IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

DEEP LEARNING AND OPTIMIZATION FOR HDR IMAGE RECONSTRUCTION

[1]Balla Kalpana Devi, [2]Dr.U.V. RatnaKumari

[1]M. Tech Student, [2]Professor,

[1]University College of Engineering, JNTUK, Kakinada, Andhra Pradesh,

[2]University College of Engineering, JNTUK, Kakinada, Andhra Pradesh

Abstract: Converting a low dynamic range (LDR) image to a high dynamic range (HDR) image produces an image that closely resembles the real world without the use of expensive equipment. Recent advances in deep learning have enabled the creation of realistic and intelligent HDR photos. In this work, we present a Deep learning method for segmenting bright and dark areas of input LDR images and reconstructing HDR images with similar real-world dynamic ranges. To create HDR images, the proposed multi-level deep learning network brightens bright areas and darkens dark areas, combining features with a wider range of brightness. Information about missing overexposed and underexposed areas is efficiently implemented by splitting the LDR image into bright and dark parts, resulting in a natural HDR image with colors and appearances similar to the extracted real features. The Firefly Optimization technique is considered to minimize the features that are extracted to obtain HDR images.

I.INTRODUCTION

Significant advances in image processing in recent years have increased the demand for techniques that can generate images with real-world resolution. The most widely used image enhancement technique is High Dynamic Range (HDR). In contrast, HDR works beyond the limits of typical camera sensors and can capture a wide dynamic range . As a result, most people have limited access to his HDR images.

LDR photo was taken with multiple exposures. Then stitch the photos together to create an HDR image. However, multiple photos cannot be taken at the same time, so stitching bracketed photos together will show ghosting artifacts that project the flow of time. Missing data ranges can also be found outside a given exposure range.

An Inverse Tone Mapping (ITM) algorithm was developed to generate HDR photos from individual LDR photos. On the other hand, conventional ITM methods are not sufficient to provide functionality equivalent to the inverse camera response function (CRF), and it is difficult to obtain parameters that satisfy all image situations. A convolutional neural network (CNN)-based MEF approach and a CNN-based ITM method were developed to solve this problem using deep learning techniques. By creating and merging multiple exposure images, CNN-based MEF algorithms can restore saturated pixels uniformly. I'm having trouble restoring saturated pixels that don't exist in the brightness range of a multi-exposure photo. CNN-based ITM algorithms have emerged to reconstruct HDR images from a single LDR image. A CNN-based ITM algorithm was developed to reconstruct an HDR image from a single LDR image. These methods map LDR images to HDR images via a CNN acting as an inverse CRF. Despite efforts to collect photos as close to the source as possible, the effective recovery of saturated pixels is still limited.

The number of saturated pixels in an HDR image is limited compared to the total number of pixels. Image reconstruction typically uses a loss function based on L1, L2, or mean squared error (MSE), which modifies the weights to reduce the global pixel value deviation between the ground truth and the derived image. increase. As a result, the network weights do not change as fast as needed to restore saturated pixels.

1) The proposed method splits the input image into bright and dark regions and trains each separately. The overexposed areas are efficiently restored in the bright areas, and the overexposed areas are restored effectively in the dark areas. As a result, the proposed method can successfully restore all saturated pixels.

2) The proposed approach to generate image segmentation masks is free of artifacts. When bright and dark sections are mixed, the proposed mask provides smooth and clear image segmentation while eliminating interface artifacts.

3)The proposed multilevel network structure expands the dynamic range and evenly restores saturated pixels. Bright areas are brightened, and dark areas are darkened to increase the dynamic range. Then combine these features to create the final HDR image. Histogram of the log2 luminance of the HDR image. Brightness extremes that are likely to be clipped by the CRF are represented by green boxes.

The unwanted features of the bright image and dark image are optimized using the firefly optimization technique. After extracting the features from the process of CNN, these features are optimized to achieve good results and improve the quality of an image. The features are optimized depending on the process of the firefly algorithm.

