IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES
ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

Building an Image Enhancer using Deep Learning and SICE Techniques

Mr. V. Koteswara Rao[1], Asst. Professor, Department of CSE,
Vasireddy Venkatadri Institute of Technology, Nambur, Guntur Dt., Andhra Pradesh.

Srilatha Mathangi[2], **Vasipalli Mahitha Reddy**[3], **Tullimilli Shanmuka Sagar**[4], **Vinay Kumar Buddi**[5], **Tiyyagura Varsha**[6]
[2,3,4,5] UG Students, Department of CSE,
Vasireddy Venkatadri Institute of Technology, Nambur, Guntur Dt., Andhra Pradesh.
koteswararao@vvit.net,mathangisrilatha10@gmail.com,
mahithavasipalli@gmail.com, shanmukasagar2019@gmail.com,
vinaychowd2001@gmail.com, varshatiyyagura13@gmail.com
DOI:10.48047/IJFANS/V11/I12/185

Abstract

Most of the images captured on digital devices like cameras, mobiles are often under exposed or over exposed to light due to inappropriate lighting conditions, which is a cause of loosing detailing of the picture. To adjust the lighting in the outcome, there are various techniques like single image contrast enhancement which are trained on a single image to spotlight the defects in the image and correct it wherever needed. This could solve the irregularities to some extent, but may not give satisfactory results in all possible scenarios. Hence, we need to train a memory (algorithm in this case) on multiple images, which could memorise a defect and its corresponding resolution tactics. All of this knowledge could be used at once to identify multiple blemishes in the input and corresponding fixes could be made for each of them. For this purpose of knowledge extraction, Convolutional neural networks (CNN) are employed on the dataset which will study and identify the problems like darkness, over exposure, blurred images and apply the remedies on them. Low Light Image/ Video enhancing (LOL) dataset is used for this purpose which has 500 pairs of defective and corresponding corrected images. CNN is trained easily on the dataset to provide significantly better results over the existing SICE techniques.

Keywords: SICE, Image enhancer, Contrast enhancer, Deep Learning, Deep CurveEstimation, Convolution, Neural Net, LOL, Blemishes, Remedies, Exposure.

Introduction

The aim of digital photography is to produce high quality images by retaining the details and crisp colors of the scene. The obtained results are recurrently flooded with light or may contain underrepresented details due to lower performance specifications of the device or unsupported lighting conditions. This lower quality of images degrades the performance of computer processing algorithms or may diminish the visual experience. Thus, contrast improvisation is the crucial phase of image quality enhancement to make the details visually clear.Legacy practices employed the approaches based on histograms,

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

which would identify points of unevens in the scene and would try to extemporize them distinctly.

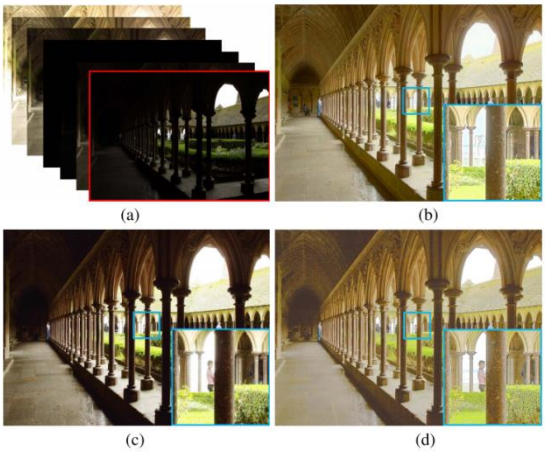


Fig. 1. (a) A sequence with under- and over-exposed images. (b) The enhanced image by a deghosting MEF method [2]. (c) The enhanced image by a SICE method [3]. (d) The enhanced image by our method. (Note that (c) and (d) are the enhanced results by a single under-exposed image.)

