

# **Neural Style Transfer**

Aryan Singh(20AR10004)

# What is Neural Style transfer?

**Neural style transfer** is in its most basic form is about transferring style from a style image onto a content image using neural nets.

In its more complex form you can:

- Transfer style to videos
- Choose whether to keep the color from the content image or take it from the style image
- Use segmentation masks to specify objects which should be styled and many cool things



Content Image

Style Image

Generated Image

## How do we get here what is the math behind this idea?

To answer these question we need to take a step back and focus on what does a convolution neural network actually learn?

### Brief of CNN

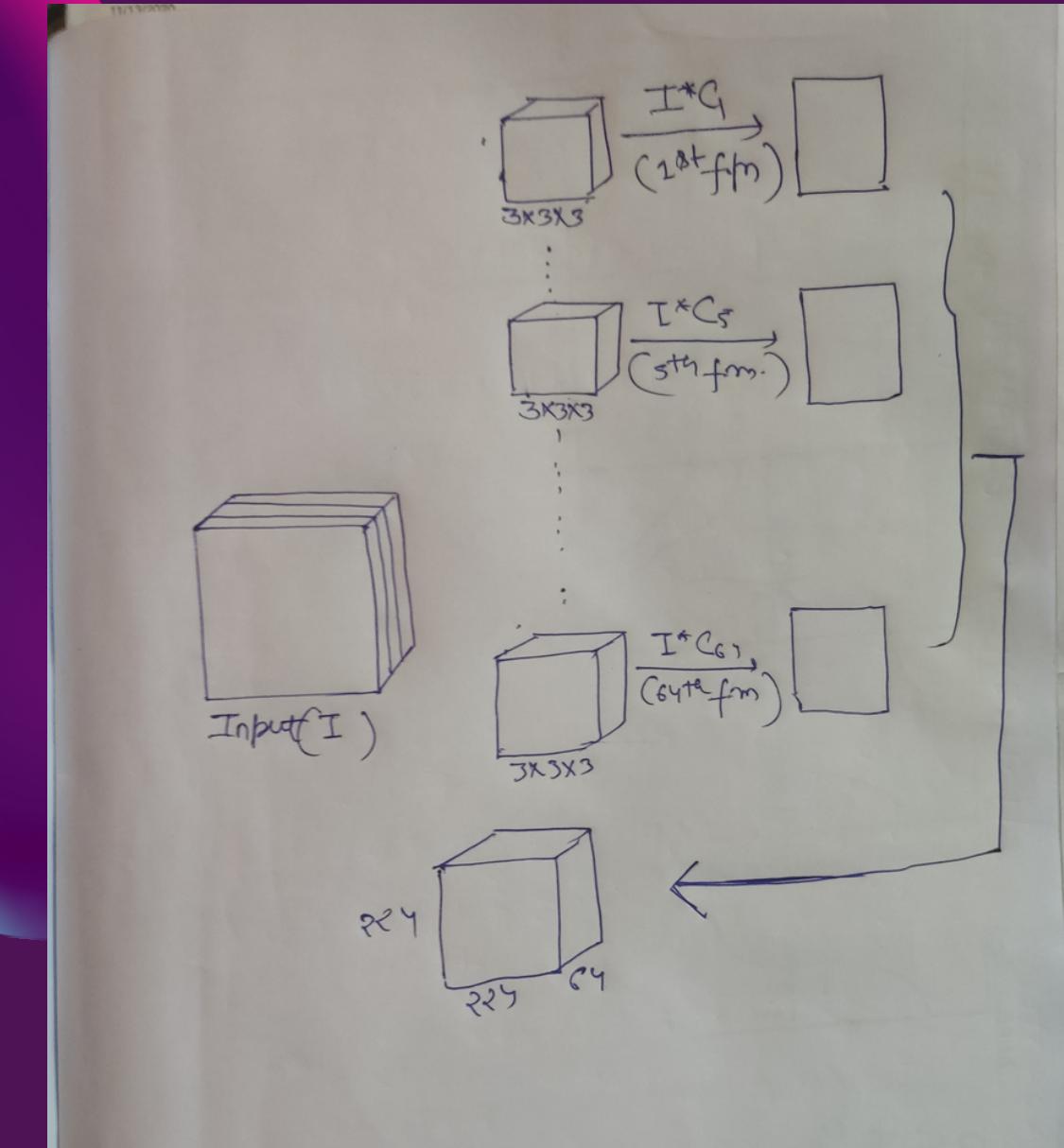
CNNs were originally created for the classification of images and have lately been used in a variety of other tasks like Image Segmentation, Neural Style, and other computer vision and NLP tasks as well

# Information we get from CNNs?

First CNN layer of VGG architecture which trains 64 feature maps and produces an image representation of 224x224x64

While training this 64 feature maps they may have learnt to detect simple patterns, such that some neural units activate when they see a specific object

CNNs does not learn what an image is ,but it learns what it represents and what it is showing.



