or patterns in input data in such a way that the model can be used to generate or output new examples that plausibly could have been drawn from the original dataset.
- GANs are a clever way of training a generative model by framing the problem as a supervised learning problem with two sub-models: the generator model that we train to generate new examples, and the discriminator model that tries to classify examples as either real (from the domain) or fake (generated).
- In this project, we use CGAN or conditional generative adversarial network ,which is a type of GAN that takes advantage of labels during the training process. It enables more precise generation, discrimination of images to train machines and allow them learn by themselves.

RESEARCH CHALLENGES

- Complex Mapping
- Size of the Dataset
- Colour Consistency
- Computational Resources
- Training Time
- Evaluation Metrices



SCOPE OF THE WORK

- Conditional GANs (CGANs), which has proven to be a powerful and flexible tool for automatic colorization, offers the potential to produce high-quality colorized images for a wide range of applications.
- Users from various disciplines can use this for effective colorization of their input data which is colorized with a accuracy score .The output image has high generative power, realism and diversity, efficient feature mapping, perceptual quality, incorporation of context.
- The scope of CGAN-based colorization hold significant potential that includes image and video restoration, artistic expression, medical imaging, forensics, cultural heritage and education. As deep learning models and algorithms become more sophisticated, the potential applications of colorization in various domains will expand, making it a versatile and valuable tool for visual enhancement and creative expression.

LITERATURE REVIEW

S.NO	PAPER TITLE	AUTHOR	REVIEW OF THE PAPER
1	Image-to-Image Translation with Conditional Adversarial Networks	Phillip Isola, Jun-Yan Zhu, Tinghui Zhou and Alexei A. Efros.	The paper delves into the synergy of GAN loss and L1 loss for image colorization with conditional GANs. By combining adversarial and perceptual loss components, the proposed optimized loss function refines colorization results. Notably, the integration of these losses allows effective training even with limited datasets, making the entire process faster and more resource-efficient. This breakthrough contributes to a more streamlined and accessible approach to image colorization.
2	Double-Channel Guided Generative Adversarial Network for Image Colorization	Kangning Du, Changtong Liu, Lin Cao, Yanan Guo	The paper introduces a groundbreaking DCGGAN network to address abnormal colorization issues in deep learning-based image colorization methods. By incorporating a reference component matching module and a double-channel guided colorization module. Results provides a novel and effective solution for achieving accurate and realistic colorization while minimizing abnormal color artifacts.
3	Image Colorization with Generative Adversarial Networks	K. Nazeri, E.Ng	The paper explores grayscale image colorization using generative adversarial networks (GANs), comparing results with existing convolutional neural networks (CNNs). They employ a novel cost function for the generator to address training challenges, achieving improved image quality. Preliminary results on CIFAR-10 demonstrate GAN's ability to produce vibrant and visually appealing colorizations.

LITERATURE REVIEW [Continued]

