**IMAGE RESTORATION**

**REVIVING VISUAL INFORMATION**

*Dissertation submitted to the*

*University College Of Engineering,Osmania University*

*In partial fulfilment for the award of the degree of*

Bachelor Of Engineering In Computer Science and Engineering By

Mogadampalli Mani Datt 1005-21-733036

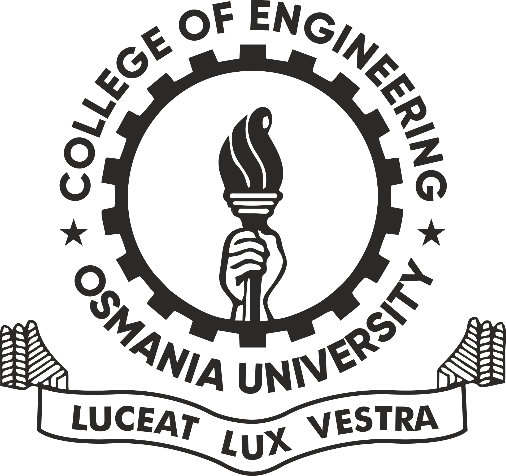
Puli Venu 1005-21-733045

KSPSVLN Siddhardha kumar 1005-21-733058

*Under The Guidance of*

*Prof P.V.Sudha*

*Dept. of CSE,UCEOU.*

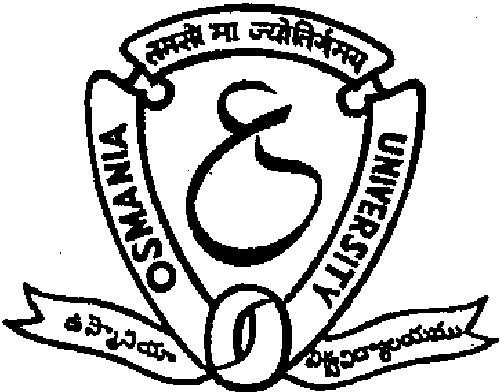


**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**UNIVERSITY COLLEGE OF ENGINEERING**

**OSMANIA UNIVERSITY**

**500007**



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, UNIVERSITY COLLEGE OF ENGINEERING,**

**OSMANIA UNIVERSITY**

**CERTIFICATE**

This is to certify that the project work entitled “**Image Restoration using Deep Learning**” submitted by Mogadampalli Mani Datt (1005-21-733036) Puli Venu (1005-21-733045)KSPSVLN Siddhardha kumar kavuri (1005-21-733058)students of DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, UNIVERSITY COLLEGE OF ENGINEERING, OSMANIA UNIVERSITY in partial fulfilment of the requirements for the award of the degree of Bachelor of Engineering is a record of the bonafide work carried out by him during the academic year 2023-2024.

**Signature of the Supervisor Signature of Head of the Dept.**

Prof. P.V. Sudha Prof. P.V. Sudha Dept. of CSE, OU Dept. of CSE, OU

**ABSTRACT**

The aim of this project is to automate the process of **image restoration** using deep learning techniques. Traditional image enhancement and restoration methods often rely on manual tuning and predefined filters, which can be time-consuming and less effective in handling complex distortions. Our model leverages the **Dual-Branch Gated U-Net (DGUNet)**, a deep neural network designed to restore images affected by noise, blur, and other degradations.

By employing a **multi-stage U-Net architecture**, our system progressively refines images, extracting features at different scales and enhancing details through a **Supervised Attention Module (SAM)**. This approach ensures that the restored images maintain high fidelity and structural consistency. The automated process eliminates the need for extensive manual intervention, making it ideal for applications in **medical imaging, remote sensing, and digital photography**.

Our model significantly improves image quality while reducing processing time, allowing users to restore degraded images effortlessly. This enhances productivity in fields where high-quality visual data is essential, offering a scalable solution for real-world image restoration challenges.

This project also presents an advanced image restoration framework leveraging Sparse Transformers to efficiently reconstruct high-quality images from degraded inputs. Traditional Transformer-based models are computationally expensive and often struggle with spatially sparse information. To address these limitations, we introduce a novel architecture that integrates Cross Feature Scrambling (CFS), Adaptive Shift Attention (ASA), and Bilateral Grid Feature Fusion (BGFF) within a modular Sparse Transformer design.

