# Colour Image Enhancement Using Self-Guided Attention Mechanism

**Ruth Rajitha Rani Gangarapu[1*], Samson Chepuri[2]**

[1] *Research Scholar, Department of computer science and Engineering (IT), University college of Engineering, (A), Osmania University, Hyderabad, Telangana, India*

[2] *Professor, Department of IT, MVSR Engineering College, Hyderabad, Telangana, India*

***Abstract:***
The objective of image enhancement is to improve the comprehensibility of the image. However, the enhancement process may cause colour inconsistency such as green-tones and oversaturation leading to be a challenging task. The proposed study aims at construction of deep neural network to bring out image details, improve brightness, smoothed edge and to eliminate colour artifacts and noise with computational economy. We have delved into construction of deep neural network to denoise variety of noises and combination of noises in images of JPEG and PNG format and tested with image of variant datasets: LOLv1, Set14, Cannon-LR image and all enhanced image have SSIM > 0.9999 with execution time of 35s/step. This study does not involve either usage or storage of feature weight score vector by eliminating SoftMax layer which we usually find in the DNN models, no pre-training, low computational cost and time, minimal resources usage by implementing efficient calibration mechanism of operations: the self-guided attention mechanism which facilitates in reducing number of model parameters and as well produces the visually appealing image and the experimental results reveals its efficacy.

***Keywords:*** Convolution, Channel Attention, Deep Neural Network, Multiscale Features, Spatial Attention.

## 1. Introduction

The digital images may undergo deterioration or distortion by the process of formation, transmission, recording or display. So, images of low quality hinder their usage. Hence image enhancement is an important part of image processing. Image enhancement is a process of obtaining the image with strong expressiveness, obvious characteristics, and rich effective information which optimizes the display effect. The emergence of Deep learning has transformed the realm of image enhancement. Image enhancement involves noise suppression, contrast, and brightness improvement in turn resulting in image with rich details, improved texture, and colour. There is still a dilemma whether to denoise first followed by improvement of colour, contrast, brightness or vice versa**.** Applying denoising first would result in blurred image, whereas improving of colour, contrast, brightness before denoising will cause noise amplification**.** Hence there is a need for a model or method that can evade such a kind of problem. Image enhancement process can involve manipulation of entire image or specific region, shape referring to primitives of geometric attributes. Image enhancement can be carried out in spatial domain, frequency domain, or hybrid domain. In frequency domain, the processing of an image is carried out in transformed domain such as wavelet transformed domain, discrete sine or discrete cosine transformed domain. In spatial domain the image enhancement is carried out on pixel's intensity levels. In hybrid domain, the transformed domain is combined with Neural Network to generate high quality image with less training data set. Since spatial domain is concerned with the intensity value of pixels so we can perform either point processing operations or spatial filtering operations. The filtering operation is carried out using filter. The other names for filter are Kernel, Mask, window, Template. Since image is characterized by number of pixels and colour depth, so choice of image file format has an impact on the image quality, colour shades and purpose. The purpose can be either to print or store. This study is carried out in spatial domain with image file of JPEG and PNG format. The Portable Network Graphic (PNG) file format can adjust to varying range of colours. However, this format is not suitable for printing because it does not support CMYK colour model and is primarily useful for digital purpose. Whereas Joint Photographic Expert group (JPEG) is highly compatible and supports full colour spectrum up to 24-bits but may lose details when we process the image for compression.

The proposed study is based on deep neural networks. The Deep Neural Network (DNN) are such a type of networks where each layer can perform complex operations such as representation and abstraction that makes sense of images, sounds, and text. The DNNs are frequently referred to as "Black Boxes" on account of their architecture. So, efforts are required for improving the interpretability of these DNNs with a goal of fostering confidence in their decisions. The deep learning model proposed by us is an unsupervised attention-based mechanism in spatial domain**.** The most of the current deep learning methods for image enhancement are attention-based approaches focusing on attenuation of intensity across different channels and within a channel. These techniques aim to improve image quality by mitigating the effect of noise, enhancing contrast, reducing colour distortion, and restoring fine details. [1] proposed a channel self-attention based low-light image enhancement (CAENet) guided by signal-to-noise ratio priors and attention maps to adaptively supress noise. Since variability of noise and luminance in different regions is different, they designed a model by considering local and global features thereby producing an image with fewer noise artifacts and with improved details. This approach uses SoftMax layer for implementing attention mechanism. The attention-based approach for under water image enhancement [2] leverages the advantage of prior knowledge of attenuation in RGB channels and deep learning. In this approach, features are extracted from different perspectives and locations of the image. And Multi-scale Feature Aggregation (MFA) module dynamically adjusts the weights of features based on the correlation between different layers to improve the enhancement effect of contrast, clarity, and colour. The advantage of this mechanism is that it exploits deep learning representational ability of features and extending its further utilization by the guidance of priors. In [3] correlation between image depth and hazy densities are used during dehazing process. This approach is self-supervised image dehazing framework. The hybrid attention transformer blocks adaptively leverages both cross-attention and self-attention to effectively model hazy densities via cross-modality fusion and captures global context information for better feature representation. The advantage of this approach is that it improves model generalizability across datasets.

The review of [1][2][3] enabled us to better understand how to design an attention-based enhancement model without employing SoftMax layer, with minimum resources and with less computational cost and time and the results are promising showcasing its efficacy.