However, these methodologies would hardly produce a satisfying result due to scarce insights available by studying single image alone. Fortunately, devolopment of digital devices that are capable of capturing multiple shots at a time helped to collect different exposures of a single scene at the same time hence it would rather facilitate the algorithm with more data, and thus gather sufficient knowledge after studying them. With a group of multi exposure images, it is possible for several algorithms like stack based high definition range(HDR) algorithms are used for studying the data which is generally time consuming, also with the devolopment of image generation algorithms like GANs it would be helpful to implement them to reproduce an image with out need of training larger datasets containing multiple similar images. But we use Deep Curve Estimation(DCE) algorithm that aims at improving the brightness and contrast of low light images using a deep learning approach based on CNNs. DCE is pleasing because, multiple input-output image pairs are not required during training a Zero DCE. This is acheived through a set of carefully designed loss functions which implicitly measure the enhancement quality and guide the training of the network. For this study a popular dataset called LOL data set is used. This research uses DCE algorithm which was devoloped by 'Ivangonchorav' specifically for the image enhancement which does not require any particular dataset.But, there are several datasets available publicly like the 'League of Legends(LOL)' dataset, which contains about 60000 records for training and testing, 'Sony Image dataset(SID)' which contains 4000 records of raw images taken under low light using a Sony camera, and the LOL dataset with 500 records. Since we train the algorithm over 50 epochs to gain a good accuracy, LOL dataset is used due to smaller size than its rivals.It would facilitate to gather enough knowledge as well as train on more number of epochs. Hence we chose the LOL dataset for training and testing. It consists of 500 low light images and their corresponding well light images, among

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

which 485 images (300 for training and 185 for validation purposes) and 15 records for testing purpose. Along the way, we need to formulate the loss functions since this wasn't a classification task we wouldn't be able to obtain the performance metrics directly, rather we measure the loss of the model for every epoch and after completing the epoch cycles, we would generate the accuracy of the model as a function of total loss obtained. Along with this model, several no code tools like Weights & Biases platform is used to reduce the effort of devolopment with pre-existing visualization tools, loaded datasets and controlled access. This makes it easier for devolopment,which is handy to explore the project in a research oriented way,without need of concentrating more on coding.

Literature Survey

A.LIME: Low-light image enhancement via illumination map estimation

Fig 2: Single image contrast enhancement results on an under-exposed image by different methods (PSNR/FSIM). (a) Original. (b) Reference. (c) CVC (d) AGCWD (e) NEP (f) SRIE (g) LIME (h) Li (i) DN (MSE) (j) DN (k) DN (DSSIM) (l) stack based HDR.

Low-light image enhancement[1] via illumination map estimation (LIME) is a simple yet effective enhancement method applied on the images that are captured in low-light conditions. The key objective is to estimate the illumination map and a structure aware smoothing model has been devoloped to impove the illumination consistency. In this research work, two algorithms were designed. One to obtain an optimal solution to the target problem and the other is to find an alternative solution for saving time. 'HDR dataset' is used along with the training models like LIME,HE,AHE which acheived a decent 80% accuracy.

B.Robust high dynamic range imaging by rank minimization [2]

For capturing multi exposure images, cameras are relatively expensive and mobile cameras were still in the budding stage. Hence, an effective HDR deghosting fusion network

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

(DDFNet) is proposed, which operates on motion attention and image correlation attention. For this work, 'Kalantari's' dataset is used along with DDFNet model for training which takes a low light image as input to give a HDR image as output. This model secured 69% accuracy.

C.Correcting over-exposure in photographs

Firstly, the over exposed region(OER) must be identified in the image for correction. Rate of detecting depends on the human visual saturation sensitivity. The proposed OER detection method can improve correction rate of excessive exposure of image. The proposed OER is able to precieve human eye and improve the OER correction performance. Gray scale images were taken as input and produces colorised images with a quality rate of 4.3.[3]

D.Deep Bilateral Learning for Real-Time Image Enhancement

The problem statement identified in this research[4] is "when given a human adjusted image pairs, the goal is to reproduce them with enhancements". A neural network is introduced for this work, whose architecture is inspired by bilateral grid processing and local affine color transforms. The algorithm is able to process high resolution images upto 1080p and reproduce them on a smartphone itself. 'FiveK' dataset is used along with a custom CNN which acquired an accuracy of 31.8 db.

