भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

CS6140: Video Content Analysis

Under the Guidance of
Prof. C. Krishna Mohan
&
(TA) Kondapally Madhavi

Self Gated Memory Recurrent Network for Efficient Scalable HDR Deghosting

---

Authors:

Rishav KUMAR AI22MTECH12003@IITH.AC.IN

Contents

1 Paper Summary

1.1 Motivation

- Current deep learning-based HDR deghosting methods face limitations, often being trained for a fixed set of three low dynamic range (LDR) images, requiring re-training for sequences with different input image numbers.
- Shortcomings include a lack of training datasets for diverse sequence lengths, making the construction of large, labeled datasets for various scenarios a laborious and resource-intensive task.
- The research introduces a novel recurrent network-based HDR deghosting method, aiming to handle dynamic sequences of arbitrary length, eliminate the need for re-training with different input image quantities, and address the scarcity of training data for varying sequence lengths.

1.2 Problem Statement

The paper addresses the challenge of developing a scalable HDR deghosting method capable of fusing sequences of dynamic images with arbitrary lengths. Existing methods, including CNN-based approaches, often lack scalability and require re-training when handling variable-length input sequences. The authors propose a novel recurrent network-based HDR deghosting method that utilizes a new recurrent cell architecture, the Self-Gated Memory (SGM) cell, outperforming traditional LSTM cells in terms of parameters and runtime. Their approach achieves state-of-the-art performance across multiple datasets, demonstrating scalability to handle variable-length input sequences without the need for re-training. The key contribution lies in offering an efficient and scalable solution for HDR deghosting, addressing the limitations of existing methods.

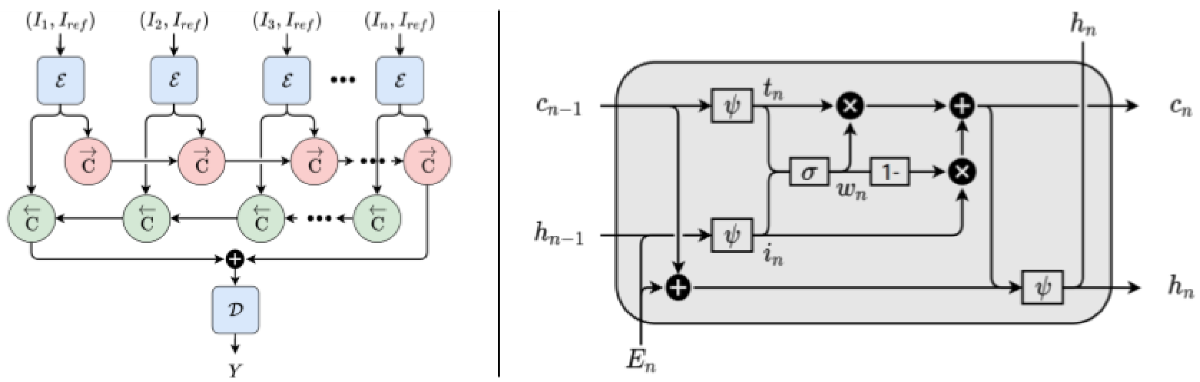

Figure 1: Architecture of the proposed approach. The current input data X_n concatenated with the reference data X_ref is passed as input to the encoder(E)

2 Reproduced Results

2.1 Dataset

The dataset contains a set of TIFF image stacks of multi-exposure image sequences that could be used for training and testing HDR de-ghosting methods. The dataset contains both the input stack and the ground truth HDR images.

The Kalantari dataset (USSD) and the Prabhakar dataset (ICCP).

2.2 Metrics

For evaluating the model, we have PSNR(peak signal-to-noise ratio) and MSE loss.