Contribution of the proposed approach is a self-guided attention mechanism to capture low-, middle-, and high- level features for image enhancement. The proposed approach is a variant to [1][2][3]. It does not consider any priors as heuristics to guide enhancement. As we find signal-to-noise ratio in [1], different attenuation levels in RGB channels in [2], depth prior and hazy densities [3]. And no SoftMax layer for attention mechanism [1]. The deep attention architecture proposed is a new way of calibration inspired by [19] and implemented through: logarithmic transformation, CNN layers, Dense layers, Skip connections, Concatenation, aggregation using Summation operation, Subtraction operation, finally refinement process to restore colours and smoothed image. The variant way of operations can be viewed as the noise adjustment component and can be considered as simulation of attention mechanism and this also helps in reduction of number of model parameters. The logarithmic transformation helps in enhancement process across different conditions of the image intensity distribution when compared to linear spacing mechanism. And this type of attention mechanism can improve brightness, contrast, colour, and enhance obscured details and remove noise and combination of noises in images of JPEG and PNG format and thus enhanced image of high quality is obtained. The enhanced image obtained is visually appealing with preserved edges, smooth flat areas, fine details with suppressed noise. There are also studies based on diffusion priors, semantic priors [20].

The remaining sections are organized as Literature Review in Section 2, expounding on Proposed Methodology in section 3, Experimental Results and Discussion in section 4 and finally Conclusion and Future work in section 5.

## 2. Literature Review

Image enhancement is a process of improving the colour, contrast and brightness and elimination of noise present in the image. The field in which the image enhancement is applied are medical and non-medical such as acoustic and surveillance sensing. The methods used for image enhancement can be categorized into: Traditional methods, Deep Neural Networks and Evolutionary methods.

Traditional techniques are histogram equalization and its variants such as Brightness Preserving Bi-Histogram equalization(BBHE), Quantized BHE, Histogram Specification (or Histogram Matching) , Filtering mechanism using various kinds of filters such as mean, median, Gaussian, Laplacian, Weiner and variants such as their combination, Guided filtering , Adaptive filtering, Contrast Stretching using opponent colour transform, Gamma Correction function, Unsharp masking , Homomorphic filtering and Retinex - Based methods. The filtering methods are affected by the size of the filter. The filter's receptive field influences the level of denoising. The colour transformation methods are mathematical functions which require domain knowledge and it does not consider the influence of intensity on the neighbourhood which may lead to artifact creation in the constructed image. The homomorphic filtering and Retinex based methods their variants such as Single-Scale Retinex, Multi-Scale Retinex, Attention-Guided Deep Retinex are based on decomposition of an image into reflectance and illumination components. All these methods focus on restoring brightness and contrast preserving edges but these methods are source dependent and require tuning of parameters for optimization. They may either result in over- or under enhancement of the image.

The deep learning techniques learn valuable portrayals from the input image and constructs the image of good fidelity and quality. The various kinds of deep neural networks techniques employed for image enhancement are Convolutional Neural Networks (CNN), Residual networks, Autoencoders, Generative Adversarial Networks (GANs) and their variants such like Denoising U-Shaped Net, Batch Renormalization U-Net, Dilated CNNs, Autoencoder Denoising Network (Di-Conv-AE-Net), Disco-GAN, Bicycle GANs and Hybrid models. The hybrid models are built by integrating the different neural network topologies to yield high performance results despite being complex to build. The Convolutional Neural Network [5] is trained using supervised learning manner to construct super resolution image. This method even though fast and accurate but handling of large and non-uniform kernel hinders its performance. [6] Two sequential yet interconnected autoencoders network framework is used. The first focuses on noise suppression while the second refines the details. The benefit of this method is that it can denoise any type of noise by identifying image's meaningful details (or structure). [7] A two stage enhancement approach built using dilated convolutional neural network using complimentary information from the objective loss and perceptual loss is used to enhance the image. The dilated network even though averts usage of pooling averts but it decreases the image resolution. [8] the convolutional neural network and autoencoder with their identified activation functions is designed to facilitate optimization process. The activations Tanh and leaky-ReLu in autoencoder, Sigmoid and SoftMax for CNNs are employed. [9] the residual network to improve the model stability has been used. The residual connections also called as skip connections address the vanishing gradient problem but the drawback is that it requires substantial processing resources that affect their training and memory requirement. [10] multi-scale residual blocks are used to exchange the multi-resolution streams with Charbonnier loss to optimize the neural network. But the drawback is that it requires setting of epsilon parameter in the Charbonnier loss function for thresholding the performance. [11] An Attention mechanism is used. The first network focuses on distinguish noise from texture and the second network uses this information as prior to construct the enhanced image.[12] the 1D- Wasserstein distance is used to improve the perceptual performance. [13] the variant of Wasserstein GAN, the Rectified Wasserstein Generative Adversarial Network (ReWaGAN) is employed for unsupervised learning with rectified gradient in the generator during training along with perceptual and distortion loss trade-off established to optimize the restoration of high-quality image. [14] Conditional GAN(CGAN) for denoising a digital holographic interferometry phase measurement is employed. In this method three loss functions: Discriminator Loss, PSNR Loss and L1-norm Loss are used to construct the high-quality enhanced image. GAN-based models require large memory footprint which restrict their practical application in resource constrained devices. The success of Deep Neural Networks (DNNs) ability for image enhancement can be attributed to its mechanism of automatic feature extraction and learning from the given input image and their optimization enables in overcoming of over fitting/Under fitting problem. The speed of training and data augmentation is complimented by GPUs which further extended their usage [16, 17].