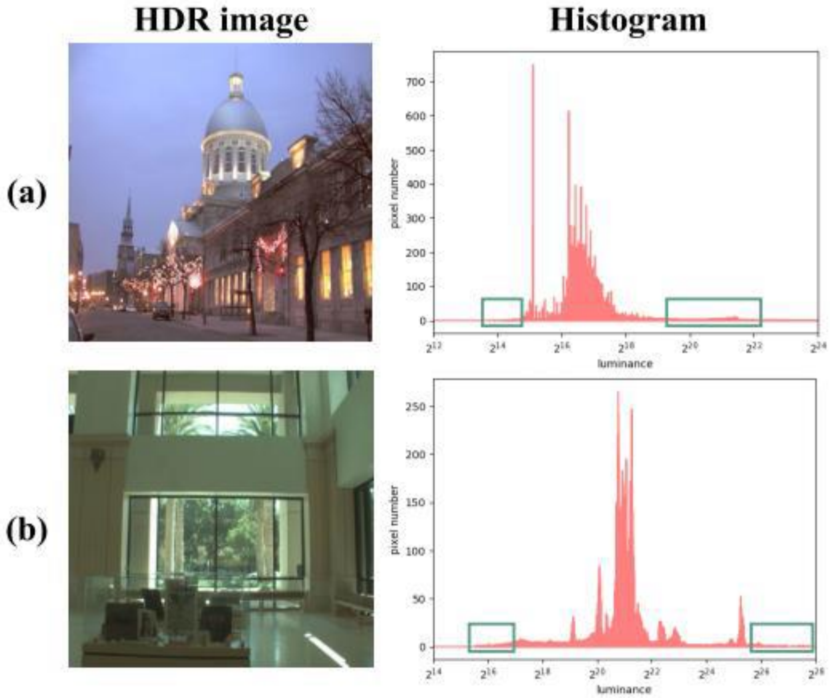

Figure 1: Histogram calculation

II. RELATED WORKS

2.1 MULTIPLE FUSION METHODS OF EXPOSURE:

By mixing bracket photos, create an HDR image. Martens et al. fused contrasting elements in multi-exposure photos using a Palladian pyramid and a Gaussian pyramid. The MEF approach is described in detail in the references. However, throughout the image merging process, the values of the bracket images were not properly combined, resulting in a halo art effect. To extract the weight values of bracket photos and generate HDR images with equally blended brightness levels, CNN-based MEF algorithms have been presented. Endo et al. used U-Net to bracket photos with different exposure values to create HDR images with large brightness ranges.

Because MEF approaches typically use more than three LDR pictures, they necessitate a lot of processing time, power, and storage space. To address these issues, Ma et al. developed a rapid MEF technique for generating weight maps for down-sampled bracketed pictures using a convolution neural network.

Two exposure fusion (TEF) methods that use only two images have been developed to mitigate the drawbacks of CNN-based MEF methods. Prabhakar et al. suggested a network that merges the brightness information of two LDR images with different exposure times.

However, because TEF approaches necessitate two exposure photographs; the images generated have a significant impact on the outcomes. Furthermore, similar to the CNN-based MEF method, restoring information that may be seen at exposure values other than the present exposure values is problematic.

2.2 METHODS FOR INVERSE TONE MAPPING:

Aksum et al. create HDR photos with gamma correction and linear expansion. However, unless the input parameters are appropriately specified, efficient range expansion is difficult to achieve. Bantered et al. created HDR photos with a power function and an expansion map, as well as an expanded map and the median cut procedure.

These methods, however, necessitate parameters for density estimation, and HDR image inferences have limits when it comes to recovering overexposed areas.

Huo et al. used a physiological ITM method based on the human visual system (HVS) to generate HDR images with fewer parameters, and Masia et al. suggested an automatic global ITM based on gamma expansion. However, the functions used were not sufficient to reconstruct the data.

Some proposed solutions, on the other hand, handle a single image by segmenting it into specific ranges. Didyk et al. divided photos into diffuse, reflection, and light zones, then applied different enhancements to each to produce HDR images with a wide brightness range.

Rumple et al. extracted pixels with high luminance and assigned information to saturated pixels using Gaussian filtering and an image pyramid. These methods were successful in restoring saturated pixels, but not in obtaining functions that matched the ground truth level colours and texture.

By applying a function to an LDR photo, the traditional ITM algorithm produces an HDR image. Finding a function that is similar to the inverse CRF, on the other hand, is extremely challenging. To overcome this problem, several deep learning algorithms have recently emerged.

Eilertsen et al. used an auto-encoder network with a loss function that took into account luminance and reflectivity to create HDR photos. By mixing photos to convert network outputs and LDR images to linear ranges, they were able to create HDR photographs with a higher dynamic range.