2.3 Results

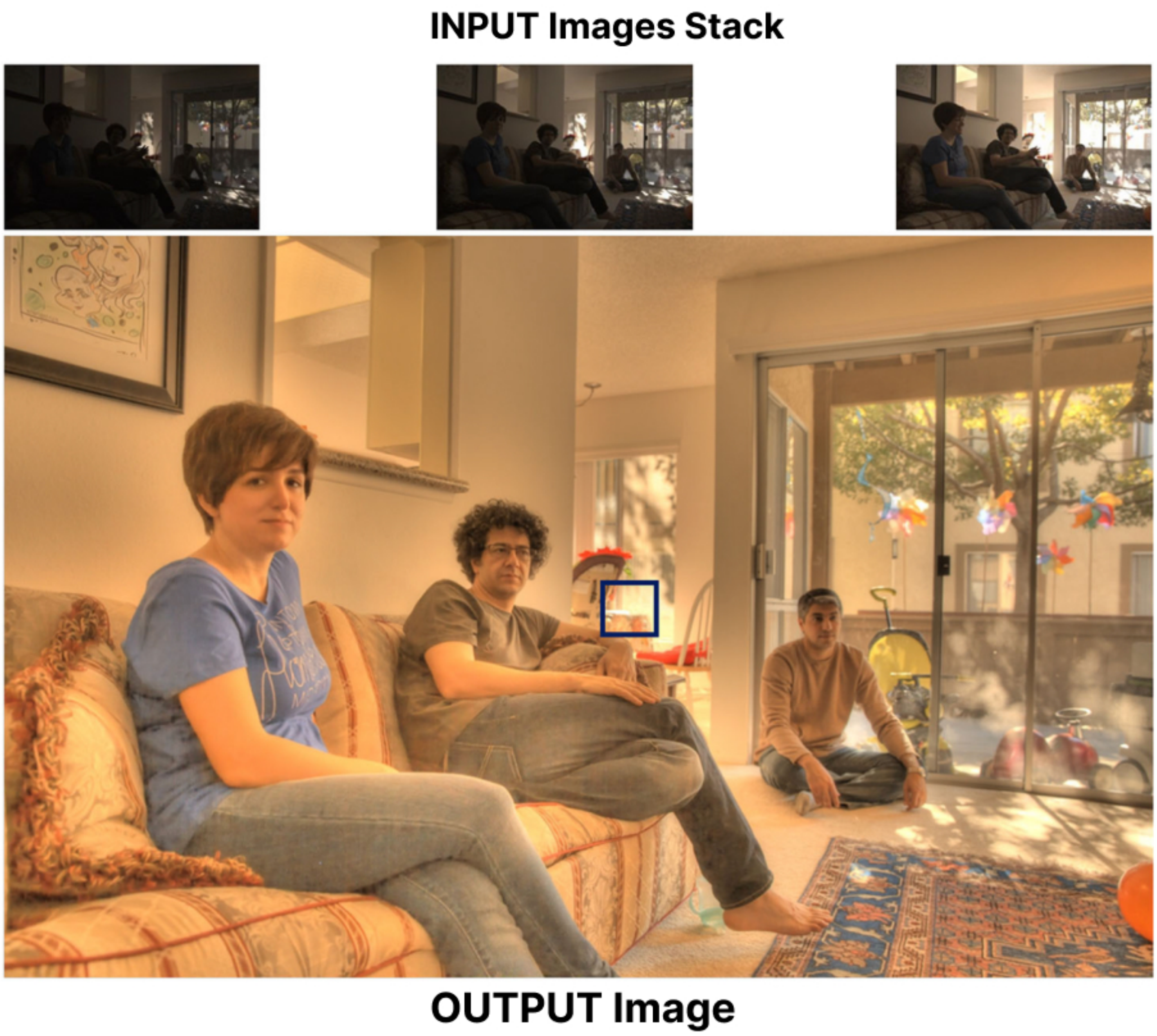

Figure 2: Sample result from Kalantari dataset

3 Novelty Implemented

3.1 Motivation

Navigating adverse weather conditions presents a significant challenge for computer vision algorithms, hindering their accuracy in object detection, tracking, and classification tasks. To address these shortcomings, researchers advocate for training these models on datasets specifically curated for adverse weather scenarios. This approach not only enhances the algorithms' adaptability to complex environments but also strengthens their overall resilience and reliability. Moreover, enabling seamless transitions across diverse adversities, such as from cloudy to rainy conditions, is paramount for developing robust AV perception models. Therefore, incorporating a preprocessing step to mitigate the impact of adverse weather scenes prior to feeding information to clear-weather-trained models is crucial for achieving optimal performance.

3.2 Objective

To remove the noise, such as fog and rain, from the Input image and generate a single image as close as the ground truth image. The method of HDR de-ghosting is successfully applied to the Domain Adaptation Problem.