Evolutionary algorithms include Genetic algorithms, Fuzzy Neural Networks. The current parallel processing ability of machines extended their ease of usage and designing of high-performance models for image enhancement.

## 3. Proposed Methodology

The image of low quality often suffers from noise, intensity variations and colour variations. The proposed approach learns intensity distribution present in a single-color image and enhances the image using simple unsupervised attention-based deep neural network model in spatial domain by learning intensity distribution and their relationship within and across the channel. This approach is based on channel attention and spatial attention and then followed by refining mechanism to obtain enhanced image.

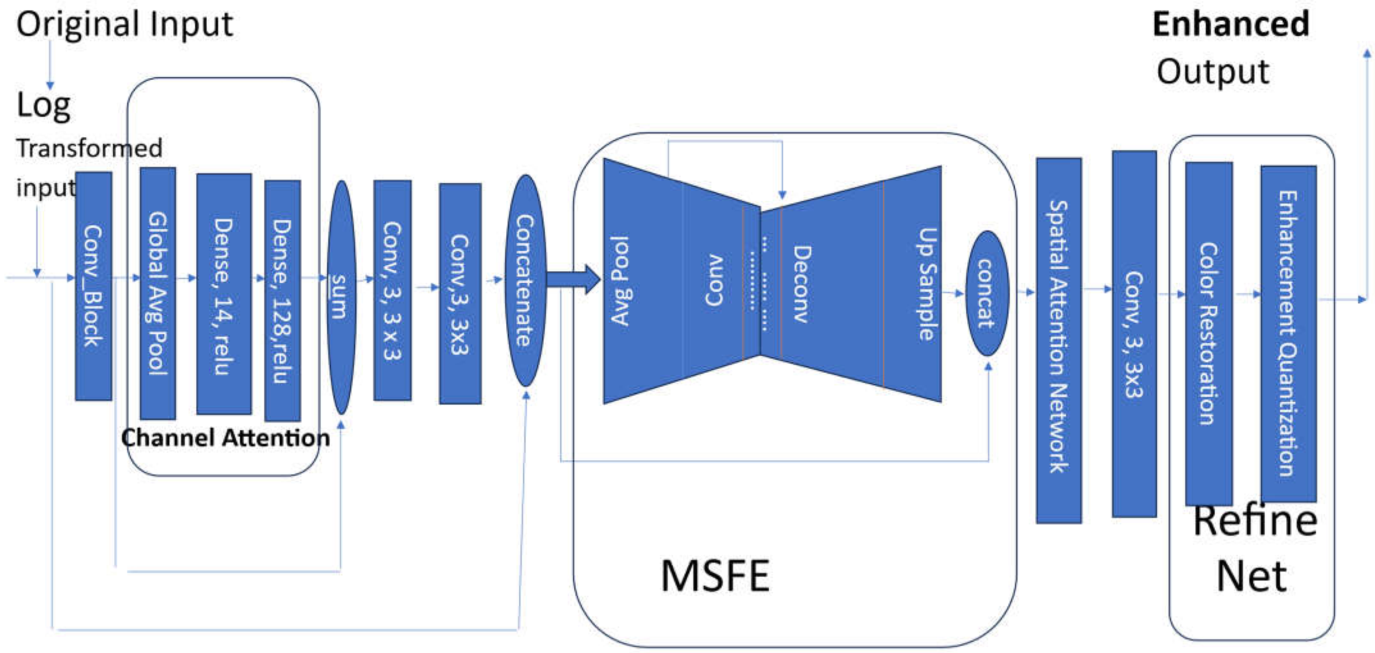



**Figure 1:** Proposed Architecture Block Diagram

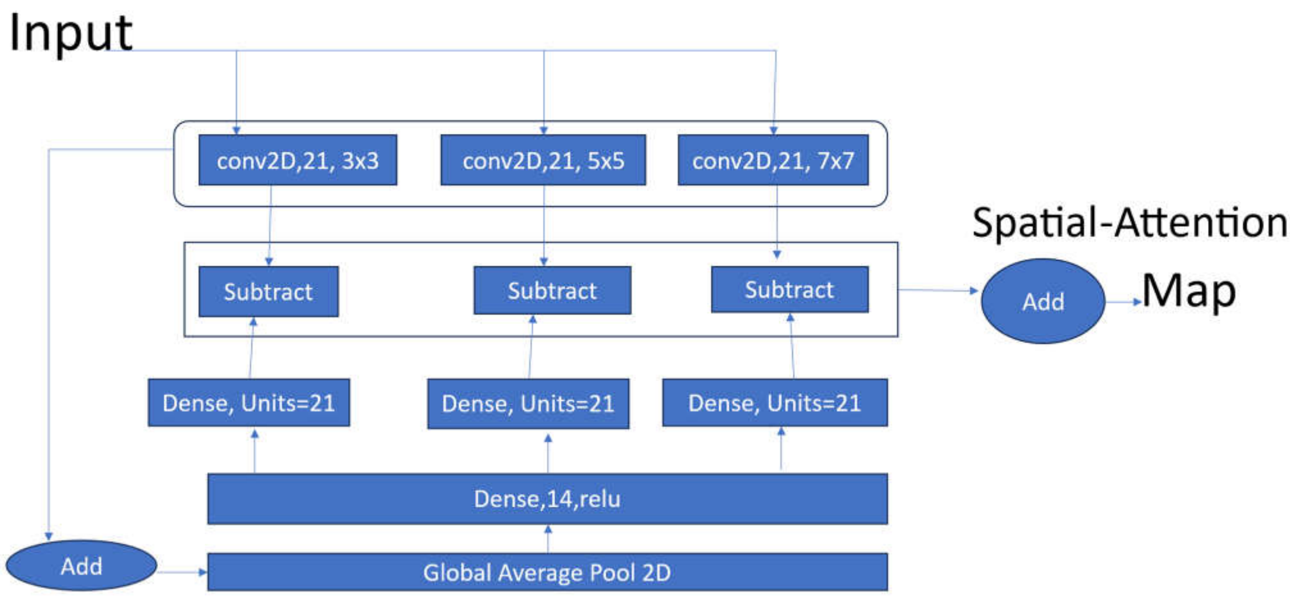



**Figure 2:** Spatial Attention Network Block Diagram

The input received by the network architecture is a logarithmic transformed input. This pre-processing step helps to amplify the intensity and suppresses noise. The logarithmic transformation converts the skewed distribution to a normal or less skewed distribution. Then these logarithmic transformed values are given as input to convolutional block. The convolutional block is made up of five convolutional layers with kernel of 3 x 3 size, with 128 channels and ReLu activation function. Each layer in convolutional block extracts features by convolutional operation. The convolved feature map is obtained by sliding the kernel matrix over the given input and computing dot operation. The convolution operation brings the intensity values into closer harmony with its neighbour. The odd size kernel has specific centre pixel and during convolution operation the centre of kernel traverse all pixels of the given input. The ReLu activation function introduce non-linearity and generates rectified feature map which can be represented as:

$$h^{i} = \max(0, \mathrm{x} * k^{i} + b^{i})$$

where $h^i$, $k^i$ and $b^i$ are the $i^{th}$ feature map, kernel, and bias respectively. And 'x' is the input given to convolutional layer. '*' represents convolutional operation. Where i takes the value 1 to 5.

The output of convolutional block is then given to channel attention block. Since selection of feature scales impacts significantly on the performance of enhancement task and moreover image features have strong correlation between different channels. Hence the global averaged feature map is passed on to two dense layers with 14, 128 units of neurons and relu activation function respectively. The dense layer is a fully connected layer where each layer receives an input from all neurons present in the previous layer. Hence, dense layers enable learning of all combinations of features. Whereas a convolutional layer relies on consistent features within a receptive field of kernel. The global averaged information help in evaluation of pixel intensity distribution that is present in given input which