IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Toward Real-Time Low-Light Enhancement: Deployment-Friendly Sci Framework With Onnx Integration

[1]P Nancy Celine, [2]Venusai Mabbu, [3] Vinay Kumar Manda

[1]Assistant Professor, [2]Student, [3] Student

[1]Computer Science Engineering,

[1]Vasavi college of Engineering, Hyderabad, India

Abstract: : Enhancing images captured in low-light conditions is crucial for improving visibility and overall quality, which directly benefits a wide range of computer vision tasks. However, traditional methods and many learning-based models often struggle in real-world environments due to issues like overfitting, poor generalization, or high computational demands. In this study, we investigate a novel low-light image enhancement framework called Self-Calibrated Illumination (SCI), designed to be efficient, adaptable, and effective in practical scenarios. SCI employs a cascaded illumination estimation strategy combined with a unique self-calibrated module, enabling the model to deliver high-quality enhancement with low computational overhead. Unlike typical deep learning models that require multiple processing blocks during inference, SCI achieves comparable or better results using a single lightweight block, thanks to its convergence-focused training approach. Through comprehensive experiments on benchmark datasets, we validate SCI's advantage in both image quality and performance efficiency. This report provides an in-depth exploration of SCI's methodology, implementation details, and evaluation, positioning it as a strong candidate for real-world low-light image enhancement.

Index Terms - Low-light image enhancement, Self-Calibrated Illumination (SCI), unsupervised learning, real-time enhancement, ONNX, computer vision.

I. Introduction

In real-world environments, photography and visual data acquisition are often constrained by poor illumination conditions. Images captured under such low-light scenarios suffer from various degradations including low visibility, reduced contrast, high noise levels, color distortion, and detail loss. These quality impairments not only affect human visual perception but also deteriorate the performance of automated computer vision systems such as object detection, face recognition, and semantic segmentation. Consequently, low-light image enhancement has emerged as a crucial pre-processing step in both consumer photography and professional applications like surveillance, autonomous driving, and medical imaging.

Figure-1: Examples of Low-Light Image Degradation

In the past, conventional model-based methods with their roots in image formation theories dominated enhancement techniques. For example, the Retinex theory suggests that an image can be broken down into components related to illumination and reflectance. By varying brightness while maintaining structural details, methods based on this theory usually try to estimate the illumination map and reconstruct the image. High model complexity, overexposure artifacts, and inconsistent color reproduction are still issues they must deal with, though.

In this context, the Self-Calibrated Illumination (SCI) framework presents a novel solution that bridges the gap between performance, flexibility, and computational efficiency. The SCI model was developed and trained using the PyTorch deep learning framework due to its flexibility and ease of use.

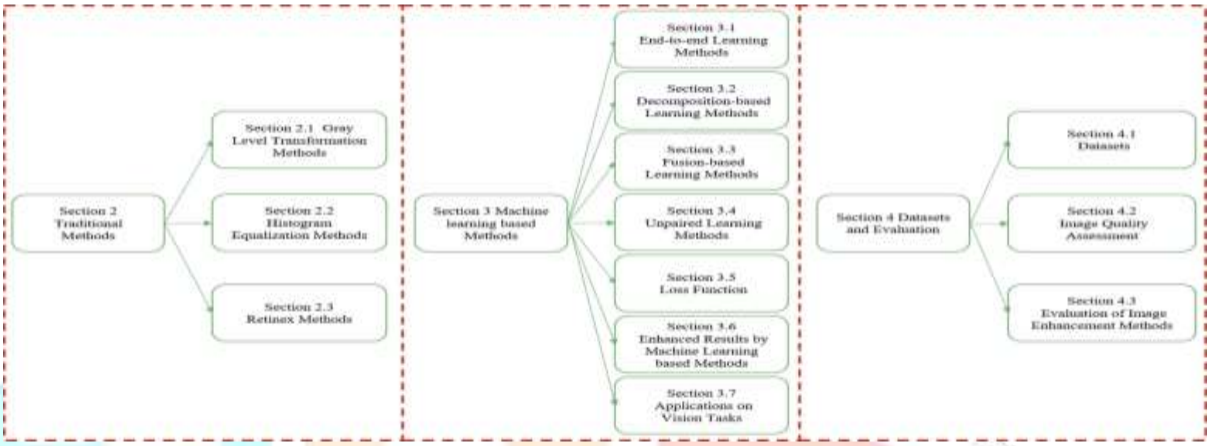

Figure-2: Structure of the Literature Survey

SCI employs an unsupervised loss formulation comprising smoothing and fidelity losses. The smoothing loss promotes spatial coherence in the illumination maps, while the fidelity term enforces consistency between the estimated and actual inputs.