Most previous CNN-based ITM methods reconstruct HDR images by inputting LDR images to the network without significant pre-processing. segmented the image through pre-processing but based on only the frequency component. divided the image based on the brightness value, but divided the image after learning, and segmented it just before the final HDR image was generated. Efficient reconstruction of saturated pixels becomes difficult using raw LDR image inputs.

III PROPOSED METHOD:

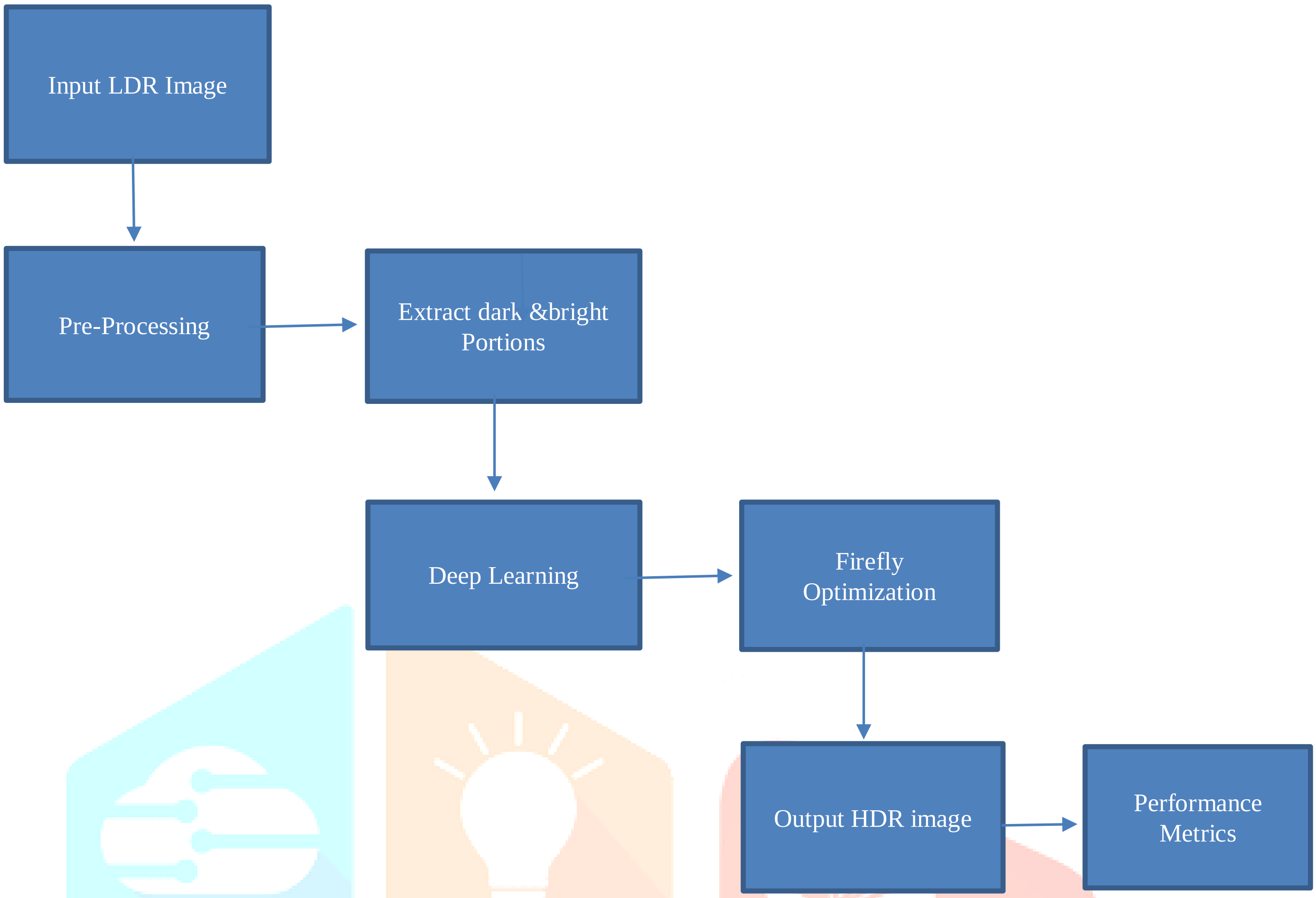

Figure 2: Block Diagram of proposed Method

The image threshold is calculated during pre-processing, and the Mask is generated. The LDR image is then segmented into bright and dark regions (Image _Bright and Image _Dark, respectively) using the Mask, and Image _Bright and Image _Dark become the inputs of the proposed multi-stage network. Image _Bright is learned to produce the increased brightness value in the Brightening block, and Image _Dark is learned by the Darkening block to reduce the lower brightness value. Finally, the two feature maps with enhanced dynamic ranges are combined to obtain the final HDR image. The original image Image _Original is also input to the network to help the calculations in the Brightening and Darkening blocks and feature blending. The bright and dark unwanted features of the image are optimized using the Firefly algorithm.

3.1PRE-PROCESSING:

Bright regions have information for only the bright pixels, and dark regions have information for only the dark pixels. As a result, the proposed network can effectively focus on over-exposed and under-exposed pixels when bright regions and dark regions are fed into the network, respectively. Therefore, the input