# **Image Colorization and Sharpening Using CNN**

1. Problem Statement

In today’s digital age, visual data has become a cornerstone of communication, entertainment, and information dissemination. However, not all images are born in full color, and not all are captured with perfect clarity. Many valuable visual resources, such as historical photographs, medical scans, or surveillance footage, often exist only in grayscale or suffer from blurriness. The transformation of such images into vibrant, detailed visuals not only improves their aesthetic appeal but also enhances their utility in various applications.

The **problem addressed** in this project revolves around two key challenges in the field of image processing:

1. **Image Colorization** – The process of converting grayscale images into color images in a semantically meaningful way.
2. **Image Sharpening** – The enhancement of image details and edges to improve visual clarity and definition.

Both of these processes require a deep understanding of image content, structure, and semantics, which traditional rule-based or filter-based methods often fail to achieve with satisfactory results. With the rise of deep learning, and more specifically **Convolutional Neural Networks (CNNs)**, it is now possible to automate these tasks by learning from a large corpus of data, enabling more accurate and visually appealing results.

This project proposes a deep learning-based approach that:

* Utilizes an **encoder-decoder CNN** architecture to colorize grayscale images.
* Applies a **2D convolutional sharpening filter** to enhance the sharpness and details of the colorized images.

## 2. Introduction

The importance of high-quality, color-rich images cannot be overstated in areas ranging from digital media and photography to scientific research and historical documentation. Color plays a critical role in human perception and interpretation of visual information, while sharpness ensures that the image content is crisp and intelligible.

### Image Colorization

Colorizing images is a complex process that involves assigning plausible and aesthetically appropriate color values to each pixel in a grayscale image. The challenge lies in the ambiguity—there is often no single “correct” way to colorize a grayscale image, especially in the absence of contextual clues. For example, a dress in a grayscale image could be red, blue, or green. Therefore, an effective model must learn semantic context to make reasonable color predictions.

### Image Sharpening

Image sharpening, on the other hand, aims to improve the visual quality of images by highlighting edges and textures. It increases the contrast between neighboring pixels, especially along object boundaries, thereby making the image appear clearer and more defined. This process is essential for applications such as document scanning, medical imaging, and computer vision.

### CNN-Based Approach

CNNs have shown remarkable success in various vision tasks due to their ability to automatically learn hierarchical representations of data. By leveraging these capabilities, we can develop a model that:

* Encodes the grayscale image into a compact, feature-rich representation.
* Decodes this representation into a color image.
* Enhances the sharpness of the output using a convolutional filter designed to emphasize high-frequency details.

## 3. Literature Survey

Numerous techniques have been proposed over the years to tackle the tasks of colorization and sharpening, each with varying levels of success and complexity.

### 3.1 Image Colorization Techniques

1. Manual Colorization  
    Historically, artists have manually colorized black-and-white images by applying color to each region. While this can yield stunning results, it is labor-intensive and subjective.
2. Heuristic and Rule-Based Approaches  
    Early automated methods used segmentation, texture recognition, and edge detection to apply predefined colors to specific image regions. However, these methods are inflexible and struggle with complex or diverse images.
3. Machine Learning Approaches  
   * *Cheng et al. (2015)* introduced a deep CNN model that directly learns color distributions from training images.
   * *Zhang et al. (2016)* proposed a model that predicts ab channels in the Lab color space from grayscale input using a classification loss.
   * *Iizuka et al. (2016)* presented a fully automatic colorization network that incorporates global and local image features.

These models demonstrated that CNNs could capture semantic information to colorize images more realistically.

### 3.2 Image Sharpening Techniques

1. Classical Filters  
    Traditional sharpening techniques use filters such as the Laplacian filter, Unsharp Mask, or Sobel operator. While effective for basic edge enhancement, they are prone to amplifying noise and lack adaptability.
2. CNN-Based Enhancements  
    More recent approaches use deep learning to perform image enhancement tasks, often using CNN layers to learn edge features and reconstruct high-frequency components. These methods tend to be more adaptive and produce higher-quality results.

## 4. Methodology

The proposed system is composed of two sequential modules:

1. Colorization using CNN-based Encoder-Decoder Architecture
2. Sharpening using 2D Convolutional Filter

### 4.1. Colorization Module

#### a. Input Preparation:

* Input grayscale images are normalized and resized to a fixed size (e.g., 256x256).
* The grayscale image is typically converted into the Lab color space where only the L channel (lightness) is used as input.

#### b. Encoder:

* A series of convolutional layers extract low-level and high-level features from the grayscale input.
* Layers include convolution + ReLU activation + max-pooling to downsample and abstract the image.

#### c. Decoder:

* The decoder resamples the encoded features using transposed convolutions (deconvolutions).
* It reconstructs the a and b color channels corresponding to chrominance.
* The model learns to map grayscale features to plausible color values.

#### d. Output Reconstruction:

* The original L channel is combined with the predicted a and b channels.
* The Lab image is then converted back to RGB format for visualization.

### 4.2. Sharpening Module

After colorization, the image is passed through a sharpening filter to enhance its clarity.

The filter used is a 3x3 kernel:

[ [ 0, -1, 0],[-1, 5, -1],[ 0, -1, 0] ]

* This filter works by emphasizing the central pixel while subtracting the surrounding pixels.
* When convolved with the image, it increases the contrast along edges, making the image appear sharper.

#### Process:

* Apply the kernel to each channel of the colorized image using a 2D convolution operation.
* The result is a sharper, more detailed image suitable for further analysis or display.

## 5. Results

The model was trained using a dataset composed of thousands of colored images (from CIFAR-10 and ImageNet), converted to grayscale for training. After training, it was tested on unseen grayscale images to evaluate colorization quality and sharpening effectiveness.

### Colorization Results:

* The encoder-decoder network successfully added realistic and coherent colors to grayscale images.
* Objects like sky, vegetation, and skin were colorized semantically, demonstrating the network's understanding of context.

### Sharpening Results:

* The sharpening filter enhanced the edges and fine details without significant artifact introduction.
* The output images had visibly better clarity and depth.

### Performance Metrics:

* PSNR (Peak Signal-to-Noise Ratio): Higher after sharpening, indicating better fidelity.
* SSIM (Structural Similarity Index): Improved, showing that the structural information was preserved.
* Visual Comparison: Side-by-side comparisons of input grayscale, colorized, and sharpened images showed progressive improvements in realism and clarity.

## 6. Conclusion and Future Scope

### Conclusion:

This project demonstrates the powerful capabilities of CNNs in addressing fundamental image processing challenges. By combining a deep learning-based colorization model with a convolutional sharpening filter, we achieved a comprehensive solution for enhancing grayscale images.

* The encoder-decoder CNN effectively captured the semantic content of images to generate plausible color representations.
* The convolutional sharpening filter provided a lightweight yet effective means of improving image sharpness and detail.
* The end-to-end pipeline produced results that were both aesthetically appealing and structurally accurate.

### Future Scope:

The current system lays the foundation for further enhancements:

* Use of Attention Mechanisms: Introduce attention layers in the encoder-decoder to focus on key image regions during colorization.
* GAN Integration: Employ Generative Adversarial Networks to produce more vibrant and diverse color outputs.
* Edge-aware Sharpening: Implement edge-aware deep learning models to enhance images without noise amplification.
* Real-Time Application: Optimize the model for real-time image processing on mobile or embedded devices.
* Video Colorization and Sharpening: Extend the method to sequential frames while maintaining temporal consistency.