Low-Light Image Classification

Objective

- Vision-based applications often struggle in low-light conditions due to poor illumination, affecting tasks like object detection and recognition.

- Develop a self-supervised method that enhances visual quality and adapts models to low light environments without extensive labeled data.

Key Contributions

The contributions of the work are:

1. **New Low-Light Vision Model**: A lightweight model that enhances images for high-level tasks in low-light conditions.

1. **Self-Supervised Framework**: A self-supervised approach using enhancement curves and pseudo-labeling, requiring limited annotated data and compatible with multiple tasks.

Dataset - CODaN

- Common Objects Day and Night (CODaN) is an image classification dataset for zero-shot day-night domain adaptation / generalization. A large dataset specifically created for low-light adaptation tasks.

- The CODaN dataset consists of 15,500 224x224 colour images in 10 classes, with 1,550 images per class. Low Light Images = 2500, Normal Light Images = 13000. Test partition = 3000 (500 + 2500)

- The dataset is collected from the excellent COCO, ImageNet and ExDark datasets. All images are filtered and cropped such that they have the same dimensions and are completely mutually exclusive, i.e. do not contain objects of different classes, nor do belong objects to multiple classes.

Dataset Highlights: Includes diverse scenes, making it ideal for generalizing across different types of low-light environments.

Data Augmentation

1. transforms.RandomResizedCrop(224)

- This transformation randomly crops the image, then resizes it back to fixed size of (224, 224, channels).
- **Purpose**: By focusing on different parts of the image, it encourages the model to learn features across various portions, enhancing robustness against variations in scale and framing.

2. transforms.RandomHorizontalFlip()

• This randomly flips the image horizontally with a probability of 50%.

• **Purpose**: Horizontal flipping creates mirror images of the input, allowing the model to learn features in both left-right orientations. This is helpful in preventing overfitting to one specific direction, improving generalization.

3. transforms.RandomRotation(15),

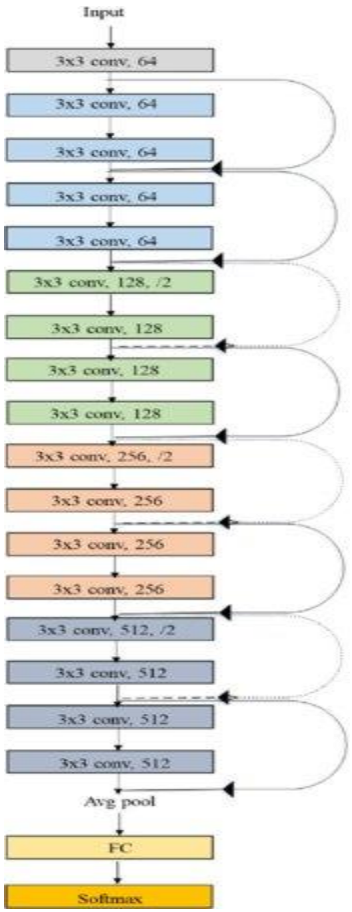
• Rotates the image randomly within a specified degree range, here between **-15 and +15 degrees**.

• **Purpose**: Rotation introduces slight variations in angle, making the model invariant to small rotations. This is especially useful in scenarios where objects may not always appear perfectly aligned, helping the model recognize features from multiple perspectives.

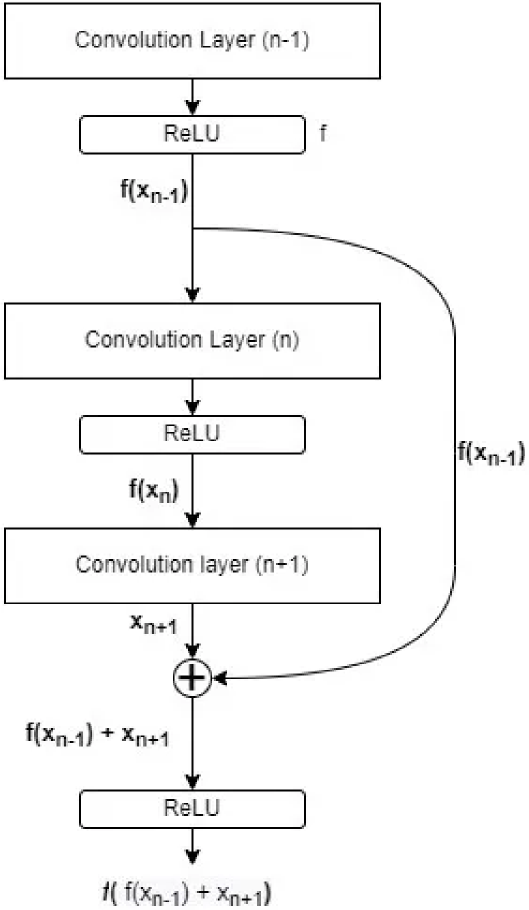
ResNet-18 - Model Architecture

ResNet-18, pretrained on ImageNet, is a widely-used convolutional neural network architecture known for its efficiency and performance in visual recognition tasks