image is segmented into bright and dark regions during pre-processing to enhance the dynamic range. Otsu proposed the representative image segmentation method that separates the image into two classes using the threshold value that provides a maximum pixel variance over the image. To clarify the difference between the two classes (Img_{Bright} and Img_{Dark}), it is necessary to find the threshold that maximizes the sum of the variances between the two classes. Based on the ostu model, a threshold is set for further processing. The proposed *Mask* generated based on the threshold can be expressed as,

$$Mask\ (n,m) = \left(\frac{max\ (0, V\ (n,m) - thr)}{(1 - thr)}\right)^2 \tag{3.1.1}$$

where $V(n,m)$ is image brightness; n, m are horizontal and vertical pixels, respectively. Since the square function is used for *Mask*, the pixel with a brightness value slightly larger than threshold has a very small mask value, which makes the boundary surface smooth when segmenting the image. The generated *Mask* is then multiplied by the original LDR image and segmented into Img_{Bright} and Img_{Dark}.

Bright Image: $Img_{Bright} = Img \times (Mask)$ **(3.1.2)**

Dark Image: $Img_{Dark} = Img \times (1 - Mask)$ **(3.1.3)**

3.2 DEEP LEARNING:

The features are extracted using the multi-stage CNN model. Network inputs include Img_{Bright} and Img_{Dark}, generated in pre-processing, and the original LDR image $Img_{Original}$. Img_{Bright} expands the dynamic range using the Brightening block increasing brightness, and Img_{Dark} expands the dynamic range using the Darkening block, reducing brightness. The two feature maps from distinct learning processes are subsequently combined properly to prevent artifacts.