E.Automatic Exposure Correction

The aim of this research work is to estimate an image specific tone curve whish is in S shape that fits best on the image. For the purpose of exposure correction, an automatic method is proposed in this [5] work. Optimised evaluation of exposure and an algorithm to adjust a curve to preserve the details were heart of this work. The Adaptive lighting – delighting model used 4000 images which were directly captured by a camera were used for training which gained 92% accuracy.

Problem Identification

The intention of SICE is to amplify the legibility of the image without loosing the minor detailing in it. Historical approaches like that use histograms were broadly used due to its straightforwardness.It is a naïve methodology of identifying the points of dullness through the x and y axes and amplifying them alone, this would rather create differentiable blemishes like watermarks on the image inspite of providing a brighter output. Such minimalistic redistributions produce highly unrealistic effects as they ignore the strucutural organization of the image. Stack based HDR algorithms employ tone mapping operator to compress the dynamic range of HDR irradiance map so that high contrast images can be rendered on low spec displays. [11-19]

Disadvantages:

- The histogram approach produces unrealistic images since they do not concentrate on the broader prospect of the image.

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

- Scope of training is limited to single image which wouldn't provide enough knowledge to work on new set of images.
- Whereas, in stack based HDR, despite their well alignment of image sequences, in the cases employing camera shakes, moving objects produces ghostic effects in the final enhanced image.

Methodology

CNNs are used for image and text based deeplearning tasks. In this view, the DCE algorithm which is built on the Zero DCE framework is considered, as the CNN for this task to train the SICE enhancer.

Zero DCE Framework

Estimation of best fitting light enhancement curves set is the key responsibility of DCE network. A curve that automatically maps low light image to the corresponding enhanced version is called light enhancement curve. Conventional approaches involved manual fitting of these curves with different permutations, hence it was a time consuming process and it did not produce efficient results. DCE is different from its competitors where the light enhancement curve is automatically learned from the training data through deeplearning approach. A light enhancement curve is typically a mathematical function which is predicted for its best fit through trail and error methods in the conventional techniques. When designing a light enhancement curve, three objectives should be considered:

a) Each pixel should be normalized in the range of 0 through 1 to avoid data loss.
b) It should preserve contrast between neighbouring pixels.
c) The shape of curve should be simple so that it would allow for back propagation, which is the process of calculating the gradient of a loss function with respect to the weights of a neural network and using it to update the weights during training.

Advantages:

- Learns the light enhancement curves automatically from the training data through deep learning.
- Trains on multiple different images at the same time which would equip the model with more knowledge compared to the former approach.

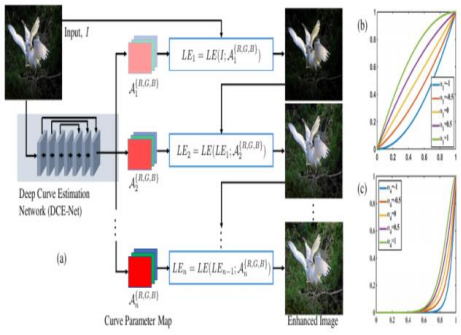


Fig 3. Overview of Zero DCE framework architecture

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

DCE Net

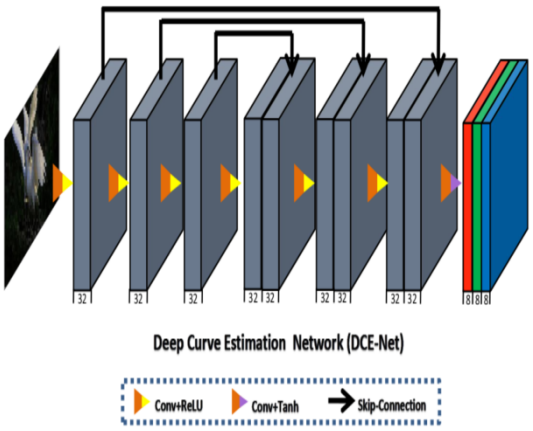


Fig 4. Architecture of DCE Net

- It is a deeplearning architecture for learning mapping between input image and its best fitting curve parameter maps.
- A low light image is taken by the DCE Net which inturn produces a set of pixel wise curve parameter maps for corresponding higher order curves.
- The network consists of 7 convolutional layers, each layer has 32 kernels of 3*3 large, and stride of size 1 followed by a ReLU activation function. The last layer is followed by a Tanh activation function.