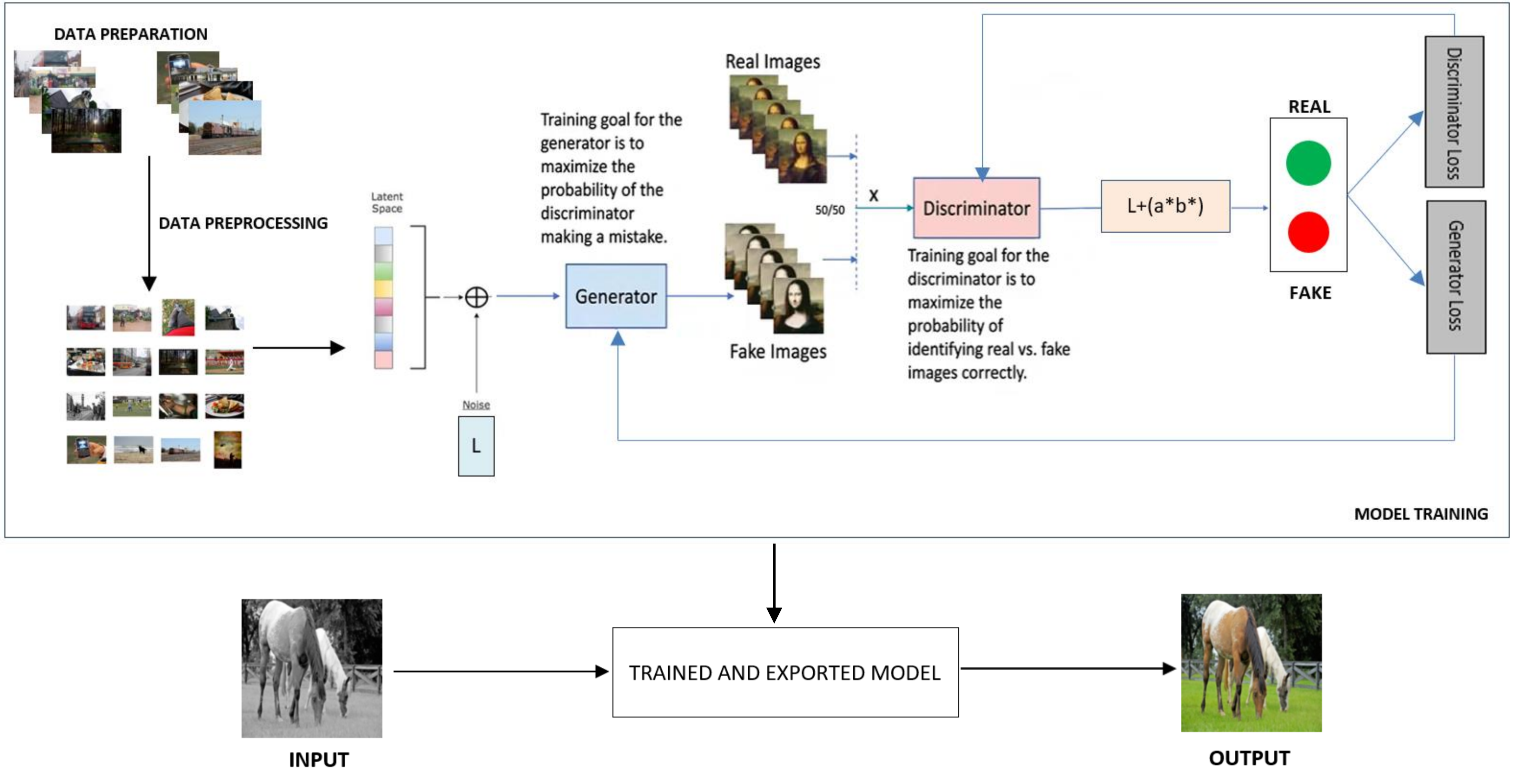
S.NO	PAPER TITLE	AUTHOR	REVIEW OF THE PAPER
4	Recovering Old or Damaged Images using GAN	Harshith J L, Chandan Kumar M, Sanjana N Dhanush B L, K S Mahesh	The ideology of the paper discusses the use of generative adversarial networks (GANs) for image restoration and colorization, highlighting the effectiveness of GANs in solving image-to-image translation problems. It addresses the limitations of traditional methods that rely on hand-coding loss functions and presents a novel approach that leverages GANs for more satisfying results.
5	Image Colorization with Palette Generative Adversarial Networks	Yi Wang, Menghan Xia, Lu Qi, Jing Shao, and Yu Qiao	PalGAN Image Colorization with Palette Generative Adversarial Networks. It has a probabilistic palette estimation mechanism, which breaks away from traditional deterministic colorization and allows for a more diverse and controllable color assignment and other feature chromatic attention mechanism, which is crucial in tackling the problem of color bleeding. This module aligns color affinities with both semantics and low-level characteristics, ensuring that the generated colors adhere to the semantic boundaries of the objects in the image.

PROPOSED WORK

The Dataset handled in this project are publicly available images ranging from a test size of 16-24GB. The model is trained with about 80% of the dataset and the rest is used for testing the model. Training a Conditional Generative Adversarial Network (CGAN) involves several steps. Here's a workflow for training a CGAN:

- **Data Preparation**
- **Data Preprocessing**
- **Model Architecture**
- **Loss Functions**
- **Model selection and Training**

ARCHITECTURE



HARDWARE AND SOFTWARE REQUIREMENTS

- **HARDWARE REQUIREMENTS:**

Operating System	:	Windows 10 or Higher
Processor	:	I5 11th Gen Processor or Higher
RAM	:	Minimum 16GB is recommended
GPU	:	NVIDIA GeForce with CUDA

- **SOFTWARE REQUIREMENTS:**

Programming Language	:	Python v3.11 or above
Deep Learning Framework	:	Pytorch v2.1.0 or above
Machine Learning Libraries	:	Skimage, numpy, pandas, matplotlib, PIL
Development Environment	:	Jupyter Notebook

MODULE 1 : DATA SELECTION, PREVIEW AND DATA PREPARATION

INPUT:

A diverse image dataset from various scenarios.