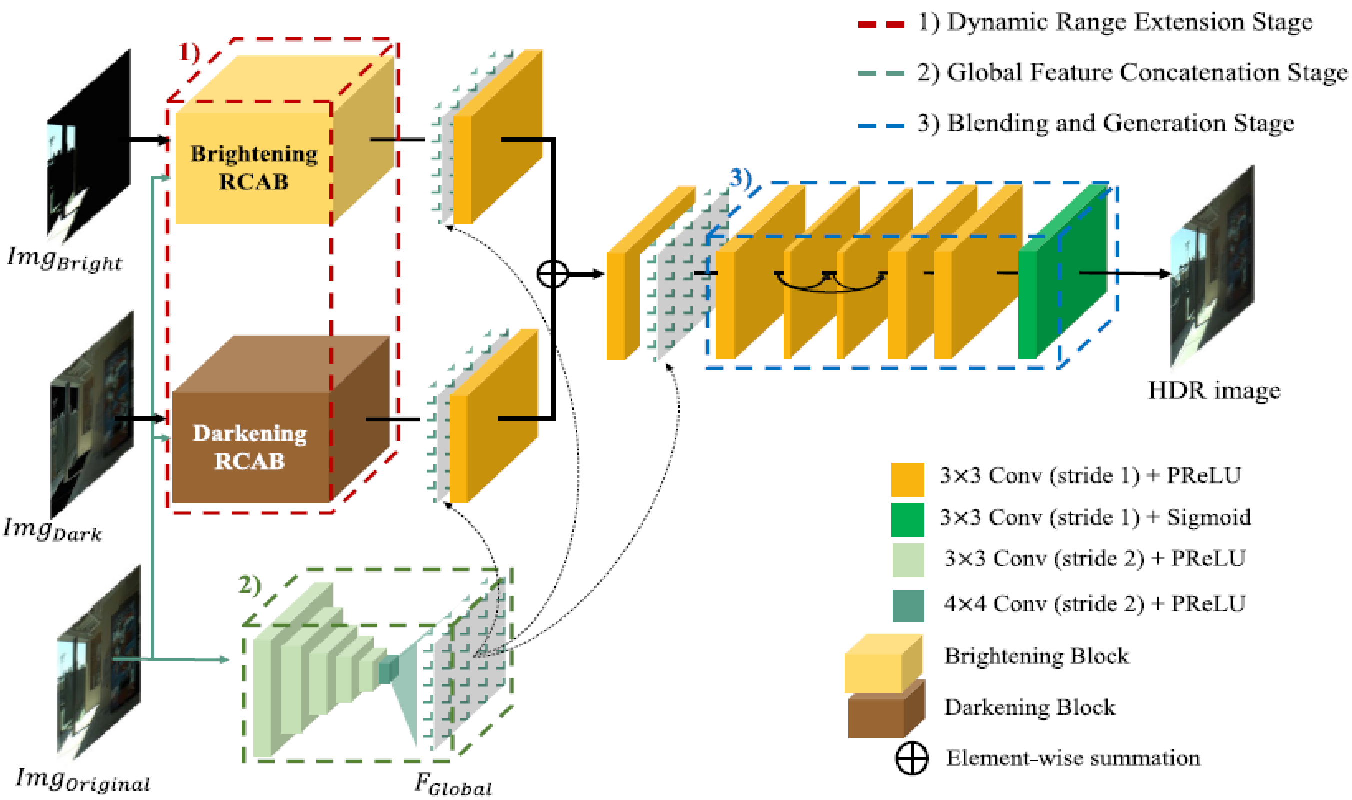

Figure3: Multi-stage deep learning network

It is Three stage model,
Stage1.
The dynamic range extension stage increases the bright range brightness and decreases the dark range brightness.
Stage2.
The global feature concatenation stage prevents outliers.
Stage3.
The blending and generation stage combines the learned feature maps.

3.3 FIREFLY OPTIMIZATION:

The undesirable highlights of the bright picture and dark picture are optimized utilizing the firefly optimization procedure. After extracting the highlights from the method of CNN, these features are optimized to attain great comes about and move forward the quality of the picture. The highlights are optimized depending on the method of firefly calculation. The point-by-point preparation of the firefly optimization is talked about. Fireflies are unisex so that one firefly will be pulled into other fireflies not withstanding their sex. Firefly's allure is corresponding to the light concentrated, and their engaging quality diminishes as their separate increments. In this way for any two flashing fireflies, the less shining one will move towards the brighter one. If there's no brighter one than a specific firefly, it'll move randomly.

Algorithm

Here are the steps to implement the proposed Firefly Algorithm (FA):

- Step 1: Generate the initial population of fireflies $\{x_1, x_2, \ldots\ldots, x_n\}$.
- Step 2: Calculate the brightness value for each firefly
- Step 3: Update the step of every firefly.
- Step 4: Fireflies intensity is given as $\{I_1, I_2, \ldots\ldots, I_n\}$
- Step 5: Line up the fireflies and find the perfect moment
- Step 6: Move each of the fireflies I to the other brighter ones.
- Step 7: Stop when a criterion is fulfilled; otherwise go to Step 2