3.3 Dataset

Image1 and Image2 from the AIWD6 Dataset, Cloudy to rainy Transition variant

3.4 Approaches

- **Input as series of Images : {0-9}.png & Ground Truth : Image1.png**
  Remarks: In this approach, the model is considering only Image 9.png as other images are redundant due to frames from the same angle. Using LSTM cell was not helpful when images are not time series and have redundant features with different perturbation levels.
  PSNR : 23.899384

- **Single Input Image1.png & Ground Truth : Image2.png**
  Remarks: This is giving better results as compared to the previous approach, as the image stack is not redundant.
  PSNR : 26.455849

3.5 Results

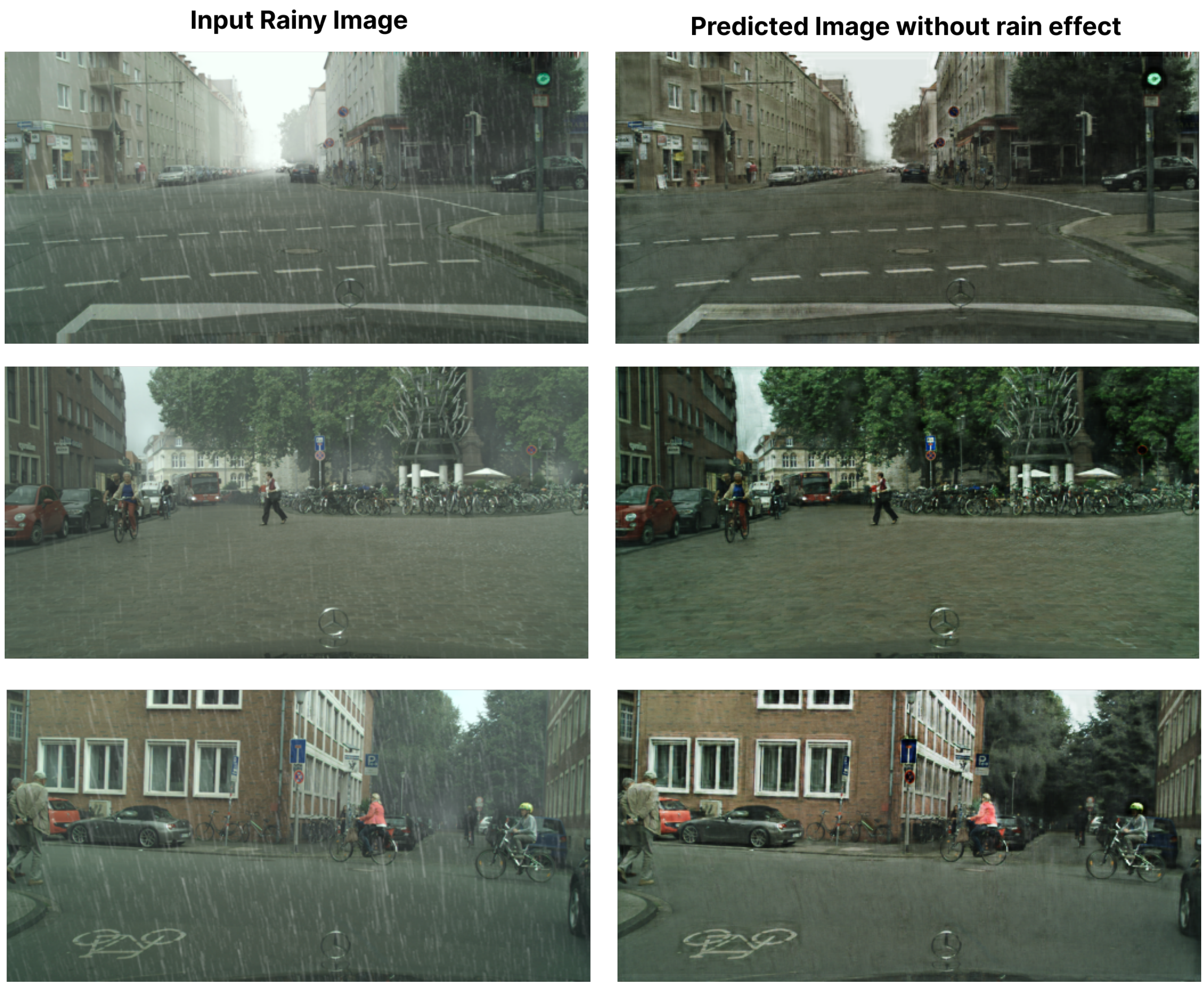

Figure 3: Removed the rain effect from the Images of AIWD6