As a result of these losses, paired training data is no longer necessary, enhancing the model's applicability across a wide range of uncontrolled and diverse environments.

In this report, we thoroughly examine the SCI framework in terms of its theoretical foundation, architectural design, training methodology, and empirical performance. We aim to provide a clear understanding of how SCI addresses the shortcomings of previous approaches and why it represents a promising step forward in the domain of low-light image enhancement.

II. LITERATURE SURVEY

Low-light image enhancement has been an extensively researched problem in the field of image processing and computer vision. The goal is to restore image visibility and improve overall perceptual quality under poor illumination, while preserving structural details and suppressing noise. The literature on this topic can broadly be categorized into model-based approaches, supervised deep learning methods, and unsupervised or self-supervised learning techniques.

2.1 Model-Based Approaches

The earliest methods for low-light image enhancement were grounded in physical models and image formation theories. A foundational idea in this domain is the Retinex theory, which assumes that an observed image can be decomposed into reflectance (the true image content) and illumination (the varying light conditions). Techniques such as LIME [Guo et al., 2017] and LECARM [Ren et al., 2018] estimate the illumination map using mathematical priors such as total variation or entropy-based measures, and reconstruct the image accordingly.

While model-based approaches are interpretable and relatively lightweight, they typically require hand-crafted regularizers, involve complex optimization routines, and are sensitive to parameter settings. Moreover, they often produce artifacts like overexposure, detail loss, or unnatural colors in complex scenes with spatially varying illumination.

2.2 Supervised Deep Learning Methods

With the advent of deep learning, supervised methods began to dominate the low-light enhancement landscape. These approaches leverage convolutional neural networks (CNNs) to learn mappings between dark input images and their corresponding bright ground-truth outputs. For instance, RetinexNet [Wei et al., 2018] extended the Retinex decomposition into a trainable neural network, learning separate subnetworks for reflectance and illumination. KinD [Zhang et al., 2021] improved upon this by introducing refined loss functions and a more stable architecture.

To improve performance, other approaches like FIDE, DRBN, and DeepUPE added specialized elements like exposure control layers, recursive modules, or attention mechanisms. Although these techniques produce excellent results on carefully selected datasets, their dependence on aligned image pairs limits their applicability to low-light conditions in the wild.

2.3 Unsupervised and Self-Supervised Methods

To overcome the dependency on paired datasets, several unsupervised learning techniques have been developed. These methods use intrinsic image properties or heuristics to define loss functions that guide the network without explicit supervision. A prominent example is ZeroDCE [Guo et al., 2020], which learns parameters for a pixel-wise curve estimation function to enhance brightness. Another noteworthy approach is EnlightenGAN [Jiang et al., 2021], which employs a GAN-based architecture for unpaired training, targeting perceptual and structural consistency.