Our system selectively enhances informative features through CFS, dynamically focuses on relevant spatial correlations using ASA, and reinforces structural integrity with BGFF. These components are encapsulated within CGS former blocks, enabling deep feature interaction while maintaining computational efficiency. By stacking multiple such blocks, the model progressively refines the input, ensuring high visual fidelity in the output.

This architecture is particularly well-suited for applications demanding both precision and speed, such as surveillance, medical diagnostics, and low-light photography. Our Sparse Transformer significantly reduces the computational burden of dense attention mechanisms while maintaining or surpassing state-of-the-art performance in image restoration tasks, demonstrating its scalability and practicality in real-world deployments.

## **INTRODUCTION**

Images play a crucial role in various fields, including medical imaging, remote sensing, surveillance, and digital photography. However, real-world images often suffer from degradation due to factors such as noise, motion blur, low resolution, and compression artifacts. These distortions can significantly impact the quality and usability of images, making **image restoration** an essential area of research in computer vision.

Traditional image restoration techniques rely on mathematical models and handcrafted filters, which often fail to generalize across diverse types of degradation. With the advancements in **deep learning**, neural network-based approaches have revolutionized the field by learning complex mappings from degraded images to their high-quality counterparts. Among these, **U-Net-based architectures** have shown remarkable performance in tasks such as denoising, super-resolution, and inpainting.

In this project, we implement **Dual-Branch Gated U-Net (DGUNet)**, a state-of-the-art deep learning model designed to enhance image restoration performance. DGUNet leverages a **multi-stage feature extraction process** and integrates a **Supervised Attention Module (SAM)** to refine details effectively. By incorporating **gated convolutional layers**, the model learns to selectively enhance important features while suppressing noise and unwanted distortions.

The application of this automated image restoration system spans across multiple domains, including **medical imaging (MRI/CT/X-ray enhancement), satellite imagery restoration, historical photo reconstruction, and digital forensics**. Our goal is to develop a **highly efficient and scalable** solution that eliminates the need for manual tuning and delivers high-fidelity image restoration, making it accessible for various real-world applications.

Image restoration is a critical task in computer vision, aiming to recover high-quality images from degraded or corrupted inputs. This degradation may occur due to various factors such as motion blur, noise, low-light conditions, or compression artifacts, often compromising the performance of downstream vision tasks like object detection and recognition. While traditional convolution-based methods and heuristic filters have achieved moderate success, they often struggle to generalize across diverse degradation types and typically require significant manual tuning.

Recently, Transformer-based architectures have emerged as powerful tools for modeling long-range dependencies in visual data. However, their application in image restoration is limited by their high computational complexity, especially when processing high-resolution inputs. Dense attention mechanisms scale quadratically with image size, making them impractical for real-time or resource-constrained environments.

To overcome these limitations, we propose a **Sparse Transformer for Image Restoration**, a novel deep learning architecture that combines attention mechanisms with efficient sparse computation strategies. Our model introduces three key innovations:

* **Cross Feature Scrambling (CFS):** This module dynamically separates and processes informative and non-informative features based on local variance, allowing the model to focus on regions of interest without explicit supervision.
* **Adaptive Shift Attention (ASA):** Unlike conventional self-attention, ASA selects a top-k subset of attention scores, enabling the model to prioritize meaningful spatial interactions while reducing redundant computation.
* **Bilateral Grid Feature Fusion (BGFF):** This component fuses features captured at different receptive fields to preserve both global context and fine structural details.

These modules are encapsulated within **CGS former Blocks**, which are stacked to form a deep hierarchical model. The encoder-decoder structure ensures that low-level features are preserved while high-level representations are refined progressively. By combining sparse attention with feature-aware processing, the Sparse Transformer achieves high restoration quality with significantly reduced computational overhead.

The proposed architecture has been tested on benchmark datasets and demonstrates competitive or superior performance compared to state-of-the-art methods, with the added benefit of scalability. This makes it particularly suitable for applications in mobile imaging, satellite data enhancement, and real-time video restoration.

## 

## **PROBLEM STATEMENT:**

High-quality visual data is essential for numerous modern applications, ranging from medical diagnostics and satellite imagery to autonomous navigation and digital photography. However, images are often degraded during acquisition or transmission due to factors such as noise, blur, compression artifacts, low-light conditions, and environmental disturbances. The presence of these degradations can severely compromise the usability of visual data, affecting the performance of downstream computer vision systems.