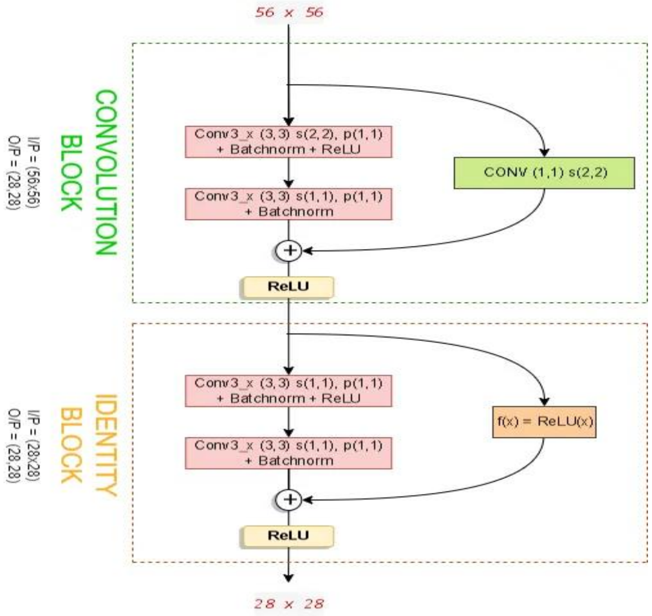
- **Residual Learning**: ResNet-18 uses skip connections to enable deeper networks without vanishing gradients.
- **Efficient Architecture**: With only 18 layers, it's lightweight and suitable for real-time applications.
- **ImageNet Pretraining**: Pretraining on ImageNet helps it learn diverse features, making it adaptable to various tasks.
- **Transfer Learning**: It performs well as a feature extractor, allowing quick adaptation to new domains with minimal data.
- **High Performance**: ResNet-18 achieves strong accuracy on many tasks while being computationally efficient.



Input
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 128, /2
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 256, /2
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 512, /2
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
Avg pool
FC
Softmax



Convolution Layer (n-1)
ReLU
f
$f(x_{n-1})$
Convolution Layer (n)
ReLU
$f(x_n)$
Convolution layer (n+1)
x_{n+1}
$f(x_{n-1})$
$f(x_{n-1}) + x_{n+1}$
ReLU
$f(\, f(x_{n-1}) + x_{n+1})$



56 x 56
CONVOLUTION BLOCK
I/P = (56x56)
O/P = (28,28)
Conv3_x (3,3) s(2,2), p(1,1) + Batchnorm + ReLU
Conv3_x (3,3) s(1,1), p(1,1) + Batchnorm
CONV (1,1) s(2,2)
ReLU
IDENTITY BLOCK
I/P = (28x28)
O/P = (28,28)
Conv3_x (3,3) s(1,1), p(1,1) + Batchnorm + ReLU
Conv3_x (3,3) s(1,1), p(1,1) + Batchnorm
f(x) = ReLU(x)
ReLU
28 x 28

Training Strategy

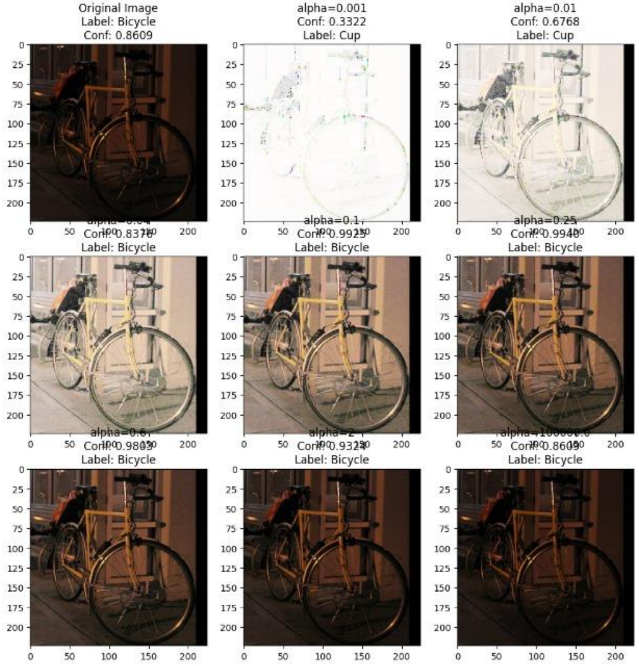
1. Finetune ResNet-18 (pretrained on imageNet dataset) on the Normal Light Dataset.
2. Formulate the base enhancement curve family as follows:

$$\mathcal{T} = \bigcup_{\alpha \in \mathcal{A}} \{f(x; \alpha)\},$$

$$f(x; \alpha) = \frac{(\alpha + 1)x}{x + \alpha}$$

$$\mathcal{A} = \{0.001, 0.01, 0.04, 0.1, 0.25, 0.6, 2, 10^5\}.$$

Reference: Wang, Wenjing, et al. "Unsupervised Illumination Adaptation for Low-Light Vision." *IEEE Transactions on Pattern Analysis & Machine Intelligence* 01 (2024): 1-15.



Base Enhancement

3. Generate pseudo labels for Low Light Images using base enhancement technique, fine tuned ResNet-18 and thresholding.

Number of Pseudo-labels generated = 548, with pseudo-label accuracy of 94.71% using confidence threshold t1 = 0.99, and voting threshold t2 = 2.

4. Find best alpha with which ResNet gives highest accuracy for pseudo-labels. Obtained alpha = 0.6.

4. Use normal light and Pseudo-labelled low light images again to fine-tune the model.

Results

Accuracy = 92.37%