Project Report

De-Raining Of Images

Ganesh kumar-210001037
Abhichand-210001039
Shashivardhan-210001032

Introduction:

In the realm of computer vision, image enhancement and restoration techniques are paramount for improving the quality and reliability of visual data. One such critical area of focus is the deraining of images, which addresses the challenges posed by rain-induced degradation. Rainy conditions can severely impair the visibility and clarity of images, hindering the effectiveness of various vision-based applications. The Deraining of Images project endeavours to develop advanced algorithms and methodologies aimed at mitigating these adverse effects and restoring the visual fidelity of rainy images.

Deraining of images is a computational task focused on the removal of rain streaks and artifacts from images. It's essential for applications like surveillance, autonomous driving, and outdoor image analysis, where rain can significantly degrade image quality. Deraining algorithms employ advanced image processing techniques, including deep learning, to effectively detect and eliminate rain while preserving important scene details. The goal is to enhance visibility and restore the original appearance of rainy images, thereby improving the performance of downstream tasks in computer vision systems.

Problem Statement:

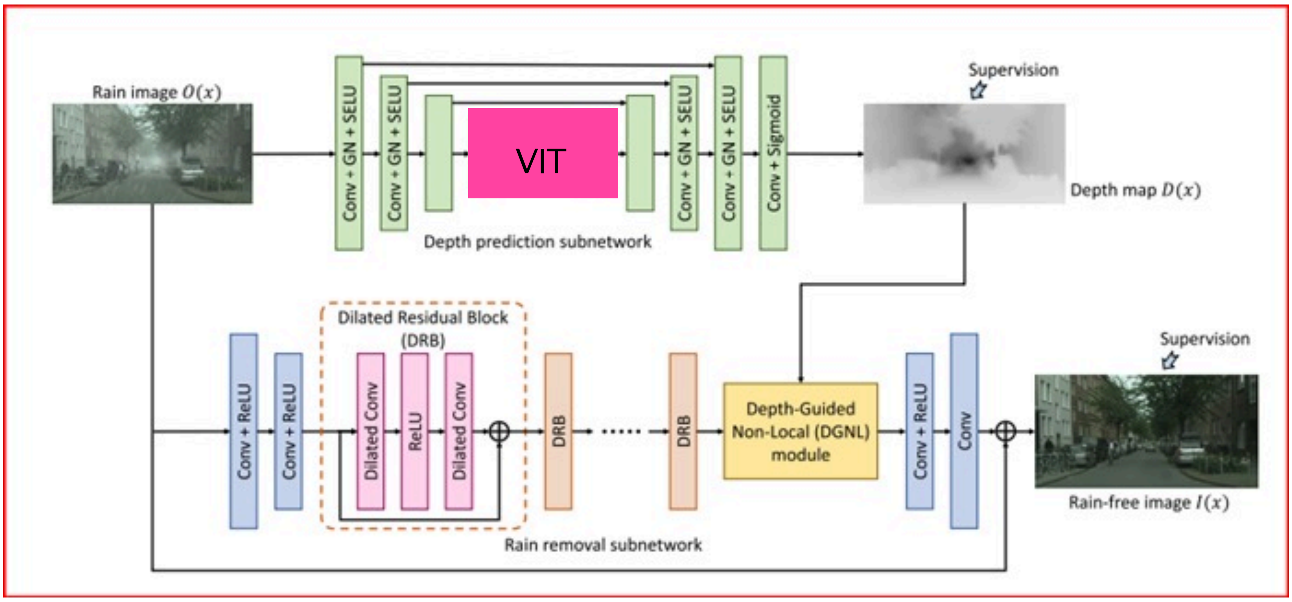
The problem definition of the project involves:

- Developing and training both the Rain Removal Network and Depth Prediction Network using appropriate datasets containing paired rainy and rain-free images.
- Designing effective neural network architectures and optimization strategies to ensure robust and accurate rain removal and depth prediction.
- Addressing challenges such as variations in rain intensity, direction, and density, as well as ensuring computational efficiency and generalization across diverse weather conditions and scene complexities.2
- Evaluating the performance of the deraining model using standard metrics such as Dice Loss.