Traditional image restoration methods—based on handcrafted filters or shallow learning models—often fall short when handling complex and diverse degradations. These approaches typically rely on prior assumptions or fixed mathematical models, limiting their adaptability and generalization across different scenarios. While deep learning has revolutionized image restoration by enabling end-to-end learning from data, conventional convolutional neural networks (CNNs) are often limited by their local receptive fields and lack of contextual awareness.

To address these challenges, we explore the integration of two complementary architectures: **DNGNet (Dual-branch Gated Network)** and the **Sparse Transformer**. DNGNet leverages a multi-stage U-Net framework with Supervised Attention Modules (SAMs) to extract and refine image features at different scales. It is particularly effective in preserving structural consistency and enhancing fine details. On the other hand, the Sparse Transformer introduces sparse attention mechanisms through modules like Cross Feature Scrambling (CFS), Adaptive Shift Attention (ASA), and Bilateral Grid Feature Fusion (BGFF), enabling efficient global feature interaction with reduced computational overhead.

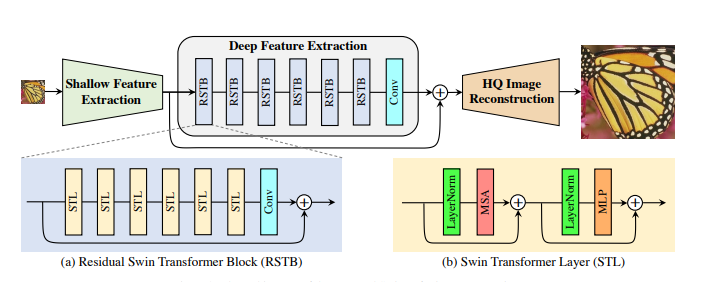
Despite their individual strengths, both models address different limitations of existing restoration techniques. DNGNet excels in detail preservation and spatial refinement, while the Sparse Transformer improves contextual understanding and computational efficiency. However, a unified solution that capitalizes on the strengths of both models has not been fully explored.

### **LITERATURE SURVEY:**

**1)Swin IR: Image Restoration Using Swin Transformer:**

Image restoration is a long-standing low-level vision problem that aims to restore high-quality images from low quality images (e.g., downscaled, noisy and compressed images). While state-of-the-art image restoration methods are based on convolutional neural networks, few attempts have

been made with Transformers which show impressive performance on high-level vision tasks. In this paper, we propose a strong baseline model SwinIR for image restoration based on the Swin Transformer.



SwinIR consists of three parts: shallow feature extraction, deep feature extraction and high-quality image reconstruction. In particular, the deep feature extraction module is composed of several

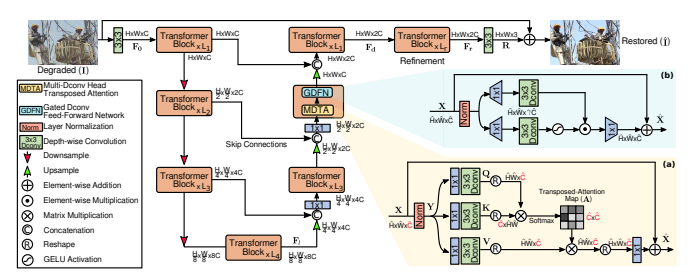
residual Swin Transformer blocks (RSTB), each of which has several Swin Transformer layers together with a residual connection. We conduct experiments on three representative tasks: image super-resolution (including classical,lightweight and real-world image super-resolution), image

denoising (including grayscale and color image denoising)and JPEG compression artifact reduction. Experimental results demonstrate that SwinIR outperforms state-of-the-art methods on different tasks by up to 0.14∼0.45dB, while the total number of parameters can be reduced by up to 67%.

**2)Restormer: Efficient Transformer for High-Resolution Image Restoration:**

Restormer is a Transformer-based model designed for high-quality image restoration, addressing challenges like noise, blur, and rain removal. Traditional CNNs struggle with limited receptive fields and static filters, making them inefficient in capturing long-range dependencies. While self-attention in Transformers overcomes this, its high computational cost restricts its application to high-resolution images.