Loss Functions

- A loss function is designed to measure the difference between colorised image and true image.
- A set of distinguishing zero reference loss functions are formulated to evaluate the quality of enhanced images to facilitate zero reference learning, which is the process of training a model to recognize the features without any prior knowledge or reference to specific set of training data.
- The model is trained on a large set of unannotated images which targets to extract generalizable features that would be helpful in recognizing a new, unseen image.
- Such approach would be useful when it is difficult to obtain labelled data.
- The loss functions used in this research are listed below:
  - ✓ Color constancy loss
  - ✓ Exposure loss
  - ✓ Illumination smoothness loss
  - ✓ Spatial consistency loss
- **Color constancy loss:** firstly, color constancy is the ability to retain or adjust colors of objects under various lighting conditions. It is pretty challenging when it comes to computer vision. It is a task that involoves color correction of image to make it appear as if it was captured under normal lighting conditions.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

- Defining the function involves several steps:
  - ✓ Calculate the mean RGB value for each image in the batch:
    mean_rgb_{i,j,k} = \frac{1}{N}\sum_{l=1}^{N} x_{l,i,j,k} where N= no. of pixels in the image, I and j are spatial dimensions, and k is the channel dimension.
  - ✓ Calculate squared differences between mean red, blue and green values:
    d_{rg} = (mean_r - mean_g)^2
    d_{rb} = (mean_r - mean_b)^2
    d_{gb} = (mean_g - mean_b)^2
  - ✓ Sumup the squared differences and square root the result to obtain the color constancy loss:
    loss = \sqrt{d_{rg}^2 + d_{rb}^2 + d_{gb}^2}
- The loss function guides the model to learn to remove the color noise or over colorised parts of the image while preserving its natural appearance.
- **Exposure loss:** exposure is the most important aspect of an image which affects the brightness, contrast and overall appearance. Correction of exposure is important to make it pleasing visually.
- The loss function guides the image to adjust the exposure of input image while minimizing the difference between corrected image and ground truth image in terms of pixel values.
- **Illumination smoothness loss:** illumination smoothness is the measure of rate of smoothness at which the illumination of the image is changing across the image.
- The mathematical representation of the function is:
  h_x = height of image
  w_x = width of image
  *count_h = (w_x - 1) * h_x * batch_size*
  *count_w = w_x * (h_x - 1) * batch_size*
  h_tv = sum_i sum_j (x[i, j+1, :, :] - x[i, j, :, :])^2
  w_tv = sum_i sum_j (x[i, :, j+1, :] - x[i, :, j, :])^2
  *loss = 2 * (h_tv / count_h + w_tv / count_w) / batch_size.*
  Where h_tv and w_tv represents the sum of squared differences in illumination between adjacent pixels along the width and height respectively.
- The illumination smoothness loss function is used during training to guide the model to produce the images with smooth illumination.
- **Spatial consistency loss:** spatial consistency is the measure of consistency of adjacent regions in the image.
- The loss function guides the model to produce the images with spatial consistency during training.

There are various phases involved in this research namely:

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

- Data collection: using this module, collects the data required for training.
- Data visualization: using the module, collected data is visualized.

