IMAGE COLORING

Presented by:
PC23 Abhay Nayak
PC41 Rohit Saini
PC33 Chandan Sah
PC31 Aliasgar Thanawala

INTRODUCTION

The colorization problem involves adding color to black and white images, making them visually appealing and realistic. It's a complex task because it requires understanding the content of the image and selecting appropriate colors. Image colorization has numerous applications, including historical photo restoration, entertainment, and improving visual comprehension in fields like medical imaging.

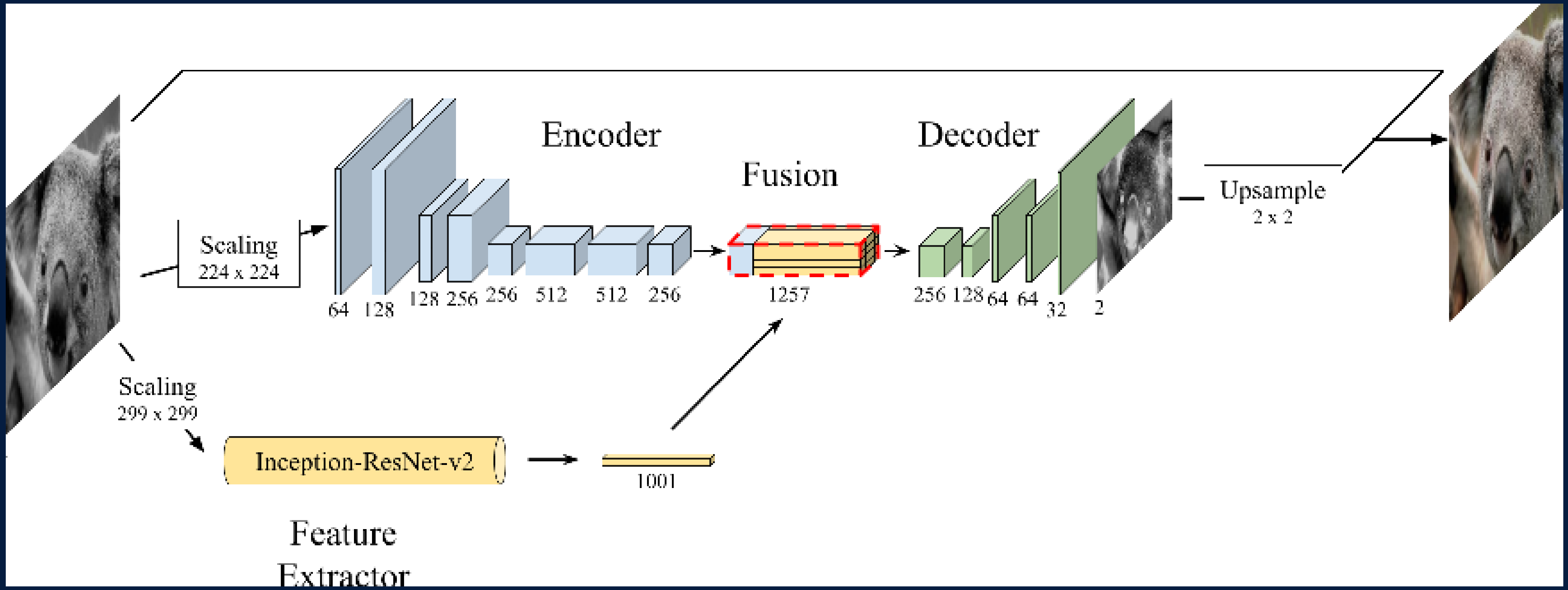
RGB vs Lab

RGB Color Space: RGB stands for Red, Green, Blue, and it's a common color model where colors are represented using combinations of these primary colors. In this color space, each pixel in an image is represented by three values (R, G, B) that specify the intensity of each color channel.

Lab Color Space: The Lab color space is designed to be perceptually uniform, meaning that the perceptual difference between colors corresponds to the Euclidean distance between points in this color space. It consists of the following components:

L*: Lightness (brightness) channel.
a*: Chrominance channel (from green to red).
b*: Chrominance channel (from blue to yellow).
Using Lab color space is often preferred for image colorization because it separates the color and brightness information. This makes it easier for models to predict color while preserving the image's structure and detail.