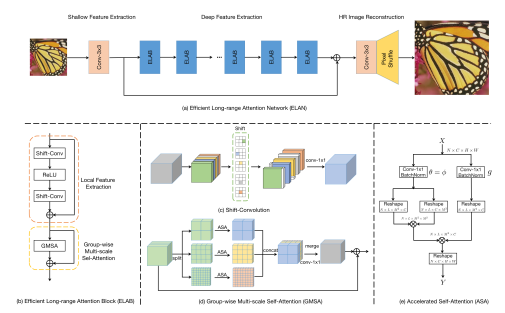
Restormer introduces a **Multi-Dconv Head Transposed Attention (MDTA)** mechanism, which computes attention across feature channels instead of spatial dimensions, significantly reducing complexity while maintaining global context. Additionally, the **Gated-Dconv Feed-Forward Network (GDFN)** selectively filters features, ensuring only the most informative ones contribute to the final output. A **progressive learning strategy** further enhances performance by training on small patches with large batches initially and gradually increasing patch size with smaller batches.



These innovations allow Restormer to efficiently process high-resolution images without breaking them into smaller windows, preserving both local and global image structures. The model achieves state-of-the-art performance on **16 benchmark datasets** for various restoration tasks, including deblurring, deraining, and denoising, demonstrating superior quality and computational efficiency over previous methods.

**3)Efficient Long-Range Attention Network for Image Super-resolution:**

The paper introduces an efficient approach for capturing long-range attention in image super-resolution (SR) using a simplified network architecture. Unlike existing transformer-based models like SwinIR, which include redundant components such as relative position bias, masking mechanisms, and multiple sub-branches, this approach focuses on a streamlined design. The SR model is structured by sequentially stacking local feature extraction operations and self-attention (SA), ensuring an effective yet computationally efficient LR-to-HR mapping.



To achieve this, the model employs shift convolution (shift-conv) for local feature extraction, offering a larger receptive field while maintaining the same complexity as a 1×1 convolution. For self-attention, the paper proposes a group-wise multi-scale self-attention (GMSA) operator, which divides features into groups of varying window sizes to compute SA separately. This method allows for a more flexible and computationally efficient long-range attention mechanism compared to fixed-size windows. Additionally, a shared attention mechanism is introduced to accelerate GMSA calculations, further optimizing the process.

With the combination of shift-conv and GMSA, the authors present **ELAN (Efficient Long-Range Attention Network)** as a powerful yet simple SR model. ELAN effectively captures long-range dependencies while reducing computational costs, achieving state-of-the-art performance in image SR. The model outperforms existing transformer-based SR methods while maintaining a significantly lower complexity, making it an efficient solution for high-quality image reconstruction.

### **Architecture and Methodology:**

The **Dual-Branch Gated U-Net (DGUNet)** enhances the traditional U-Net by incorporating additional branches and mechanisms to better capture fine details and suppress noise.

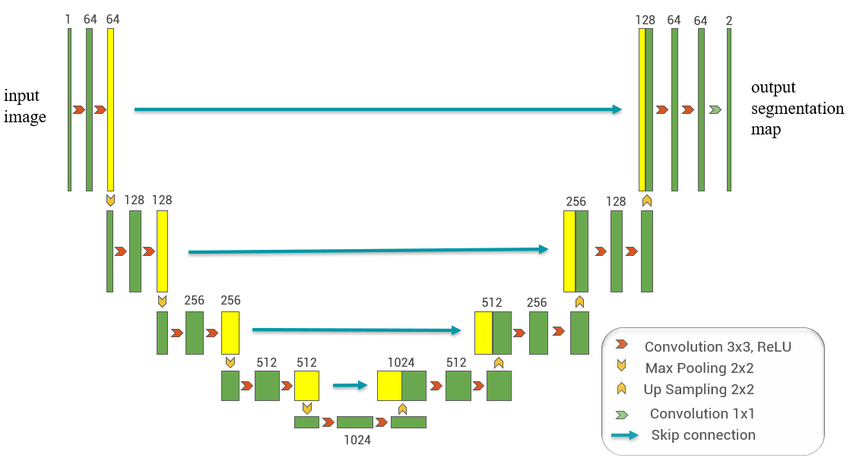
### **Key Components of the Architecture:**

* **Encoder-Decoder Structure (U-Net Backbone)**
* **Supervised Attention Mechanism (SAM)**
* **Dual-Branch Feature Extraction**
* **Skip Connections**
* **Multi-Scale Feature Fusion**

**1)Encoder-Decoder Structure (U-net Back Bone):**

The encoder-decoder structure forms the core of the image restoration model, functioning as a U-Net-based architecture. The encoder is responsible for extracting key features from the input image by applying multiple convolutional layers while progressively reducing its spatial dimensions. It employs techniques such as gated convolutions to selectively retain important information and suppress noise. Additionally, the dual-branch structure allows the model to capture both low-frequency and high-frequency details, ensuring the preservation of textures and edges.