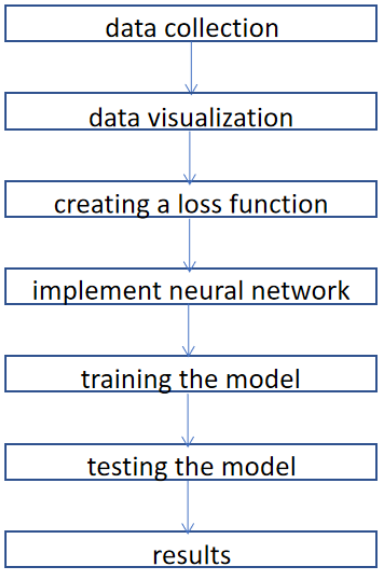


Fig 5: System Architecture

- Loss function creation: using this module, models will be guided during training to minimize various types of losses in image enhancement.
- Implementing deeplearning model: this module involves creation of deeplearning model, typically a CNN for the learning purpose using a library like Keras.
- Training the model: model learns from available data, perform feature extraction and mapping parameters.
- Testing and validation: trained model is tested with unseen set of inputs.

Implementation

For acheiving the above stated objectives, we create and train a DCE Net algorithm using keras library. The implementation involved various steps:

- Download the LOL dataset into devolopment environment and visualize the raw data to get a glance at the dataset.
- For optimised computing, a tensorflow dataset object is used for training and evaluation. The images are initially 512*512 px from the dataset we considered which are resized to 256*256 px or only the width is resized manually and height is adjusted to preserve the aspect ratio.
- The tensorflow dataset object facilitates in efficient loading, data preprocessing, batching, as well as shuffling the data and paralell processing.
- As it is stated before, DCE net is a network of convolutional layers with ReLU and Tanh activation functions.

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

- It is implemented in the form of code using the keras library. It consists of 7 convolutional layers, each having kernels of size 32 and strides of size 1*1 with first 6 layers having the ReLU activation

```
def build_dce_net():
    input_img = keras.Input(shape=[None, None, 3])
    conv1 = layers.Conv2D(
        32, (3, 3), strides=(1, 1), activation="relu", padding="same"
    )(input_img)
    conv2 = layers.Conv2D(
        32, (3, 3), strides=(1, 1), activation="relu", padding="same"
    )(conv1)
    conv3 = layers.Conv2D(
        32, (3, 3), strides=(1, 1), activation="relu", padding="same"
    )(conv2)
    conv4 = layers.Conv2D(
        32, (3, 3), strides=(1, 1), activation="relu", padding="same"
    )(conv3)
    int_con1 = layers.Concatenate(axis=-1)([conv4, conv3])
    conv5 = layers.Conv2D(
        32, (3, 3), strides=(1, 1), activation="relu", padding="same"
    )(int_con1)
    int_con2 = layers.Concatenate(axis=-1)([conv5, conv2])
    conv6 = layers.Conv2D(
        32, (3, 3), strides=(1, 1), activation="relu", padding="same"
    )(int_con2)
    int_con3 = layers.Concatenate(axis=-1)([conv6, conv1])
    x_r = layers.Conv2D(24, (3, 3), strides=(1, 1), activation="tanh", padding="
        int_con3
    )
    return keras.Model(inputs=input_img, outputs=x_r)
```

Fig 6: implementation of DCE Net using keras library

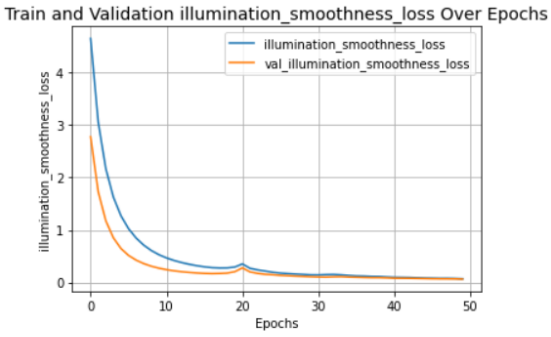
function and the last one using Tanh activation function.