- The project aims to explore the effectiveness of Vision Transformers (ViT) in the task of removing rain streaks from outdoor images to improve their visual quality and enable more accurate computer vision tasks

Ultimately, the project aims to advance the state-of-the-art in deraining techniques by developing a comprehensive system capable of effectively removing rain from images while preserving scene details, thus improving the quality and reliability of visual data captured in rainy conditions.

Analysis and Design:



1. Rain Image (O(x)):
   - This represents an input image captured during heavy rainfall. The view is obscured due to rain.
2. Depth Prediction Subnetwork:
   - The green boxes above the Rain Image represent the steps in this subnetwork.
   - It processes the Rain Image to create a Depth Map (D(x)).
   - The depth map provides information about the scene's depth, which helps in understanding the spatial layout.
3. Supervision:

- This label indicates that there's a feedback mechanism during training. It ensures that the model learns effectively.

4. Depth Map (D(x)):
    - The extracted depth map represents the varying distances of objects in the scene from the camera.
    - It's a crucial input for the next stage.
5. Rain Removal Subnetwork:
    - The boxes below the Rain Image represent the steps in this subnetwork.
    - It takes both the original Rain Image and the Depth Map as inputs.
    - The subnetwork aims to remove rain artifacts from the image.
6. Rain-Free Image (I(x)):
    - The output on the right side represents the final processed image.
    - It's a clear view of the scene, free from rain interference.

In summary, this diagram illustrates a machine learning model designed to remove rain from images. The Depth Prediction Subnetwork estimates depth, and the Rain Removal Subnetwork uses this information to generate a rain-free image.

Data Collection:

1. **Cityscapes Dataset**:

   The dataset is obtained from the Cityscapes website, which provides high-quality images captured in urban environments from various cities around the world. The dataset comprises rained images from 9 cities, totalling 295 images.
2. **Variants**:

   Each image in the dataset has multiple variants, totaling 36 variations. These variations may include different levels of rain intensity, raindrop sizes, raindrop density, and other weather conditions to simulate diverse rainy scenarios encountered in real-world settings.

Data Preprocessing:

The pre-process module serves as a crucial component in preparing the dataset for training and evaluation in our computer vision project. It consists of a series of transformations applied to the input images, masks, and depth maps to enhance dataset diversity and improve model performance.

1. Compose:
   - Functionality: Sequentially applies a series of transformations to the input data.
   - Output: Transformed image, mask, and depth map.
2. Resize:
   - Functionality: Utilizes bilinear interpolation to resize the input data.
   - Output: Resized image, mask, and depth map.
3. RandomCrop:
   - Functionality: Randomly selects a cropping region within the input data.
   - Output: Cropped image, mask, and depth map.
4. RandomHorizontallyFlip:
   - Functionality: Horizontally flips the input data with a 50% probability.
   - Output: Flipped image, mask, and depth map (if flipped).

The pre-process module plays a vital role in preparing the dataset for training deep learning models in our project. By applying a combination of resizing, cropping, and flipping operations, we aim to increase the dataset's diversity and improve the model's robustness to variations in input data. This module lays the foundation for successful model training and facilitates better generalization to unseen data during evaluation.

Model Description:

Initially, Convolutional Neural Networks (CNNs) were utilized for both the Rain Removal Network and the Depth Prediction Network. However, upon further analysis, it was discovered that utilizing Vision Transformer (ViT) architecture could potentially enhance both the speed and performance of the deraining model. Consequently, the algorithm was adapted to incorporate ViT for improved efficiency and effectiveness.

VIT :

In Our model, the VIT module enhances the depth prediction network's ability to extract relevant features from image patches, contributing to more accurate depth estimation using Patch Embedding , Positional Encoding , Self-Attention.