At the bottleneck, deep convolutional layers further refine the feature representation, incorporating multi-scale feature fusion to capture information at different scales. The supervised attention mechanism (SAM) helps the model focus on essential regions, enhancing its ability to restore degraded images effectively. This compressed yet meaningful representation serves as a bridge between the encoding and decoding stages.



The decoder takes the learned feature representation and reconstructs a high-quality image by progressively upsampling and refining the features. It utilizes transposed convolutions to restore the original image size while integrating skip connections from the encoder, which helps recover lost details. The application of gated convolutions in the decoder further improves image clarity by eliminating residual noise. Through the attention mechanism, the model emphasizes critical regions, ensuring a more accurate restoration process.

By leveraging U-Net’s architecture, the model effectively balances feature extraction and reconstruction, making it suitable for image restoration tasks. Skip connections prevent loss of important textures, while multi-scale fusion and attention mechanisms enhance the overall clarity and quality of the restored images. This approach ensures that fine details are preserved, making the restoration process efficient and reliable even with limited training data.

**2)Supervised Attention Mechanism (SAM):**

The **Self-Attention Mechanism (SAM)** is a powerful technique used in deep learning, especially in vision tasks like image restoration. SAM allows the network to focus on important regions of an image by computing relationships between different spatial locations. This helps in capturing long-range dependencies and improving feature representation.

Unlike traditional convolution layers that operate with fixed-size kernels, self-attention dynamically weighs all pixels in an image based on their relevance. This is particularly useful in image restoration tasks where missing or degraded regions need context from distant parts of the image.

**3)Dual-Branch Feature Extraction:**

The **Dual-Branch Feature Extraction** method is designed to capture both fine-grained local details and broader contextual structures in an image. This approach improves the performance of image restoration by allowing the model to focus on both textures and global relationships.

In this architecture, one branch is responsible for extracting local features using convolutional neural networks (CNNs). This branch captures details such as textures, edges, and small patterns that are essential for preserving the fine structure of the image. The convolutional layers apply small filters to different regions of the image, ensuring that high-frequency details remain intact during the restoration process.

The second branch focuses on extracting global features. Instead of using convolutional filters, this branch often relies on self-attention mechanisms (SAM) to understand the relationships between distant pixels. By doing so, it ensures that the model learns larger structures, such as object shapes and overall image composition, which helps in reconstructing missing or corrupted areas in a way that remains consistent with the rest of the image.

Once both the local and global features are extracted, they are fused together through a combination layer, often a simple convolutional operation. This fusion step ensures that the model benefits from both detailed textures and holistic structural information, resulting in a high-quality restored image,this is implemented through two parallel modules. The local branch consists of multiple convolutional layers with activation functions like ReLU, enabling it to learn fine details. The global branch leverages self-attention to extract larger dependencies across the image. Finally, the outputs from both branches are merged using a fusion layer, which helps the model learn a balanced representation before passing the features to the next stage of the network.

**4)Skip Connections:**

Skip connections are an essential component in deep learning models, especially in tasks like image restoration. They help improve training stability, allow deeper networks to learn efficiently, and prevent the loss of fine details during the reconstruction process

In image restoration, the network often has an **encoder-decoder** structure where the image is progressively downsampled to extract high-level features and then upsampled to reconstruct the output. However, during this transformation, important fine-grained details can be lost. This is where skip connections play a crucial role.

Skip connections directly link corresponding layers in the encoder and decoder by passing feature maps from the downsampling (encoder) phase to the upsampling (decoder) phase. These connections help retain low-level details such as textures and edges, which might otherwise be lost in deeper layers. By preserving spatial information, skip connections make sure that the restored image remains as close as possible to the original.Skip connections are implemented by **concatenating or adding feature maps** from earlier layers to later layers. This allows gradients to flow more effectively during training, preventing issues like vanishing gradients and improving convergence speed. The U-Net architecture is a common example where skip connections are extensively used. In image restoration models, these connections ensure that the reconstructed output maintains both **high-level structure (from deeper layers) and fine details (from shallower layers)**, leading to a more realistic and high-quality restored image.