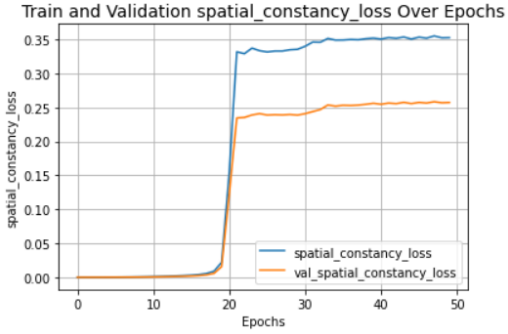
- Tensorflow API is used for the purpose of implementing above stated loss functions.
- The model is trained on the dataset in 50 epoch cycles each taking about 10 minutes to attain minimum loss in each iteration and obtain a decent accuracy at the end.
- Finally, the model is tested with various test data and also validated on various performance metrics.
- Accuracy of the model is calculated as a function of total loss obtained.

  avg_loss_per_image = total_loss / num_images

  *accuracy = (1 - avg_loss_per_image) * 100*

Results





Research paper **© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022**

Fig 7, 8: visualization of illumination smoothness loss function, spatial contancy loss function during training and validation

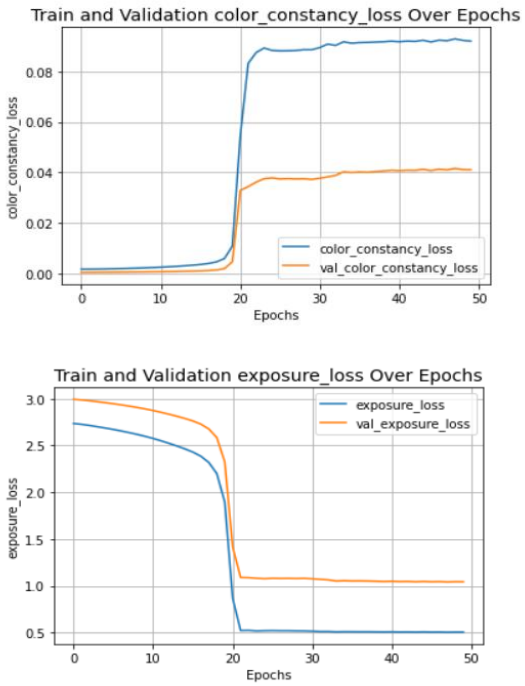
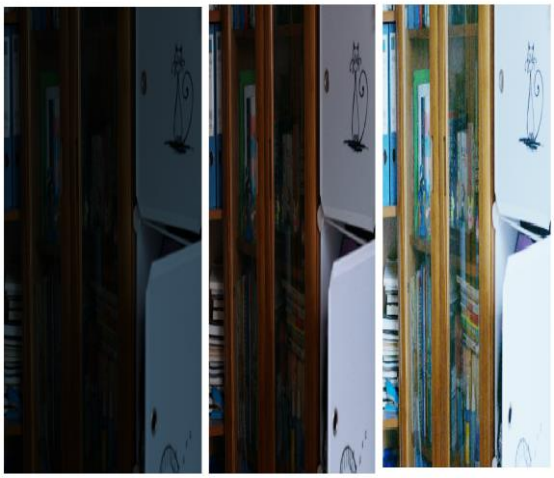


Fig 9, 10: visualization of illumination smoothness loss function, spatial contancy loss function during training and validation



(a) (b) (c)

Fig 11: test sample of image; (a) original image, (b) PIL auto contrast, (c) enhanced image

Conclusion

The results were seen for various learning rates, epoch counts and data splits. With a 99% accuracy rate, the algorithm was able to upscale the image quality based on the knowledge acquired by DCE-Net. This accuracy is attained by considering the images taken in the pitch black conditions also. The performance parameters could be adjusted to further increase the accuracy.

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

Limitations and Future Scope

- One of the considerable limitations of this approach is that the time taken for training is very high, typically 10 minutes per epoch.
- To overcome this, It will be better to explore similar algorithms for faster training rate and would also train on larger sized datasets.
- The performance of the model can be improved by hyper parameter tuning.
- After training the model with more data further, this model will be deployed on a cloud service like Sagemaker to make it accessible to all.

References

[1] X. Guo, Y. Li, and H. Ling, "LIME: Low-light image enhancement via illumination map estimation," IEEE Trans. Image Process., vol. 26, no. 2, pp. 982–993, Feb. 2017.