Architecture: Inception-ResNetv2 based CNN



Inception ResNetV2 is a deep convolutional neural network architecture that combines elements from both the Inception architecture and the ResNet architecture. It was designed to achieve high performance in image classification and other computer vision tasks while addressing issues like vanishing gradients and enabling the training of very deep networks.

Inception-ResNetv2:

- Inception-ResNetv2 is a pre-trained deep convolutional neural network architecture known for its effectiveness in image classification.
- In the context of image colorization, it can be used to leverage the high-level feature representations learned during pre-training.
- These features are extracted from the black and white image and provide valuable information for the colorization task.

Encoder:

- The encoder is responsible for extracting relevant features from the input black and white image.
- It typically uses convolutional layers to down sample the image and capture its essential information.
- The encoding process reduces the spatial dimensions while increasing the depth of the features.

Fusion Layer:

- The fusion layer is where the information from the Inception-ResNetv2 network is combined with the feature maps obtained from the encoder.
- This fusion process is critical because it merges the high-level semantic information from Inception-ResNetv2 with the low-level spatial features extracted by the encoder.

Decoder:

- The decoder is responsible for generating the color channels for the image.
- It takes the fused information and uses it to predict the Lab color values for each pixel in the image.
- The L channel (lightness) comes from the input black and white image, while the a* and b* channels are predicted by the decoder.