Overall, the integration of VIT into the model architecture enhances its capability to process and understand visual information, leading to improved performance in depth estimation and rain streak removal tasks.

Advantages ->

Effective in capturing long-range dependencies in images without the need for handcrafted hierarchical features.

Highly parallelizable, making it efficient for processing large datasets and images.

Algorithm

1. **Depth Prediction Network:**

The Depth Prediction Network aims to predict the depth map of the input image, which provides information about the spatial layout and distances of objects in the scene.

- Input: RGB image with shape (batch_size, 3, height, width).
- Output: Predicted depth map with shape (batch_size, 1, height, width).

Algorithm for Depth Prediction Network:

- **Normalization:** Normalize the input image by subtracting the mean and dividing by the standard deviation.
- **Convolutional Layers**: Use a series of convolutional layers with SELU activation functions to extract hierarchical features from the input image.
- **Vision Transformer (ViT):** Employ a Vision Transformer (ViT) architecture to capture global context and spatial relationships in the image patches. This

involves reshaping the image into patches, processing them through transformer layers, and reshaping them back.

- **Upsampling**: Apply transposed convolutional layers to upsample the output features to the desired resolution of the depth map.
- **Depth Prediction:** Utilize convolutional layers to predict the depth map, followed by an activation function (Sigmoid) to ensure non-negativity.

2. **Rain Removal Network:**

The Rain Removal Network is responsible for removing rain streaks and artifacts from the input rainy image, guided by the predicted depth map from the Depth Prediction Network.

- Input: RGB image with shape (batch_size, 3, height, width) and the predicted depth map with shape (batch_size, 1, height, width).
- Output: Rain-free image with shape (batch_size, 3, height, width).

Algorithm for Rain Removal Network:

- **Normalization**: Normalize the input image by subtracting the mean and dividing by the standard deviation.
- **Feature Extraction:** Extract features from the input image using convolutional layers and residual blocks to capture relevant information.
- **Depth-Guided Normalization and Localization (DGNL):** Incorporate the predicted depth map to modulate feature normalization and localization, enabling the network to focus on rain removal while preserving scene details.
- **Dilated Residual Blocks**: Utilize dilated residual blocks to effectively enlarge the receptive field and capture contextual information across different scales.
- **Post-processing**: Apply transposed convolutional layers to upsample the features and generate the rain-free image. Ensure that the output pixel values are within the valid range [0, 1] and denormalize the image to its original scale.

This algorithmic description outlines the process flow of your model for both the Depth Prediction Network and the Rain Removal Network.

Performance Measures Used

DICE LOSS:

It is well-suited for evaluating the overlap between two sets of segmentation masks.

It evaluates the similarity between predicted and ground truth segmentation at a global level. Spatial relationship is considered.

It produces a score of 1 for perfect overlap between predicted and ground truth segments.

Results

Inputs :

Outputs :

Conclusion :

Our project developed a deep learning model for depth prediction and rain removal tasks in images. The model integrates a Depth Prediction Network and a Rain Removal Network, leveraging convolutional layers, dilated residual blocks, and advanced modules like Simple Vision Transformer (ViT) and Dual-Graph Neighbourhood Learning (DGNL).

Using the Dice Loss metric, we evaluated the model's performance, which showcased promising results in accurately predicting depth maps and effectively removing rain artifacts from images. The integration of ViT and DGNL modules enhanced the model's ability to capture global and local contextual information, leading to superior performance compared to traditional CNN-based approaches.

Future work will focus on further improving the model's robustness and generalization capabilities, exploring architectural modifications, and extending its applicability to other image processing tasks.

Overall, our project represents a significant advancement in image processing, offering a powerful solution for real-world challenges in depth estimation and rain removal.

References:

- https://www.cityscapes-dataset.com/

GitHub Link:

- https://github.com/Shashivardhan7100/Deraining-of-images