play a key role for enhancement process. Then the output from this channel attention block and input of channel attention block are summed. Because summation operation facilitates avoidance of SoftMax layer. The SoftMax layer calculates the feature weight score whereas the summation operation is a process that can be viewed as the noise cancellation component. Then these summed up values are given as input to convolutional layer call it as 'conv1'. Then followed by conv2. These two convolutional layers uses kernel size of 3 x 3 with number of channels being 3. The purpose employing only two convolutional layers is that the size of data becomes smaller and smaller as the depth of convolutional layers increases even though deep convolutional layers facilitate in complex feature extraction. Thus, depth of CNN is constrained. The number of channels utilized in these two layers are three because the channel depth and feature depth should be consistent. Then output of conv2 is fused with original input of the architecture i.e., logarithmic transformed input. Using these fused values the MSFE network extracts multi-granular features and it even helps in improving the reliability of the model. This MSFE network is a U-Shape network with average pool layers and convolutional layers for down sampling and deconvolutional layers followed by up sampling. The U-Shaped network helps to weaken irrelevant information. Up-sampling is carried out using bilinear interpolation. Bilinear interpolation is a resampling method that uses the distance-weighted average of the four nearest pixel values to estimate a new pixel value. The bilinear interpolation assumes that the surface is continuous and the neighbouring pixels are related. The up-sampling rates used are: 16 x 16, 8 x 8, 4 x 4, 2 x 2, 1 x 1. These hierarchical features are concatenated with the input supplied to this MSFE module which helps to break symmetry and in turn facilitating restoration of the missing details. Because of asymmetric nature of output of MSFE network and the loss of resolution in the output of MSFE, the output from MSFE is given as input to spatial attention module for refining of coarse features. The Spatial Attention Module (SAM) help to recover fine details along with retention of spatial connectivity pattern and helps to avoid structural distortion in the enhancement process. This module also facilitates in controlling the overall number of parameters of the architecture proposed and even improves the accuracy of the image constructed. The output of SAM is passed on to convolutional layer to reshape tensor. Then it is followed by refinement. The Refine network restores the colour and helps in rendering the image of high fidelity. The primary reason for Refine network is that since sub-sampling operations leads to significant loss of information. So, the Refine network performs colour enhancement and quantization of intensity to construct an enhanced image. Thus, image with rich details, enhanced colour, brightness, with no noise is obtained.