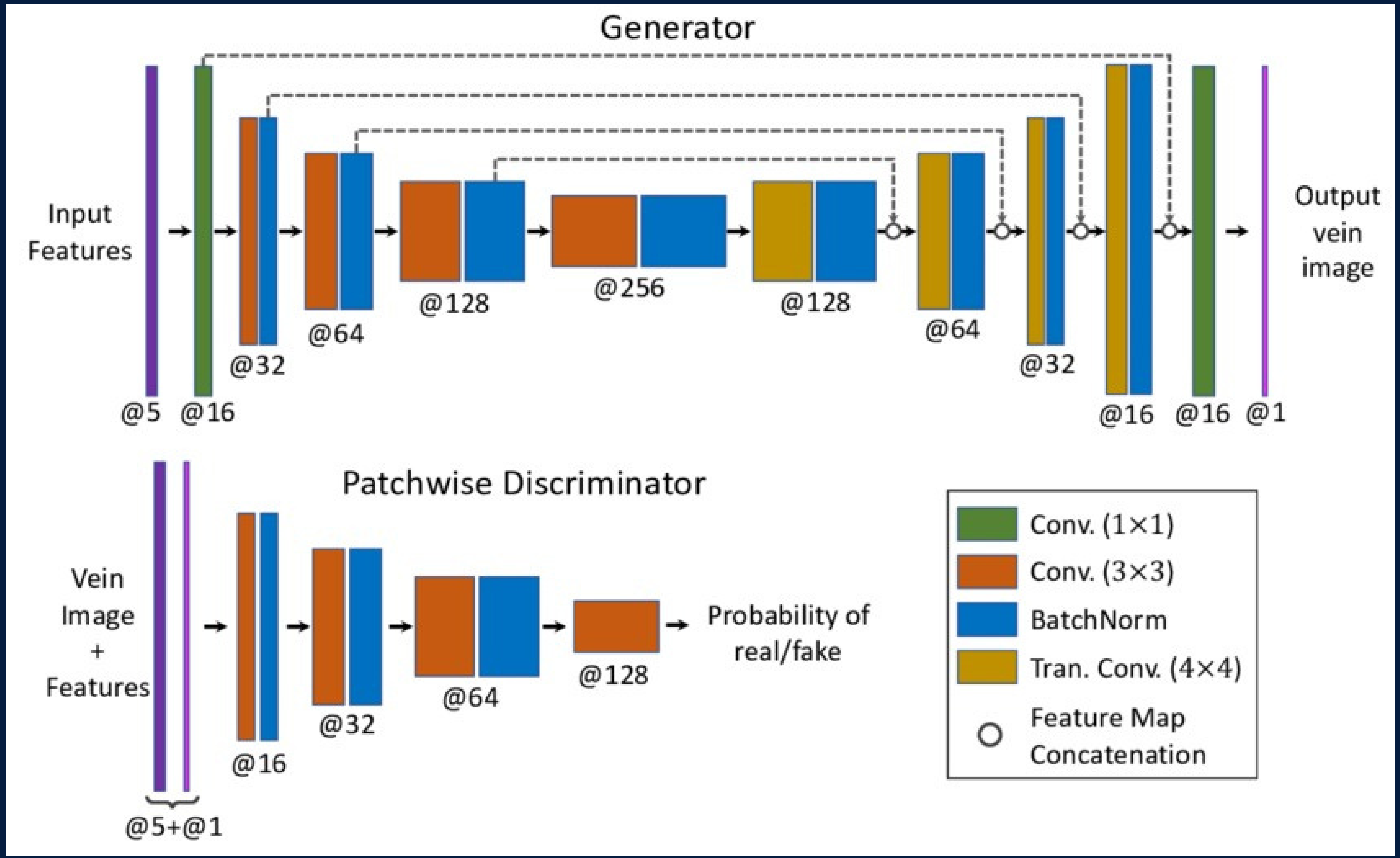
Pix2pix GAN

Pix2Pix is a type of Generative Adversarial Network (GAN) architecture used for various image-to-image translation tasks, which can include image coloring. The Pix2Pix GAN is particularly well-suited for tasks that involve converting one image into another while preserving the underlying structure or content. In the context of image coloring, Pix2Pix can be used to take a black and white image as input and produce a colorized version of that image as output.

In the Pix2Pix framework, there are two main components: a generator and a discriminator.

Generator: The generator takes a black and white image as input and aims to transform it into a colored version that is as realistic as possible. It does this through a series of convolutional layers and other neural network operations. The generator tries to map the input image to the output image.

Discriminator: The discriminator's role is to distinguish between real color images (from the dataset) and fake colorized images produced by the generator. It provides feedback to the generator, helping it to improve its colorization performance over time.



Loss Functions: Pix2Pix uses multiple loss functions to guide the training process:

- Adversarial Loss: This loss encourages the generator to produce colorized images that the discriminator can't distinguish from real ones.
- Perceptual Loss: To ensure the colorization is semantically meaningful, a perceptual loss, often based on pre-trained networks like VGG, is used to measure the difference between the colorized image and the ground truth color image.
- L1 Loss: This loss encourages the generator to produce images that are close to the ground truth images in terms of pixel-wise differences.

Challenges

Complexity of Color Perception

Human color perception is a highly complex process, and it can be challenging for deep learning models to accurately replicate the nuances of color in different scenes and lighting conditions.

Lack of Ground Truth Data

Creating a large and diverse dataset of black and white images with corresponding color ground truth can be difficult and expensive.

Semantic Understanding

Understanding the semantics of an image, such as identifying objects and their contextual colors, is a complex task.

Computational Resourcesl

Training deep neural networks for image colorization requires significant computational resources..

Memory and Speed Constraints

Applying deep learning models for real-time colorization on low-resource devices, such as smartphones or embedded systems, can be challenging .

LIMITATIONS

Color Bleeding

Colors from one object spreads into adjacent areas

Over-reliance on Context

It may struggle to accurately colorize objects in isolation

Gray or Dull Tones

Lack vibrancy and realism in the produced output

Limited Creativity

They tend to stick to common color associations

B&W Image Quality

The quality of input can significantly impact the results

Object-Specific Limitations

Transparent or reflective surfaces, can be challenging

References

- https://towardsdatascience.com/colorizing-black-white-images-with-u-net-and-conditional-gan-a-tutorial-81b2df111cd8
- https://www.irjet.net/archives/V8/i10/IRJET-V8I10231.pdf
- https://www.geeksforgeeks.org/black-and-white-image-colorization-with-opencv-and-deep-learning/
- https://medium.datadriveninvestor.com/coloring-black-white-images-using-deep-learning-984e6f4ddf14