OUTPUT:

Dataset sizes, a visual preview, and data loaders for efficient handling.

IMPLEMENTATION:

This module encompasses dataset selection, sizing, preview, and data preparation for colorization tasks. It utilizes a COCO dataset, randomly selecting 10,000 images, shuffling, and allocating 8,000 for training and 2,000 for validation. The dataset sizes are reported. Additionally, it includes a 4x4 image grid for visual data inspection. To facilitate data handling, this module seamlessly prepares the dataset by resizing, data augmentation, and normalization, creating data loaders for both training and validation. This robust preparation ensures the dataset is ready for use in machine learning models.

MODULE 2 : GENERATOR – UNET ARCHITECTURE FOR COLORISATION

INPUT:

Configuration parameters for the UNet architecture, such as the number of down-sampling layers, filters, and input/output channels.

OUTPUT:

A UNet model tailored for colorization.

IMPLEMENTATION:

In this module, we introduce the UNet architecture, referred to as the generator, designed for colorization tasks. The UNet comprises building blocks, each encapsulating convolutional, activation, and normalization layers, allowing flexibility for down-sampling, up-sampling, and inner layers. This adaptable structure can serve as the outermost, innermost, or intermediate block, with optional dropout layers for regularization.The Unet model assembles these blocks to create a complete UNet architecture for colorization. It offers control over the number of down-sampling layers and filters for task-specific customization. The model takes grayscale images as input and produces colorized results, demonstrating its effectiveness for image-to-image tasks.

MODULE 3 : DISCRIMINATOR (PatchGAN)

INPUT:

Configuration parameters for the PatchGAN discriminator, specifying the number of filters and down-sampling layers.

OUTPUT:

A PatchGAN discriminator optimized for adversarial learning in colorization.

IMPLEMENTATION:

This module introduces the PatchGAN discriminator, a key component for adversarial learning in the colorization task. It comprises a sequence of layers, with each layer encompassing convolutional operations, optional normalization, and activation functions. The design allows customization of filter counts and down-sampling layers. The Patch Discriminator effectively evaluates the authenticity of colorized images at a patch level. It leverages LeakyReLU activations and convolutional layers for real versus generated image discrimination. This discriminator plays a pivotal role in the adversarial training process, enhancing the colorization model's performance and the quality of colorized outputs

MODULE 4 : GAN LOSS

INPUT:

The CGAN loss module requires specifying the GAN mode, which can be set as 'vanilla' or 'lsgan', along with labels for real and fake data (default values: 1.0 for real, 0.0 for fake).

OUTPUT:

It produces the CGAN loss, representing the adversarial loss incurred in CGAN training.

IMPLEMENTATION:

Based on the chosen GAN mode, it employs binary cross-entropy loss ('vanilla') or mean squared error ('lsgan'). The get_labels method determines the target labels for real or fake data. Upon invocation, the module calculates and returns the loss by comparing predicted values with target labels. This GAN loss module is crucial for effective CGAN training.

MODULE 5 : TRAINING THE MODEL

INPUT:

For training, key inputs include the generator network, learning rates, beta parameters, and the L1 loss weight.

OUTPUT:

The primary output is a well-trained colorization model, with tracked loss metrics for performance evaluation.

IMPLEMENTATION:

This module sets up the training process, configuring the generator network architecture, optimizer parameters, and L1 loss weight. A structured training loop is established to process data for a defined number of epochs. The module tracks GAN and L1 losses, offering insights into training progress. Visualizing model outputs provides real-time assessment of colorization quality, and training continues until the specified number of epochs is completed.