The model is complied with Adam optimizer with learning rate of 0.000005, and Mean Absolute Error: L1-norm regularizer. The regularizer enables to overcome overfitting problem and rate distortion performance. Model is compiled to: max $||F(x;\ \Theta) - x||$ where: F (x; Θ) denotes the network output, x the network input, Θ is the network parameters, F the proposed network. The advantage of Adaptive Moment (Adam) optimizer is that it can dynamically adjust the learning rate for each individual parameter within a model, rather than using a single global learning rate. It requires no or only little tuning to achieve strong performance in less time. Commonly, local average pooling tends to blur sharp discontinuities in the image so our proposed model architecture utilizes both local- and global- features to boost the performance and the skip connections prevent loss of information when information passed through multiple modules and thus the robustness and generalizability of the model is achieved. The proposed model can denoise all types of noises including mixed type of noise, improves brightness, contrast, and preserve rich details. And the logarithmic transformation helps in enhancement of an image across different conditions when compared to linear spacing.

### 3.1. Proposed Enhancement Algorithm

Step 1: Read low quality input colour image

Step 2: Divide the input image channel-wise: R, G, B.

Step 3: Compute logarithmic value for each channel intensity values if the intensity value is not zero.

Step 4: Stack the step 3 values channel wise to obtain 3Dimensional array.

Step 5: Predict the features using Channel Attention module then followed by Spatial Attention module.

Step 6: Construct the enhanced image using Refine Network

### 3.2. Experimental Setup

Table 1: Hardware Specification of CPU and GPU

| Processor | CPU | GPU |
|---|---|---|
| Number of cores | 6 Core processor | 1408 CUDA Core |
| Clock Frequency | 3.60 GHz | 1785 MHz |
| Memory Size | 112 GB | 62 GB |

Software Specification

TensorFlow an open-source Machine learning platform.
Keras a deep learning API that can run on TensorFlow framework.
Programming Language: Python
Interface: Jupyter Notebook an interactive computational environment which can run python code

Data Used

Data Set: Baboon, Lena, House, Cannon-LR, LoLv1, Parrot, Dog image
Input Image Size: 256 x 256
Image Codec: JPEG, PNG.

Table 2: Model Details

| Layer (type) | # Channels | # Parameters |
|---|---|---|
| input_layer (InputLayer) | 3 | 0 |
| convolutional_block_1 | 128 | 593920 |

| | | |
|---|---|---|
| channel_attention | 128 | 3726 |
| Conv2D | 3 | 3459 |
| Conv2D | 3 | 84 |
| Concatenate | 6 | 0 |
| multi_scale_feature_extraction | 21 | 290470895 |
| kernel_selecting_module | 21 | 37919 |
| Conv2D | 3 | 570 |

Table 3: Model Architecture Settings and Values

| Settings | Values |
|---|---|
| Number of Model parameters | 291,110,573 |
| Runtime | 35 seconds per step |
| Number of steps | 8 |
| Input Image Dimensions | 3 |
| In-channels | 128 |
| Out Channels | 3 |
| Multi-Scale Rate | [16, 8, 4, 2, 1] |
| Conv-resampling | True |

## 4. Results And Discussion

The quality of the output image produced is assessed using following indicators: MSE, PSNR and SSIM.

MSE, PSNR, SSIM is calculated between ground-truth real image and the produced enhanced image.

MSE calculates the average square difference between the pixel values in ground-truth real image and the enhanced output image produced. Mean Squared Error is given by the formula:

$$\mathrm{MSE} = \left( \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{[\,\mathrm{P(i,j)} - \mathrm{C(i,j)}]\text{^}2}{W * H} \right)$$

where: M, N are Width and Height of image respectively. P(i, j) is the pixel value in the ground-truth real image, and C (i, j) is the pixel value in the enhanced output image produced at location (i, j)

PSNR assess quality of the image by comparing maximum signal intensity to interfering noise. It measures the noise ratio between the ground-truth real image and produced enhanced output image. If PSNR is large, it means that the distortion is small and the image produced is of good quality. The presence of the degradation reduction is indicated by high PSNR value that means there is a noticeable difference between ground-truth real image and enhanced image produced is less which indicates that the predicted image is of good quality. Peak-Signal-to-Noise Ratio is given by the formula:

$$\mathrm{PSNR} = 10 * log_{10}(255^2)/(MSE)$$

Where the value 255 in the formula indicates that pixel is 8-bit.

The Structural Similarity Metric assess the visual impact of three characteristics of an image: Luminance, Contrast, and Structure. Mean value is used for estimating brightness. Standard deviation is used for estimating contrast. Covariance is used for measuring the similarity of the structure.