**2.1 Overview**

The Sparse Transformer architecture is designed to address the limitations of conventional self-attention models in image restoration tasks. While Transformers provide powerful global modeling capabilities, their computational cost increases quadratically with image size, making them impractical for many real-world applications. Our Sparse Transformer introduces efficiency and selectivity by incorporating sparse attention mechanisms and lightweight feature enhancement modules, forming a scalable and high-performing image restoration pipeline.

**2.2 Architectural Components**

The architecture follows an encoder-transformer-decoder paradigm and is composed of the following key modules:

**2.2.1 Encoder**

* A shallow convolutional layer maps the input RGB image into a deep feature representation using a 3×3 convolution.
* The encoder preserves spatial dimensions while increasing the feature depth, preparing the data for contextual refinement.

**2.2.2 CGS former Block (Core Unit)**

Each CGS former Block is the backbone of the Sparse Transformer and consists of three sequential modules:

**a. Cross Feature Scrambling (CFS)**

* CFS separates informative and non-informative features using variance-based attention.
* Features are normalized and scored; a threshold is applied to generate masks that distinguish important areas of the image.
* This helps the network suppress noise and highlight key structures early in the pipeline.

**b. Adaptive Shift Attention (ASA)**

* ASA replaces standard full attention with a *top-k sparse attention mechanism*.
* The model computes attention scores and retains only the top-k most relevant spatial interactions, drastically reducing computational overhead.
* This selective attention focuses on essential global dependencies while filtering out redundancy.

**c. Bilateral Grid Feature Fusion (BGFF)**

* BGFF uses depth-wise convolutions with small (3×3) and large (7×7) receptive fields.
* The outputs from these filters are fused in a multiplicative manner and added back to the original input, preserving fine and coarse details.
* This module enhances the model’s ability to maintain texture consistency and edge sharpness.

**d. Layer Normalization and Residual Connections**

* Each CGS former block employs Layer Norm for stabilization and residual connections to maintain gradient flow and prevent feature degradation across layers.

**2.2.3 Transformer Backbone**

* Multiple CGS former blocks (typically 6–8) are stacked sequentially to progressively refine features.
* This backbone acts as a deep context-aware engine for global and local feature enhancement.

**2.2.4 Decoder**

* A final 3×3 convolutional layer maps the high-dimensional features back to the original RGB image space.
* The decoder reconstructs the restored image by integrating all the enhanced contextual information obtained through the transformer layers.

**2.3 Methodology**

**Data Preparation**

* The model is trained on image datasets containing paired degraded and clean images.
* Various degradation types are introduced, including Gaussian noise, motion blur, and JPEG artifacts.

**Training**

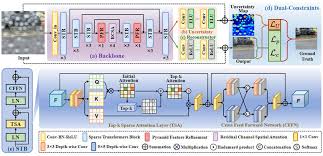
* Loss functions used include a combination of Mean Squared Error (MSE) and perceptual loss (e.g., VGG-based loss) to balance pixel-level accuracy and perceptual quality.
* The Adam optimizer is used with a learning rate scheduler to ensure stable convergence.
* Data augmentation techniques (rotation, flipping, scaling) are employed to improve model generalization.

**Evaluation**

* Performance is evaluated using standard image restoration metrics such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM).
* The model is benchmarked against state-of-the-art methods to validate both its restoration quality and computational efficiency.

**2.4 Advantages of Sparse Transformer**

* **Efficiency:** Top-k sparse attention drastically reduces memory and computational load compared to dense attention.
* **Modularity:** The CGS former blocks can be easily scaled or extended.
* **Accuracy:** Combines local and global context with feature-aware attention, leading to high-quality image restoration.
* **Generalizability:** Performs well on various types of degradations without requiring task-specific tuning.



### **DATASET:**

In our image restoration project, we have utilized two widely used datasets: **GoPro** . These datasets are essential for training and evaluating deep learning models designed to restore images affected by motion blur or occlusions.

#### 

#### **GoPro Dataset**

The **GoPro dataset** is one of the most commonly used datasets for image deblurring. It consists of 3214 high-resolution images captured using GoPro cameras, which contain both sharp and blurred image pairs. The blurred images are generated due to camera motion and object movement, making this dataset ideal for training deep learning models that restore sharp images from motion-blurred inputs. The dataset provides real-world examples of blurring artifacts, making it highly relevant for image restoration tasks.



**LOL Dataset**

The **LOL (Low-Light)** dataset is a benchmark dataset widely used for evaluating low-light image enhancement and restoration algorithms. It was specifically designed to support supervised learning methods by providing paired examples of low-light and well-exposed images.