RESULTS AND DISCUSSION

Results after 200th Iteration:

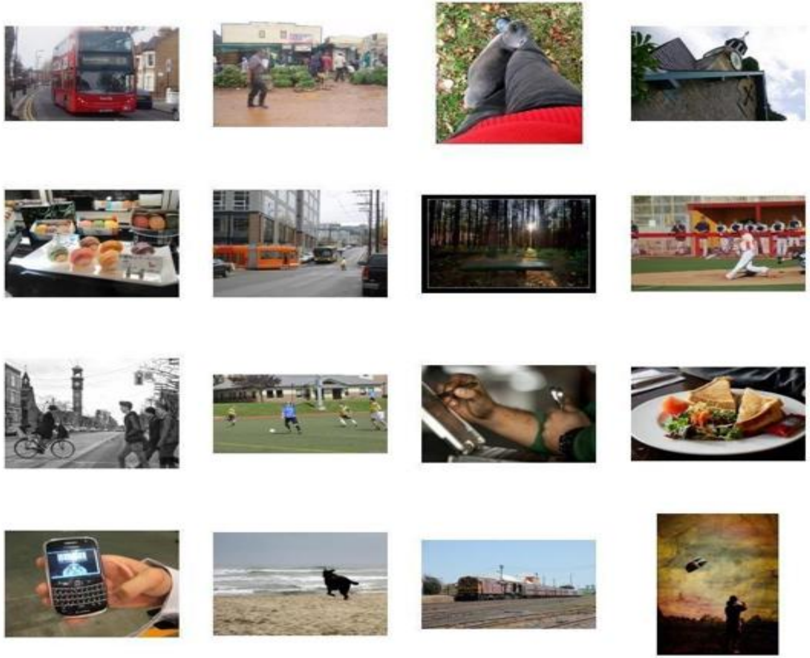
Metrics Visualization:

```
model initialized with norm initialization
model initialized with norm initialization
95% 473/500 [3:58:04<13:37, 30.28s/it]

Epoch 1/100
Iteration 200/500
loss_D_fake: 0.45000
loss_D_real: 0.46979
loss_D: 0.45990
loss_G_GAN: 1.62204
loss_G_L1: 9.64994
loss_G: 11.27198
```

RESULTS AND DISCUSSION

Results after 400th Iteration:



Metrics Evaluation:

```
Epoch 1/100
Iteration 400/500
loss_D_fake: 0.47620
loss_D_real: 0.49065
loss_D: 0.48342
loss_G_GAN: 1.55394
loss_G_L1: 9.96936
loss_G: 11.52330
```

REFERENCES

[1] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou and Alexei A. Efros. (2016), “Image-to-Image Translation with Conditional Adversarial Networks”, arXiv:1611.07004v1 [cs.CV] .

[2] Kangning du and Changtong Liu. (2021), “Double-Channel Guided Generative Adversarial Network for Image Colorization”, IEEE Access PP(99):1-1 DOI:11.09/Access.2021.3055575.

[3] K.Nazeri and E. Ng. (2020), “Image Colorization with Generative Adversarial Networks”, University of Ontario Institute of Technology Ontario, Canada.

[4] Harshith J and Chandan Kumar M. (2021), “Recovering Old or Damaged Images using GAN”, International Journal of Advanced Research in Computer and Communication Engineering Vol. 10, Issue 5,DOI 10.17148/IJARCCE.2021.105173.

[5] Yi Wang, Menghan Xia, Lu Qi, Jing Shao, and Yu Qiao. (2021), “PalGAN: Image Colorization with Palette Generative Adversarial Networks”, Shanghai AI Laboratory, Tencent AI Lab and Sense Time Research

[6] T. Sai Srinivas, Vemuri Harshitha. (2023), ”Colorizing black and white images using conditional GAN”, International Research Journal of Modernization in Engineering Technology and Science , Volume:05 ,Issue:04, e-ISSN: 2582-5208

REFERENCES

WEB LINKS:

https://www.researchgate.net/publication/310610633_Image-to-Image_Translation_with_Conditional_Adversarial_Networks

https://www.researchgate.net/publication/348887683_Double-Channel_Guided_Generative_Adversarial_Network_for_Image_Colorization

https://www.arxiv-vanity.com/papers/1803.05400/

https://ijarcce.com/wp-content/uploads/2021/06/IJARCCE.2021.105173.pdf

https://www.ecva.net/papers/eccv_2022/papers_ECCV/papers/136750268.pdf

https://www.irjmets.com/uploadedfiles/paper//issue_4_april_2023/35935/final/fin_irjmets1681479119.pdf

https://towardsdatascience.com/colorizing-black-white-images-with-u-net-and-conditional-gan-a-tutorial-81b2df111cd8

DATASET LINK:

https://www.kaggle.com/datasets/awsaf49/coco-2017-dataset/code