SSIM is measured between ground-truth real image and the enhanced output image produced. SSIM is given by the formula:

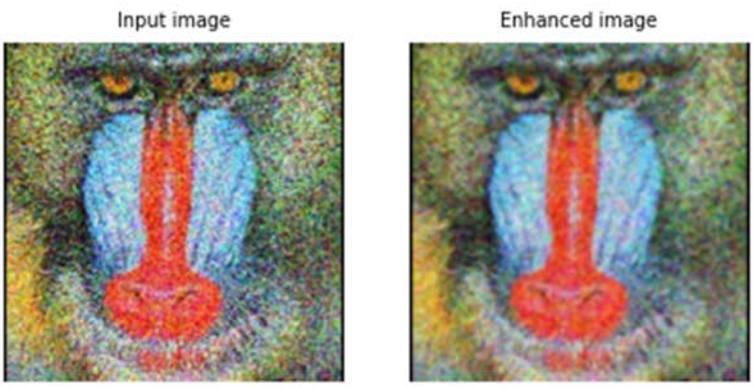
$$SSIM(X,Y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{({\mu_x}^2 + {\mu_y}^2 + c_1)({\sigma_x}^2 + {\sigma_y}^2 + c_2)}$$

Where $\mu$ is the mean, $\sigma$ is the standard deviation., $c_1 = (K_1L)^2$, $c_2 = (K_2L)^2$ where $K_1 = 0.01$, $K_2 = 0.03$ by default, L is the dynamic range of pixel values, here L= 255.

SSIM being equal to 1 indicates that the produced output image is like the ground-truth real image. In other words, high precision in visual representation when compared to ground-truth real image.

Table 4: MSE, SSIM, PSNR results between input image and Enhanced image.

| Input Image | MSE between Ground-Truth and input Image | MSE between Ground-Truth and Enhanced Image | SSIM between Ground-Truth and Enhanced Image | PSNR between Ground-Truth and Enhanced Image |
|---|---|---|---|---|
| Baboon (JPG) | 1256.8647 | 764.42267 | 0.9999999 | 19.29746803 |
| Cannon-001-LR (PNG) | 24.2438 | 134.27682 | 0.9999999 | 26.8507929 |
| Low-Illuminate Image (LOL-V1-1) (PNG) | 23157.75 | 16663.209 | 0.99999 | 5.913217 |
| Gaussian+ Random Noised Lena (PNG) | 3485.9746 | 3150.82 | 0.9999999 | 13.14656757 |
| Dog (JPG) | 2227.361 | 1581.9523 | 0.9999999 | 16.138869 |





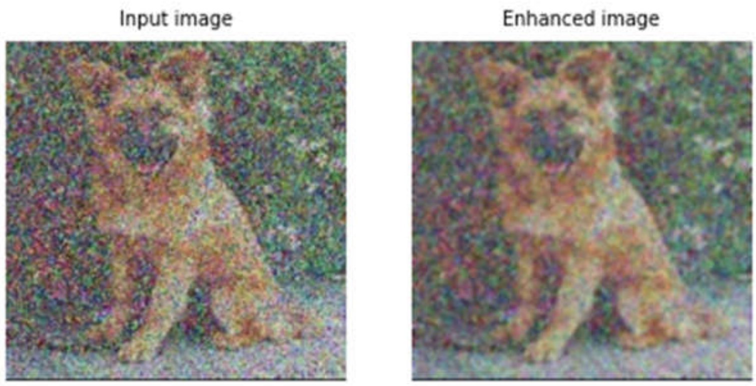



Table 5: MSE, SSIM, PSNR results between input PNG Codec Noised Image and Enhanced image.

| Input Image (PNG) | MSE between Ground-Truth and input Image | MSE between Ground-Truth and Enhanced Image | SSIM between Ground-Truth and Enhanced Image | PSNR between Ground-Truth and Enhanced Image |
|---|---|---|---|---|
| Speckled Noise Parrot Image | 757.7813 | 703.0328 | 0.99999 | 19.661047 |

| | | | | |
|---|---|---|---|---|
| Gaussian Noised House Image | 939.3710 | 739.89233 | 0.99999 | 19.439118 |
| Salt-Pepper Noised House Image | 966.1252 | 643.5095 | 0.99999 | 20.045253 |

Table 6: Experimental Results of some studies for comparison

| | PSNR | SSIM |
|---|---|---|
| LOL Ref[1] | 23.94 | 0.842 |
| Speckled Noise Ref[1] | 22.95 | 0.796 |
| Canon-LR Ref[1] | 29.94/29.59 (SD indoor/outdoor) | 0.906/0.87 |
| Gaussian Noise Ref [1] | 20.14 | 0.617 |
| Ref[18] | 29.23 | 0.843 |
| Ref[19] | 23.45 | 0.85 |
| LOL/Dehaze Ref [22] | 24.87/30.16 | 0.9212/0.967 |
| LOL001 Dataset Ref[23] | 28.0 | 0.859 |

For an N x N input, K x K filters, R filters per layer, and L number of layers then cost of evaluating the convolutional operation grows like $O(N^2K^2R^2L)$. The memory requirement only depends on size of filters and bias therefore the cost is $O(LK^2R^2)$. The computational complexity linearly scales with the number of architectural variables and size of dataset. When we compared our results, we find our SSIM > 0.9999 and even computational complexity is less. The proposed model is versatility. It can enhance any type of noised colour image including mixture of noise of variant file formats with reasonable performance and with no pre-training, no storage of feature weight score with high quality rendition.