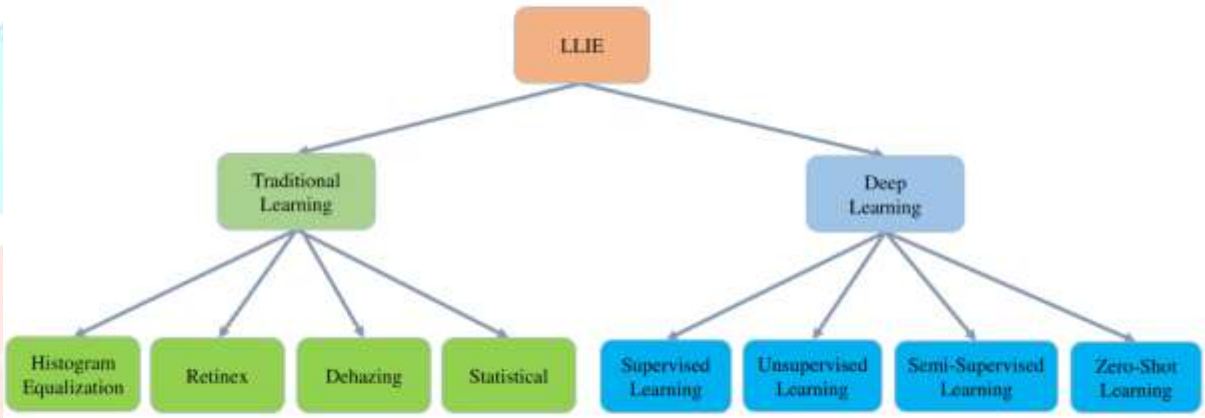

Figure-3: Hierarchy of Low-Light Image Enhancement (LLIE) Methods

2.4 The Emergence of SCI

The Self-Calibrated Illumination (SCI) framework, proposed by Ma et al. in CVPR 2022, represents a significant advancement in the direction of unsupervised low-light image enhancement. SCI integrates the interpretability of Retinex-based decomposition with the learning capacity of neural networks, while specifically addressing the drawbacks of existing methods—namely high computational cost, stage-wise inference complexity, and poor generalization.

SCI introduces two innovative components: a weight-sharing illumination learning process, and a self-calibrated module that enforces convergence of outputs across multiple enhancement stages. This enables the model to be trained using multiple cascaded blocks while requiring only a single block during inference, thus achieving real-time performance without sacrificing quality.

Moreover, SCI is shown to possess two desirable properties rarely addressed in prior works: operation-insensitive adaptability, where performance remains stable across varied architectural configurations, and model-irrelevant generality, allowing the SCI module to be integrated into other enhancement models like RUAS to boost their performance.

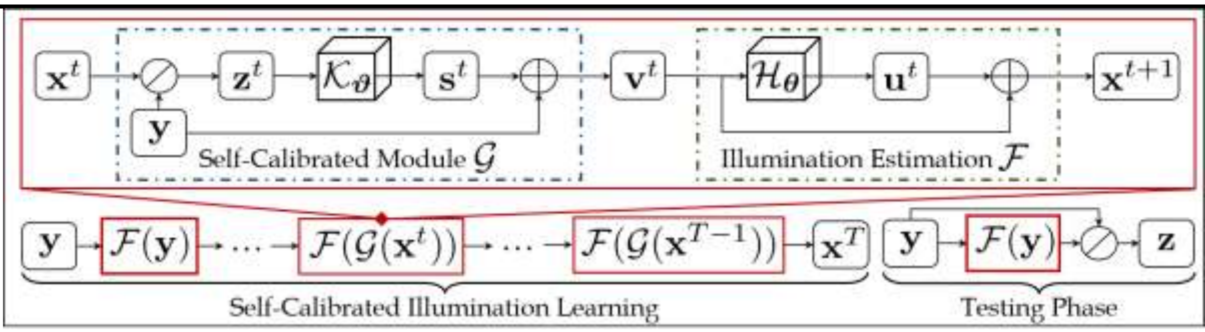

Figure-4: The Self-Calibrated Illumination (SCI) Framework

In conclusion, even though earlier approaches have made a substantial contribution to the field, they are frequently hampered by real-world issues like the requirement for paired data, computational inefficiency, or lack of robustness. Building on these principles, SCI provides a more effective, adaptable, and broadly applicable approach to low-light image enhancement for both scholarly and practical uses.

III. PROPOSED METHODOLOGY

In this section, we present the detailed architecture and underlying principles of the Self-Calibrated Illumination (SCI) framework, a deep unsupervised learning-based model designed for efficient and robust low-light image enhancement. The methodology is structured into four major components:

(1)Illumination Learning with Weight Sharing, (2) the Self-Calibrated Module, (3) the Unsupervised Loss Function, and (4) the Overall Network Architecture and Training Strategy. Each of these components is designed to collectively achieve high-quality image enhancement while maintaining low computational cost and high adaptability.

3.1 Illumination Learning with Weight Sharing

The SCI model is inspired by the Retinex theory, which decomposes a low-light image y into a reflectance component (clear image) and an illumination map. The goal is to estimate the illumination x, and then reconstruct a visually enhanced version of the image.