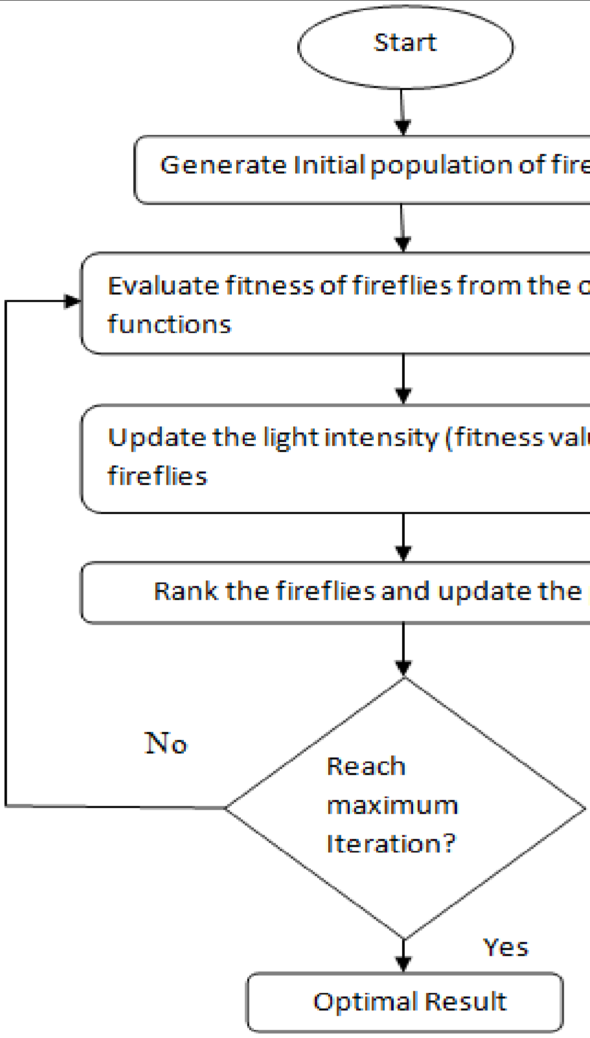

Figure 4: Firefly flowchart

Parameters used in Firefly Algorithm

Distance:
You can define the distance between any two fireflies **i** and **j** at positions $\mathbf{x_i}$ and $\mathbf{x_j}$ respectively, can be defined as a Cartesian or Euclidean distance ($\mathbf{r_{ij}}$) using the equation

$$r_{ij} = \left\| x_i - x_j \right\| = \sqrt{\sum_{k=1}^{d} (x_{i,k} - x_{j,k})^2} \qquad \textbf{(3.3.1)}$$

Where $\mathbf{x_i}$ is the intensity of a firefly and **k** is the **k**th component of the spatial coordinate $\mathbf{x_i}$ of the **i**th firefly.

Attractiveness:
Change in attractiveness β with distance r by

$$\beta = \beta o \qquad \textbf{(3.3.2)}$$

γ is an attractive coefficient.
r is the distance between two fireflies.

Movement:
The movement of a firefly ***i*** to attract another more attractive (brighter) firefly ***j*** is determined by

$$x_i^{t+1} = x_i^{\,t} + \beta_0 \, e^{-\gamma r_{ij}^2} (x_j^{\,t} - x_i^{\,t}) + \alpha_t \, \varepsilon_i^{\,t} \qquad \textbf{(3.3.3)}$$

where the second term is based on the attractiveness. The third term is randomized, with α being the randomization parameter and ε_i^t it is a vector of random numbers.

3.4 PERFORMANCE METRICS:

Parametric Evaluation
The parameters which are used to evaluate the performance of the proposed methodology are:

A. Peak Signal to Noise Ratio (PSNR)
B. Structural Similarity Index (SSIM)

C. Multi-Scale SSIM
D. HDR visual difference predictor (HDR-VDP)

Peak Signal to Noise Ratio (PSNR)

Peak signal-to-noise ratio (PSNR) is the ratio between the maximum possible power of an image and the power of noise that affects the quality of the image. To estimate the PSNR of an image, it is necessary to compare this image with an ideal sharp image with the maximum possible power. The higher the PSNR, the better the image quality is compressed or reproduced. For the input image X and the transformed image R, PSNR is expressed as

$$\mathrm{PSNR} = 10 \times \log_{10}\left(\frac{255 \times 255}{\mathrm{MSE}}\right) \tag{3.4.1}$$

Structural Similarity Index (SSIM)

SSIM can be a perceptual demonstration that considers image degradation as seen modified in the base data, also participating in obligatory perceptual wonders, including degree masking terms. brightness and the difference of the curtain. The difference with other procedures such as MSE or PSNR is that these methods evaluate outright errors.

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

Multi-Scale SSIM

The SSI estimation framework (Wang and Bovik,2002) is based on the modeling of picture luminance, differentiation, and structure. The MS-SSIM is an expansion of the SSI framework that accomplishes superior exactness than the single-scale SSI approach but at taken a toll of moderately lower handling speed.