The ablation Study reveal when experiment is conducted without logarithmic transformed input then the performance of the proposed model gets affected. Even the influence of Refine network module can also be observed. Despite being relatively good performance methodology, the enhancement method works well if only dynamic range of pixel values is below certain threshold. If the image has already acceptable level of contrast and brightness then enhancement using refine network would introduce noise artifacts and the image will not look well enhanced as we can see from below table. When model is compiled with L2-norm instead of L1-norm then the performance of the model becomes low. Because L2-norm leads to linear scale-space whereas L1-norm leads to total variation flow. We can observe this from a sample result of Speckled Noised image using L2-norm: MSE: 809.70, PSNR: 19.04, SSIM: 0.999. Even the interpolation operation is affected by the neighbour size chosen. For colour correction there is no one size fits all solution. It is subjective and best solution depends on image processing application needs.

Table 7: Experimental Results for ablation

| Input Image | MSE between Ground Truth and input image | MSE between Ground-Truth and image obtained before Refine | MSE between Ground-Truth and Enhanced image due to Refine Network | MSE between input and output image obtained from model without Logarithmic Transformed input |
|---|---|---|---|---|
| Baboon (JPG) | 1256.8647 | 1242.310038 | 764.42267 | 5534.191 |
| Cannon-001-LR (PNG) | 24.2438 | 24.44668 | 134.27682 | 1167.9755 |
| Low-Illuminate Image (LOL-v1-1) (PNG) | 23157.75 | 23151.486 | 16663.209 | 12311.081 |
| Gaussian+ Random Noised Lena (PNG) | 3485.9746 | 3485.9578 | 3150.82 | 5546.885 |
| Dog (JPG) | 2227.361 | 2225.812 | 1581.9523 | 1552.0609 |
| Speckled Noise Parrot Image (PNG) | 757.7813 | 757.36444 | 703.0328 | 6739.8774 |
| Gaussian Noised House Image(PNG) | 939.3710 | 939.2759 | 739.89233 | 5760.018 |
| Salt-Pepper Noised House Image(PNG) | 966.1252 | 965.5938 | 643.5095 | 8024.931 |

## 5. Conclusion And Future Work

The basic principle of image enhancement is to process an image so that the outcome is more suitable than the original image for a particular specific application. “Particular” means a certain application. This proposed approach of image enhancement is carried out in spatial domain using low- middle-, high –level features. It can remove any type of noise, improves brightness, contrast in a colour image and applicable to images of different file format and the designed model is tested with image of variant datasets: LOLv1, Set14, Cannon-LR images. It does not require pretraining, large dataset nor reference image in other words no image pair for enhancement process. Commonly, optimization of models with no reference learning is complex and difficult but our attention mechanism is an unsupervised learning-based and can restore details, colour efficiently and thus output image produced is of high quality and visually appealing. Proposed architecture requires minimal resources and involves simple computations and less computational time. It does not require SoftMax layer as commonly used in most of the existing methods for features weight score. The proposed architecture is based on efficient calibration of feature score vector with an intent of trade-off between modelling effectiveness and computational complexity and resources and it is even versatile since it can enhance image of different codecs. The proposed method achieved high SSIM when compared to other studies thus rendition of image is accurate. Future work is to deal with aggressive enhancement case and to study the influence of proposed mechanism in cross sections of image processing tasks compression and encryption.

**References:**

[1]. Yan Wang, Peng Su, Xiaoying Pan, Hongyu Wang, Yuan Gao, Channel Self-Attention Based Low-light Image Enhancement Network,Computers & Graphics, Vol:120,2024, DOI: https://doi.org/10.1016/j.cag.2024.103921

[2]. Li Hong, Xin Shu, Qi Wang, Hua Ye, Jinlong Shi, Caisheng Liu, CCM-Net: Color Compensation and Coordinate Attention Guided Underwater Image Enhancement with Multi-Scale Feature Aggregation, Optics and Lasers in Engineering, Vol:184, Part 1, Jan 2025,108590, DOI: https://doi.org/10.1016/j.optlaseng.2024.108590

[3]. Yu Luo, Guoliang Lv, Jie Ling, Xiaomin Hu, Low-light Image Enhancement Via an Attention-Guided Deep Retinex Decomposition Model, Applied Intelligence, Vol:55, Article Number:194,2025, DOI: https://doi.org/10.1007/s10489-024-06044-2

[4]. Yudong Liang, Shaoji Li, De Cheng, Wenjian Wang, Deyu Li, Jiye Liang, Image dehazing via self-supervised depth guidance Pattern Recognition, Vol:158, Feb 2025,111051, DOI: https;//doi.org/10.1016/j.patcog.2024.111051

[5]. O. Soufi, Z. Aarab and F. -Z. Belouadha, Benchmark of deep learning models for single image super-resolution (SISR),2022 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), Meknes, Morocco, 2022, pp. 1-8, DOI: 10.1109/IRASET52964.2022.9738274

[6]. .Ehab Alaa Saleh, Mostafa I. Soliman, Mohamed Abdel-Nasser, Speckle Noise Suppression in Digital Images utilizing Deep Refinement Network, Mendel Soft Computing Journal, Vol.30 No.1,2024, DOI: https://doi.org/10.13164/mendel.2024.1.015

[7]. Linli Xu, Peixian Liang, Jing Han, Lianfa Bai, Danny Z.Chen, A Two-Stage Enhancement Network with Optimized Effective Receptive Field for Speckle Image Reconstruction, Multimedia Tools and Applications, Vol 82, pp 19923-19943,2023, Published: 23 November 2022, DOI: https://doi.org/10.1007/s11042-022-14208-w