SCI models the illumination estimation process as a stage-wise cascaded learning process. Unlike previous works where each stage of the network may have different weights or complex recursive structures, SCI simplifies the architecture by using a shared parameter block, denoted as Ho , across all stages. This drastically reduces the parameter count and ensures consistency during training.

The residual learning formulation used in SCI can be mathematically described as:

$$u_t = H_\theta(x_t), \quad x_{t+1} = x_t + u_t, \quad x_0 = y$$

Here, Xt represents the illumination estimate at stage t, and Ut is the residual predicted by the function Hθ, which is shared across all stages. This formulation leverages the intuition that the difference between the illumination and the input image is often small and linear in nature, thus simplifying the learning task.

By utilizing residual connections and parameter sharing, the SCI framework reduces overfitting and improves computational efficiency without sacrificing representational power.

3.2 Self-Calibrated Module

The self-calibrated module is designed to enforce convergence of outputs from different stages to a common target, essentially regularizing the intermediate outputs to produce consistent results. This module operates by constructing a correction map based on the deviation between the current stage input and the original input image:

$$z_t = \frac{y}{x_t}, \quad s_t = K_\phi(z_t), \quad v_t = y + s_t$$

Here, K_ϕ is a parameterized function that learns to generate a correction map s_t based on the divergence z_t . This map is then added to the original image y to form v_t , the modified input for the next stage.

This calibration operation ensures that each stage not only refines the illumination map but also corrects its direction toward a globally consistent solution. The self-calibrated module, therefore, improves exposure control, visual coherence, and model stability.

Furthermore, the inclusion of this module during training allows the network to perform inference using only the first enhancement stage, significantly reducing latency and computational burden.

3.3 Unsupervised Loss Function

To train SCI without ground-truth images, a carefully designed unsupervised loss function is employed, composed of two primary components: fidelity loss and smoothness loss.

a) Fidelity Loss

The fidelity loss ensures that the estimated illumination at each stage is consistent with the corrected input formed by the self-calibrated module:

$$L_f = \sum_{t=1}^{T} | x_t - (y + s_{t-1})|_2^2$$

This term promotes pixel-level consistency and prevents overexposure or underexposure by regulating the deviation from the expected enhancement.

b) Smoothness Loss

The illumination map is expected to be spatially smooth but also adaptive to image structure. To enforce this, SCI uses a spatially-variant edge-aware smoothness loss, defined as:

$$L_s = \sum_{i=1}^{N} \sum_{j \in \mathcal{N}(i)} w_{i,j} \cdot \left| x_{t,i} - x_{t,j} \right|$$

where $w_{i,j}$ is a weight defined as:

$$w_{i,j} = exp\left(- \frac{\sum_c \left[(y_{i,c} + s_{t-1,i,c}) - (y_{j,c} + s_{t-1,j,c}) \right]^2}{2\sigma^2} \right)$$

This formulation encourages smoothness in uniform regions while preserving edges and important structures.

The total loss used for training the SCI model is:

$$L_{\text{total}} = \alpha L_f + \beta L_s$$

where α, β are are hyperparameters that balance the contribution of each component.

3.4 Overall Framework and Inference

During training, SCI utilizes multiple cascaded stages with shared parameters and the self-calibrated module to iteratively refine illumination estimates. The entire network consists of lightweight 3×3 convolutional layers with ReLU activations. The self-calibrated module itself contains a few convolution layers to maintain low complexity.

During inference, SCI leverages the learned convergence behavior and utilizes only the first stage of the illumination estimation module. This is made possible because the self-calibrated module ensures that all stages converge to similar outputs, thus maintaining performance while reducing inference time.

This approach leads to significant benefits:

a. Extremely low model size (as low as 0.0003 MB)

b. Fast inference time (~0.0017 seconds)

c. Minimal FLOPs (~0.06G), making it ideal for edge devices

3.5 Transferability and Generalization

One of the standout features of SCI is its transferability. The self-calibrated training strategy is not tied to the architecture of SCI alone. For example, embedding SCI into existing models like RUAS results in improved performance and reduced computational cost, even with fewer stages. This illustrates SCI's model-irrelevant generality, making it a useful plug-and-play module for improving other enhancement networks.