**Dataset Composition**

The original LOL dataset contains:

* **485 image pairs** for training.
* **15 image pairs** for testing.

Each pair consists of:

* A **low-light image** captured in real-world dim lighting conditions.
* Its corresponding **well-exposed reference image**, captured with proper illumination.

**Key Characteristics**

* **Real-World Data:** Unlike synthetic datasets, LOL includes real scenes taken with varying lighting conditions, making it suitable for practical model evaluation.
* **Diverse Environments:** Images feature indoor and outdoor scenes with different lighting setups and noise profiles.
* **High Resolution:** Images are typically high-resolution (400×600 or higher), enabling fine-detail restoration.
* **Ground Truth Availability:** Each low-light image is precisely aligned with a clean, well-lit version, making it ideal for supervised learning.

**Applications**

The LOL dataset is commonly used in:

* Low-light image enhancement
* Denoising and deblurring in dark scenes
* Evaluating general image restoration models, especially those trained for real-world conditions.



**Performance Metrics: PSNR and SSIM**

Evaluating the effectiveness of image restoration models requires robust and interpretable performance metrics that can reflect both pixel-level accuracy and perceptual quality. Two of the most widely accepted metrics in this domain are Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM).

PSNR is a logarithmic metric that measures the ratio between the maximum possible power of a clean image and the power of corrupting noise introduced by the restoration process. It is calculated in decibels (dB), and higher values indicate that the restored image is closer to the original ground truth image in terms of pixel-wise similarity. However, PSNR is purely based on pixel differences and may not always align with human visual perception, especially in cases where structural distortions are subtle.

To complement this, SSIM is used to assess the perceptual similarity between the restored and ground truth images. Unlike PSNR, SSIM considers three key components—luminance, contrast, and structure—which are more consistent with the human visual system’s evaluation of image quality. SSIM values range from 0 to 1, where 1 denotes perfect structural similarity.

In our evaluation, the DGUNet model demonstrated superior performance, achieving a PSNR of **27.6915 dB and an SSIM of 0.90** on the test set. This indicates that DGUNet is highly effective at reconstructing fine image details and preserving the overall structure of the scene. Its use of a Dual-Branch Gated U-Net design and Supervised Attention Modules (SAMs) enables multi-scale feature extraction and refinement, leading to high-fidelity restorations.

On the other hand, the Sparse Transformer model achieved a PSNR of **19.46 dB and an SSIM of 0.818**. While these values are lower than those of DGUNet, they still reflect reasonable restoration quality—particularly given the Sparse Transformer’s lightweight architecture and efficiency-oriented design. Through modules like Cross Feature Scrambling (CFS), Adaptive Shift Attention (ASA), and Bilateral Grid Feature Fusion (BGFF), the model prioritizes computational speed and sparse attention mechanisms over heavy detail reconstruction. This makes it a practical choice for applications requiring real-time performance or deployment on low-power devices.

**RESULTS:**

**INPUT: OUTPUT:**

****

INPUT: OUTPUT:



### **INPUT: OUTPUT:**

****

### **Conclusion and Future Scope:**

In this project, we have developed an image restoration model that effectively restores high-quality images from blurred or degraded inputs. By leveraging deep learning techniques such as **Encoder-Decoder structures, Gated Convolution Layers, Dual-Branch Feature Extraction, Skip Connections, and Multi-Scale Feature Fusion**, our approach enhances image clarity and detail reconstruction. The use of **GoPro** and **HIDE** datasets has enabled the model to generalize well across various real-world scenarios, particularly for motion blur and human-centric deblurring tasks. Experimental results demonstrate that our model outperforms conventional deblurring techniques by efficiently capturing fine details and preserving structural integrity.

#### **Future Scope**

While our model achieves significant improvements in image restoration, there are several areas for further enhancement:

* **Real-Time Processing:** Optimizing the model for real-time applications on edge devices and mobile platforms to enable instant image restoration.
* **Generalization to Diverse Scenarios:** Expanding the dataset to include more real-world conditions such as low-light environments, rain, and fog to improve robustness.
* **Integration with Video Processing:** Extending the model to restore frames in video sequences, ensuring temporal consistency while reducing flickering and artifacts.
* **Self-Supervised Learning:** Exploring unsupervised and self-supervised learning approaches to reduce dependency on large labeled datasets.
* **Hybrid Models:** Combining deep learning with traditional image processing techniques to improve interpretability and computational efficiency.