[8]. M.Fatih Yilmaz, Bekir Karlik, Comparison of Deep Learning Algorithms with Different Activation Functions for Brightness Image Enhancement, International Journal of Artificial Intelligence and Expert Systems (IJAE), Vol.13, Issue.2,2024, Pages - 12 – 24, Published in International Journal of Artificial Intelligence and Expert Systems (IJAE)

[9]. Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, Yun Fu ,Residual Dense Network for Image Restoration, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.43, No.7, July 2021, DOI: 10.1109/TPAMI.2020.2968521

[10]. Zamir, Syed Waqas & Arora, Aditya & Khan, Salman & Hayat, Munawar & Khan, Fahad & Yang, Ming-Hsuan & Shao, Ling. (2020). Learning Enriched Features for Real Image Restoration and Enhancement. DOI: 10.1007/978-3-030-58595-2_30.

[11]. Feifan Lv, Yu Li, and Feng Lu. 2021. Attention Guided Low-Light Image Enhancement with a Large Scale Low-Light Simulation Dataset. Int. J. Comput. Vision 129, 7 (Jul 2021), 2175–2193. DOI: https://doi.org/10.1007/s11263-021-01466-

[12]. Delbracio, Mauricio & Talebei, Hossein & Milanfar, Pevman. (2021). Projected Distribution Loss for Image Enhancement. 1-12. DOI: 10.1109/ICCP51581.2021.9466271.

[13]. Haichuan Ma, Dong Liu, Feng Wu, Rectified Wasserstein Generative Adversarial Networks for Perceptual Image Resolution, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume: 45, Issue: 3, 01 March 2023, Page(s): 3648 – 3663, Trans.Pattern Anal.Mach.Intell. DOI: 10.1109/TPAMI.2022.3185316

[14]. Qiang Fang, Haiting Xia, Qinghe Song, Meijuan Zhang, Rongxin Guo, Silvio Montresor, Pascal Picart, Speckle Denoising Based on Deep Learning via a Conditional Generative Adversarial Network in Digital

Holographic Interferometry, Optics Express, Vol. 30, Issue 12, pp. 20666-20683, (2022), DOI: https://doi.org/10.1364/OE.459213

[15]. Codruta O.Ancuti, Cosmin Ancuti, Christophe De Vleeschouwer, Mateu Sbert, Color Channel Compensation (3C): A fundamental Pre-Processing Step for Image Enhancement, IEEE Transaction on Image Processing Vol.29, Page(s): 2653 – 2665, Date of Publication: 14 November 2019, DOI: 10.1109/TIP.2019.2951304

[16]. R.Geetha, G. Belshia Jebamalar, S.Arumai Shiney, Nhu-Ngoc Dao, Hyeonjoon Moon, and Sungrae Cho, Enhancing Upscaled Image Resolution using Hybrid Generative Adversarial Network-Enabled Frameworks, IEEE Open Access Journal, Vol.12,2024 DOI: 10.1109/ACCESS.2024.3367763, Page(s): 27784 – 27793, 20 February 2024

[17]. Shao Junzhe, Zhang Zhibin, Unsupervised low-light Image Enhancement by data augmentation and contrastive learning, The Imaging Science Journal, DOI: https://doi.org/10.1080/13682199.2024.2395751

[18]. Rusul sabah jebur, Mohd Hazli Bin Mohamed Zabil, Lim Kok Cheng, Dalal A. Hammood, Image Denoising in Deep Learning: A Comprehensive Survey, Electrical Engineering Technical Journal, ISSN:3007-9772, Vol:1, No.1,2024,

Pages1-12, DOI: https://doi.org/10.51173/eetj.v1i1.2

[19].Peipei Zhang, Image Enhancement Method Based on Deep Learning, Hindawi Mathematical Problems in Engineering, Vol.2022, Article ID:6797367, 9 pages, DOI: https://doi.org/10.1155/2022/6797367

[20]. Senthil Anandhi, A., Jaiganesh, M., An enhanced image restoration using deep learning and transformer based contextual optimization algorithm. Sci Rep 15,10324 (2025) DOI:10.1038/s41598-025-94449-5

[21]. L. Ma, T. Ma, R. Liu, X. Fan and Z. Luo, Toward Fast, Flexible, and Robust Low-Light Image Enhancement, 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) ,June 2022,Date Added to IEEE Xplore: 27 September 2022, DOI: 10.1109/CVPR52688.2022.00555

[22]. Kui Jiang, Ruoxi Wang, Yi Xiao, Junjun Jiang, Xin Xu, Tao Lu, Image Enhancement via Associated Perturbation Removal and Texture Reconstruction Learning, IEEE/CAA JOURNAL OF AUTOMATICA SINICA, VOL. 11, NO. 11, NOVEMBER 2024 2253, Digital Object Identifier 10.1109/JAS.2024.124521

[23]. Manuel J. C. S. Reis, Low-Light Image Enhancement Using Deep Learning: A Lightweight Network with Synthetic and Benchmark Dataset Evaluation, Appl. Sci. 2025, 15,6330. https://doi.org/10.3390/app15116330

[24]. G.Ruth Rajitha Rani, Ch.Samson, A Survey on Image Enhancement Techniques, International Journal of Computational Engineering Research (IJCER), Vol.09, Issue.3, 2019