HDR visual difference predictor (HDR-VDP)

High Dynamic Range Visible Difference Predictor (HDR-VDP) can work inside the total run of luminance the human eye can see. Input to our metric could be a High Dynamic Range (HDR) picture, or a conventional 8-bits-per-color picture, changed over to the real luminance values.

EXPERIMENTAL RESULTS:

The proposed method is stated as shown below. During pre-processing, the picture threshold is computed, and the Mask is created. The Mask is then used to segment the LDR pictures into bright and dark portions (Img Bright and Img Dark, respectively), with Img Bright and Img Dark serving as inputs to the proposed multi-stage network. The Brightening block learns Img Bright to create the enhanced brightness value, whereas the Darkening block learns Img Dark to reduce the lower brightness value. Finally, the final HDR image is created by combining the two feature maps with expanded dynamic ranges, and by using firefly optimization we extract the best features of an image.

S.NO	INPUT	OUTPUT
PSNR	27.2	30.9
SSIM	0.65	0.74
MSSIM	0.8	0.95
HDR-VDP	47.5	50.7

Table 1: various parameters comparison of input image and output image.

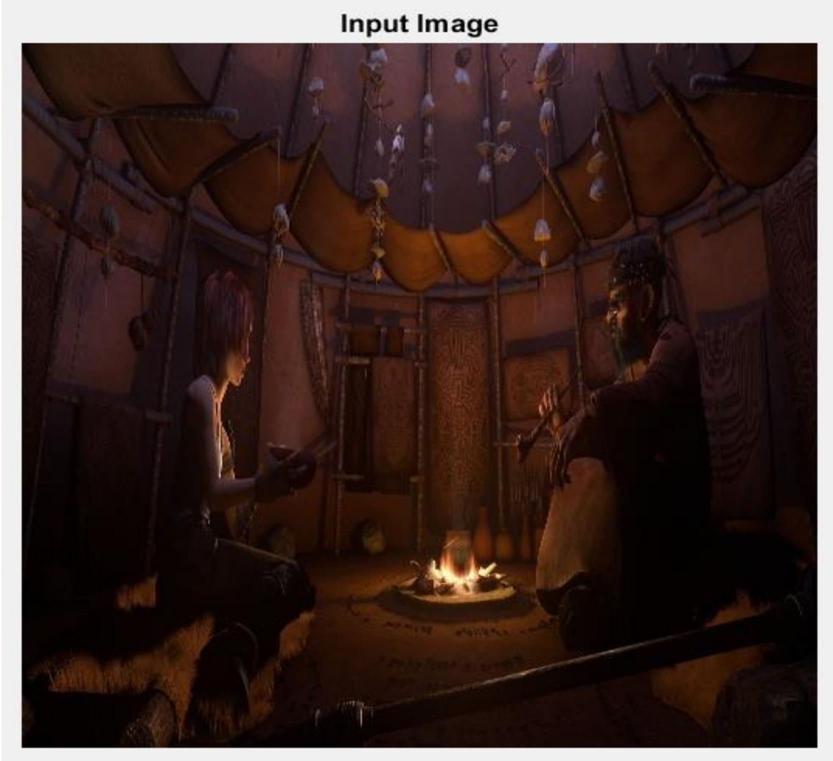

Figure5: Input Image

Figure6: HDR Image using Firefly Optimization

CONCLUSION:

Recent developments in deep learning technology have substantially advanced ITM approaches, with many studies employing deep learning to reconstruct saturated pixels and regenerate high-quality HDR images. However, the LDR image should be separated and learned according to brightness values to accurately reflect reality. Reconstructed HDR images using the proposed method achieved average PSNR, SSIM, MS-SSIM, and HDRVDP2 metrics higher than the previous methods, respectively. In particular, the proposed method provided a high mean log exposure range, confirming that the dynamic range was the most similar to the real world. In the proposed techniques optimization plays a key role in optimizing the extracted features. The firefly methodology identifies pixels and features which help in reconstructing the HDR image.