## Use of Image Representations for Style Transfer

CNN's nature of encoding representations itself is the key to style transfer it is used to calculate loss between the generated image with respect to content and style image.

Training the model over ten thousands of images per class the model is able to generate similar feature representation for many different images given they belong to same class or have similar content or style.

Use the difference in value of feature representation of generated image w.r.t content and style image to guide the iterations through which we produce the generated image.

# How do we make sure that generated image only inherits a similar style representation and not the entire style image itself?

This is solved by dividing the loss function into two parts, one is the Content loss and the other is the Style loss

$$L_{total}(S, C, G) = \alpha L_{content}(C, G) + \beta L_{style}(S, G)$$

As we can see, the overall loss is divided into content and style loss.

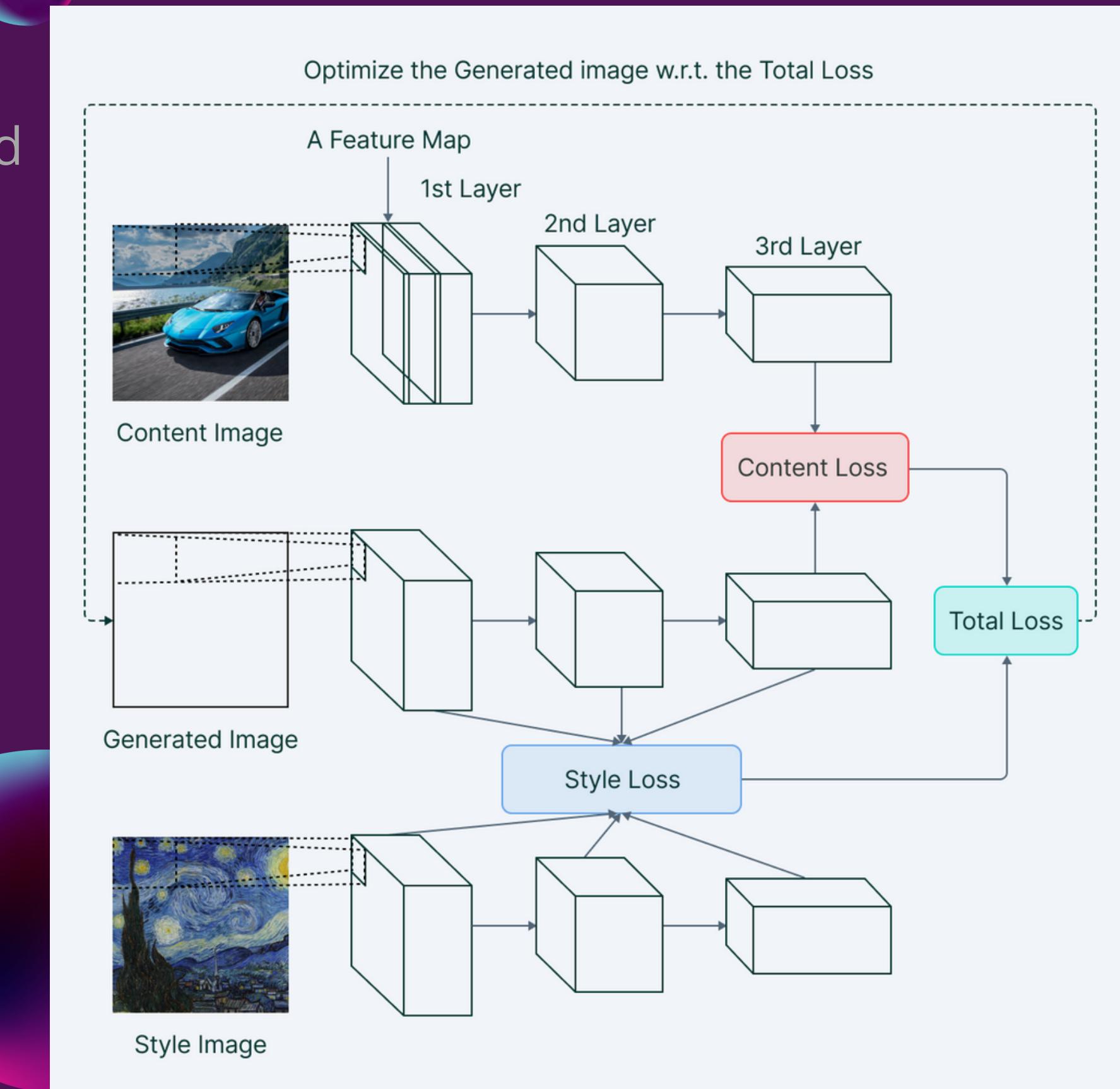
Alpha and Beta hyperparameters can be thought of simply as knobs to control how much of the content/style we want to inherit in the generated image

## Content Loss

It helps to establish similarities between the content image and the generated image.

# NST Model Architecture

The architecture of the NST can be designed in such a way that it can range from applying a single style in an image to allowing mix and match of multiple styles.



# THANKYOU