[2] T. H. Oh, J. Y. Lee, Y. W. Tai, and I. S. Kweon, "Robust high dynamic range imaging by rank minimization," IEEE Trans. Pattern Anal. Mach. Intell., vol. 37, no. 6, pp. 1219–1232, Jun. 2015

[3] D. Guo, Y. Cheng, S. Zhuo, and T. Sim, "Correcting over-exposure in photographs," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2010, pp. 515–521.

[4] M. Gharbi, J. Chen, J. T. Barron, S. W. Hasinoff, and F. Durand, "Deep bilateral learning for real-time image enhancement," ACM Trans. Graph., vol. 36, no. 4, p. 118, 2017.

[5] S. Wang, J. Zheng, H.-M. Hu, and B. Li, "Naturalness preserved enhancement algorithm for non-uniform illumination images," IEEETrans. Image Process., vol. 22, no. 9, pp. 3538–3548, Sep. 2013.

[6] Z. Li, Z. Wei, C. Wen, and J. Zheng, "Detail-enhanced multiscale exposure fusion," IEEE Trans. Image Process., vol. 26, no. 3, pp. 1243–1252, Mar. 2017.

[7] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2016, pp. 770–778.

[8] Y. Endo, Y. Kanamori, and J. Mitani, "Deep reverse tone mapping,"ACM Trans. Graph., vol. 36, no. 6, Nov. 2017, Art. no. 177.

[9] M. Gharbi, J. Chen, J. T. Barron, S. W. Hasinoff, and F. Durand, "Deepbilateral learning for real-time image enhancement," ACM Trans. Graph.,vol. 36, no. 4, p. 118, 2017.

[10] Z. Li, J. Zheng, Z. Zhu, W. Yao, and S. Wu, "Weighted guided image filtering," IEEE Trans. Image Process., vol. 24, no. 1, pp. 120–129,Jan. 2015.

[11] Sri Hari Nallamala, et al., "A Literature Survey on Data Mining Approach to Effectively Handle Cancer Treatment", (IJET) (UAE), ISSN: 2227 – 524X, Vol. 7, No 2.7, SI 7, Page No: 729 – 732, March 2018.

[12] Sri Hari Nallamala, et.al., "An Appraisal on Recurrent Pattern Analysis Algorithm from the Net Monitor Records", (IJET) (UAE), ISSN: 2227 – 524X, Vol. 7, No 2.7, SI 7, Page No: 542 – 545, March 2018.

[13] Sri Hari Nallamala, et.al, "Qualitative Metrics on Breast Cancer Diagnosis with Neuro Fuzzy Inference Systems", International Journal of Advanced Trends in Computer Science and Engineering, (IJATCSE), ISSN (ONLINE): 2278 – 3091, Vol. 8 No. 2, Page No: 259 – 264, March / April 2019.

[14] Sri Hari Nallamala, et.al, "Breast Cancer Detection using Machine Learning Way", International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878, Volume-8, Issue-2S3, Page No: 1402 – 1405, July 2019.

[15] Sri Hari Nallamala, et.al, "Pedagogy and Reduction of K-nn Algorithm for Filtering Samples in the Breast Cancer Treatment", International Journal of Scientific and Technology Research, (IJSTR), ISSN: 2277-8616, Vol. 8, Issue 11, Page No: 2168 – 2173, November 2019.

[16] Kolla Bhanu Prakash, Sri Hari Nallamala, et al., "Accurate Hand Gesture Recognition using CNN and RNN Approaches" International Journal of Advanced Trends in Computer Science and Engineering, 9(3), May – June 2020, 3216 – 3222.

[17] Sri Hari Nallamala, et al., "A Review on 'Applications, Early Successes & Challenges of Big Data in Modern Healthcare Management'", Vol.83, May - June 2020 ISSN: 0193-4120 Page No. 11117 – 11121.

[18] Nallamala, S.H., et al., "A Brief Analysis of Collaborative and Content Based Filtering Algorithms used in Recommender Systems", IOP Conference Series: Materials Science and Engineering, 2020, 981(2), 022008.

[19] Nallamala, S.H., Mishra, P., Koneru, S.V., "Breast cancer detection using machine learning approaches", International Journal of Recent Technology and Engineering, 2019, 7(5), pp. 478–481.