REFERENCES:

[1]S. Lee, G. H. An, and S.-J. Kang, ``Deep chain HDRI: Reconstructing a high dynamic range image from a single low dynamic range image," *IEEE Access*, vol. 6, pp. 49913_49924, 2018.
[2] S. Lee, G. H. An, and S.-J. Kang, ``Deep recursive HDRI: Inverse tone mapping using generative adversarial networks," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Sep. 2018, pp. 596_611.
[3] S. Lee, S. Y. Jo, G. H. An, and S.-J. Kang, ``Learning to generate multi exposure stacks with cycle consistency for high dynamic range imaging," *IEEE Trans. Multimedia*, vol. 23, pp. 2561_2574, 2021.
[4] K. Ma, Z. Duanmu, H. Zhu, Y. Fang, and Z. Wang, ``Deep guided learning for fast multi-exposure image fusion," *IEEE Trans. Image Process.*, vol. 29, pp. 2808_2819, 2020.
[5] K. R. Prabhakar, V. S. Srikar, and R. V. Babu, ``Deep Fuse: A deep unsupervised approach for exposure fusion with extreme exposure image pairs," in *Proc. ICCV*, Oct. 2017, pp. 4724_4732.
[6] J.-L. Yin, B.-H. Chen, and Y.-T. Peng, ``Two exposure fusion using the prior aware generative adversarial network," *IEEE Trans. Multimedia*, early access, Jun. 15, 2021, doi: 10.1109/TMM.2021.3089324.
[7] G. Eilertsen, J. Kronander, G. Denes, R. K. Mantiuk, and J. Unger, ``HDR image reconstruction from a single exposure using deep CNNS," *ACM Trans. Graph.*, vol. 36, no. 6, p. 178, 2017.
[8] J. Cai, S. Gu, and L. Zhang, ``Learning a deep single image contrast enhancer from multi-exposure images," *IEEE Trans. Image Process.*, vol. 27, no. 4, pp. 2049_2062, Apr. 2018.
[9] D. Marnerides, T. Bashford-Rogers, J. Hatchett, and K. Debattista, ``Expand Net: A deep convolutional neural network for high dynamic range expansion from low dynamic range content," *Comput. Graph. Forum*, vol. 37, no. 2, pp. 37_49, May 2018.
[10] Y.-L. Liu, W.-S. Lai, Y.-S. Chen, Y.-L. Kao, M.-H.Yang, Y.-Y. Chuang, and J.-B. Huang, ``Single-image HDR reconstruction by learning to reverse the camera pipeline," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.(CPR)*, Jun. 2020, pp. 1651_1660.
[11] R. K. Mantiuk, K. Myszkowski, and H. P. Seidel, ``High dynamic range imaging," in *Wiley Encyclopedia of Electrical and Electronics Engineering*. Hoboken, NJ, USA: Wiley, 2015, pp. 1_42.
[12] Y. Endo, Y. Kanamori, and J. Mitani, ``Deep reverse tone mapping," *ACM Trans. Graph.*, vol. 36, no. 6, p. 177, 2017.
[13] P. E. Debevec and J. Malik, ``Recovering high dynamic range radiance maps from photographs," in *Proc. ACM SIGGRAPH Classes*, Aug. 2008, pp. 1_10.
[14] T. Mertens, J. Kautz, and F. Van Reeth, ``Exposure fusion," in *Proc. 15th Paci_c Conf. Comput. Graph. Appl. (PG)*, Maui, HI, USA, Oct./Nov. 2007, pp. 382_390.
[15] A. Srikantha and D. Sidibé, ``Ghost detection and removal for high dynamic range images: Recent advances," *Signal Process., Image Com-mun.*, vol. 27, no. 6, pp. 650_662, 2012.
[16] A. O. Akyüz, R. Fleming, B. E. Riecke, E. Reinhard, and H. H. Bülthoff, ``Do HDR displays support LDR content? A psychophysical evaluation,"*ACM Trans. Graph.*, vol. 26, no. 3, pp. 1_7, Jul. 2007.
[17] F. Banterle, P. Ledda, K. Debattista, and A. Chalmers, ``Inverse tone mapping," in *Proc. 4th Int. Conf. Comput. Graph. Interact. Techn. Australasia Southeast Asia*, New York, NY, USA, 2006, pp. 349_356.
[18] F. Banterle, K. Debattista, A. Artusi, S. Pattanaik, K. Myszkowski, P. Ledda, and A. Chalmers, ``High dynamic range imaging and low dynamic range expansion for generating HDR content," *Comput. Graph.Forum*, vol. 28, no. 8, pp. 2343_2367, 2009.