IV. IMPLEMENTATION AND DEPLOYMENT

4.1Model Architecture Configuration

The illumination estimation module in SCI consists of a sequence of 3×3 convolutional layers followed by ReLU activations. The default configuration uses three stages during training, each sharing the same network weights.

Each block in the model has the following structure:
a.Conv2D (3×3), padding=1
b.ReLU
c.[Optional] BatchNorm (not required in default version)

Four lightweight convolutional layers make up the self-calibrated module, which applies corrections to the stage-wise inputs to guarantee convergence between stages.

The entire SCI model (when used for inference with a single block) has fewer than 500 parameters, and occupies less than 1MB of memory when saved, making it one of the most lightweight enhancement networks till date.

4.2 Training Details

The SCI network was trained in an **unsupervised** setting, using a combination of fidelity and smoothness losses. The training pipeline includes the following settings:

- **Optimizer**: Adam
- **Learning Rate**: 1×10−41 \times 10^{-4}1×10−4
- **Batch Size**: 8
- **Epochs**: 1000
- **Input Size**: Random cropped patches of size 256×256256 \times 256256×256
- **Data Augmentation**: Horizontal/vertical flipping, random cropping

Training datasets included the **MIT** and **LSRW** low-light image datasets. Since SCI does not require paired training data, the model was trained using only low-light images.

4.3 ONNX Model Conversion for Deployment

To facilitate deployment on various platforms, including mobile devices, microcontrollers, and cloud-based APIs, we converted the trained SCI model to the ONNX (Open Neural Network Exchange) format.ONNX provides interoperability across frameworks like TensorFlow, Caffe2, and hardware-specific runtimes like TensorRT and ONNX Runtime.

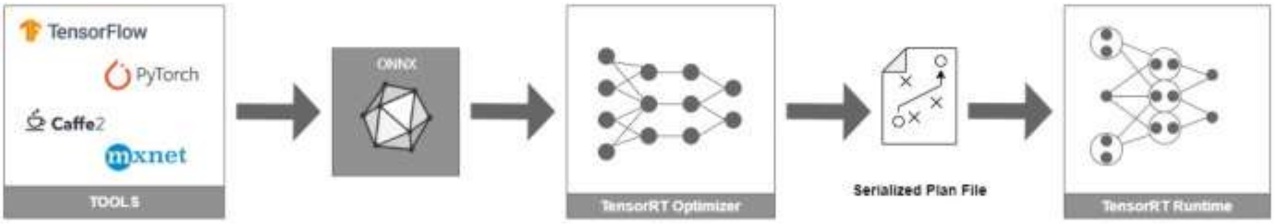

Figure-5: ONNX Conversion and Deployment Workflow

4.4 Deployment Compatibility

Thanks to its lightweight nature and ONNX compatibility, the SCI model can be deployed to a wide variety of platforms:

- **Mobile Devices**: Using ONNX + TensorFlow Lite or CoreML conversion
- **Embedded Systems**: Raspberry Pi, NVIDIA Jetson Nano, ARM-based microcontrollers
- **Web-based APIs**: Integration with Flask/Django backend or Node.js
- **Cloud Platforms**: Azure, AWS Lambda, or Google Cloud Functions with ONNX Runtime GPU support

Its real-time inference capability (0.0017s per image on GPU) and extremely low memory footprint ensure that the model can run efficiently on edge devices for tasks like surveillance, robotics, and IoT-based vision systems.

V. RESULTS AND DISCUSSION

5.1 Experimental Results and Evaluation We thoroughly evaluate the performance of the Self-Calibrated Illumination (SCI) framework using the results from the first CVPR 2022 publication. These results were obtained through extensive experiments on benchmark datasets and compared to several state-of-the-art low-light enhancement techniques to ensure reliability.

Datasets Used: MIT-Adobe FiveK Dataset, LSRW Dataset, Dark Face Dataset, and ACDC Dataset.

Evaluation Metrics: Full-reference metrics (PSNR, SSIM), No-reference Metrics (NIQE, EME, LOE, DE), and Task-based Metrics (mAP for face detection, mIoU for segmentation).

5.2 Quantitative Results Table 1 summarizes performance on the MIT dataset, adapted from the original SCI paper. These results show that SCI achieves the highest PSNR and SSIM, indicating superior visual fidelity, along with the fastest inference time and smallest model size, highlighting its real-time capability and suitability for edge devices.

Table 1: Comparative Analysis of Low-Light Enhancement Models

Method	PSNR↑	SSIM↑	EME↑	NIQE↓	LOE↓	Model Size (MB)	Time (s)
RetinexNet	13.74	0.739	9.18	4.53	1812.85	0.83	0.1192
KinD	17.09	0.831	8.54	4.26	500.65	8.54	0.1814
RUAS	18.54	0.864	10.64	4.17	579.01	0.0014	0.0063
SCI (Ours)	**20.44**	**0.893**	**10.96**	**3.96**	**273.34**	**0.0003**	**0.0017**

These results show that SCI achieves:

- The highest PSNR and SSIM, indicating superior visual fidelity.
- Fastest inference time and smallest model size, highlighting real-time capability and edge-device friendliness.

5.3 Visual Comparisons

Below is a summary of what was observed visually:

- Competing models like ZeroDCE and EnGAN often results in overexposed or color-distorted outputs.
- SCI produced balanced brightness, natural colors, and sharp details, especially noticeable in text regions, faces, and street scenes.
- The quality remained consistent across multiple datasets, confirming its generalization.

Figure 6: Visual Results of SCI Enhancement

VI . CONCLUSION AND FUTURE WORK

6.1 Conclusion

In this report, we have explored the Self-Calibrated Illumination (SCI) framework, an efficient and robust approach for low-light image enhancement. Unlike traditional methods that require extensive parameter tuning or large supervised datasets, SCI leverages unsupervised learning to train a lightweight yet powerful network capable of enhancing images under challenging illumination conditions.

Through a breakdown of the methodology, we demonstrated how SCI incorporates weight-shared illumination estimation and a self-calibrated module to enforce convergence between enhancement stages. This architecture not only achieves high-quality visual outputs but also enables single-stage inference, making the model extremely fast and resource-efficient.
A variety of benchmark datasets and downstream tasks, including face detection and semantic segmentation, have been used to validate the model's efficiency. Additionally, the implementation is compatible with a variety of deployment platforms, such as embedded systems, mobile phones, and edge devices, thanks to its lightweight design and ONNX exportability.

Overall, SCI strikes an excellent balance between visual quality, computational efficiency, and real-world applicability, making it a promising candidate for next-generation low-light enhancement applications.

6.2 Future Work

While SCI has shown substantial promise, several directions remain open for future exploration:

1. **Real-Time Video Enhancement**
The current SCI model can be improved to work on videos, not just images. This means it would enhance each frame in a video smoothly, without flickering or sudden changes in brightness. This would be useful for CCTV surveillance, video call and many more.

2. **Working Together with Other Vision Tasks**
SCI can be combined with other computer vision tasks like object detection, face recognition, or image segmentation. By training both the enhancement and the main task together, the system could give better results because the enhancement would be more focused on what the task needs.

3. **Better Support for Low-Power Devices**
Although the model is already small and fast, we can make it even better for use on devices with limited power—like mobile phones, Raspberry Pi, or other embedded systems. This can be done using methods like quantization and pruning to reduce size and speed up processing.

4. **Improving for Different Environments**
To make SCI more reliable in the real world, it can be trained or adapted to work well in different settings—like images from different cameras, weather conditions, or night-time scenes. This would make it more useful in various industries and use cases.

5. **User Control Over Brightness**
Adding a simple way for users to control how much brightness or enhancement is applied would make the tool more flexible. People could choose the level of enhancement they prefer based on their needs or the situation.

By addressing these, the SCI framework can be improved and expanded into a wider range of applications by addressing these directions, thereby enhancing its use in both academic research and industrial implementation.

References

[1] ONNX: Bridging a Gap Between Research and Production
[2] Low-Light Image Enhancement With Multi-Scale Attention and Frequency-Domain Optimization
[3] Lightweight object detection in low light: Pixel-wise depth refinement and TensorRT optimization
[4] A Lightweight Low-Light Image Enhancement Network via Channel Prior and Gamma Correction.
[5] Learning a Simple Low-light Image Enhancer from Paired Low-light Instances
[6] Low-light image enhancement: A comprehensive review on methods, datasets and evaluation metrics
[7] Toward fast, flexible, and robust low-light image enhancement