“ART, STYLE AND MACHINES”

By

Max Kaufmann

Introduction

The unquantifiability of art intrigues me. What exactly makes some brush strokes better than others? What makes an image pleasing? Why do some combinations of colour appear aesthetic to me, but not to others?

These questions are tough to answer, and any progress on them will come from someone smarter than I.

What I do find compelling (and perhaps a little insightful) is the separation between the content of an artwork, and the style it is painted in – asking Vincent van Gogh to paint a field will lead to very different results than giving Salvador Dalí the same task.

In recent times, machines have proven to be better than expected at the unquantifiable. Perhaps, by leveraging their own internal representations of the style and content of the art they view, we can gain some understanding of the nature of art.

At the very least, they make some pretty pictures.

A selection of pretty pictures

Style transfer takes the content of one image, and the 'style' of another, and produces an image combining the two. My content images are a selection of pictures sampled from my own life, and my style images are taken from a variety of artists (some of which you might be aware). I also speak about each painting and highlight an aspect of the art which interests me.

1. A Starry Max

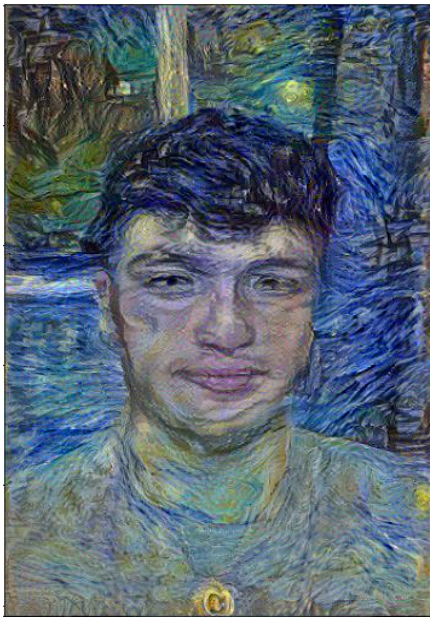
Input Content:

A photo of me

Input Style:

'The Starry Night'
By Vincent van Gogh

Output image:



I sometimes find the stories of the artists more intriguing than the art itself, and Vincent van Gogh's life is indeed an intriguing (albeit sad) story.

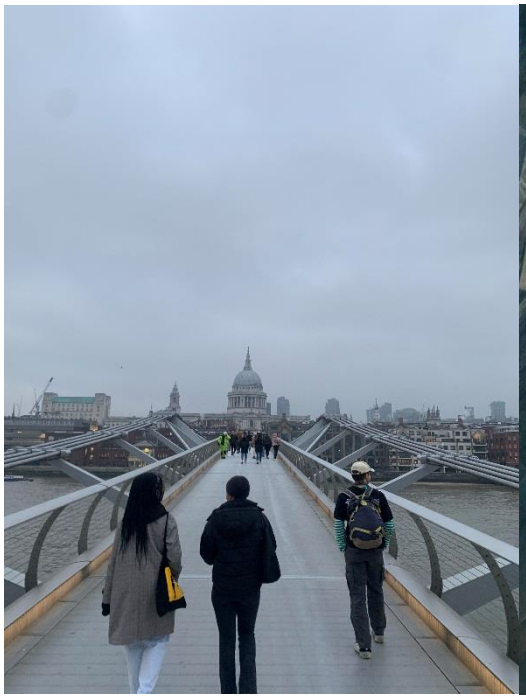
Born in 1853 in Zundent, Netherlands the most famous post-impressionist artist received no fame or accolades during his short life. In his mind, 'A starry night' (arguably his most well-known work) was a failure.

He was a tortured soul, leading a hard and thankless life. This painting was completed whilst he voluntarily stayed in the Saint Paul de Mausole lunatic asylum in Saint-Remy, France. He would shoot himself a few months after completing this work, the cypress trees (often associated with cemeteries) perhaps reflecting his dark thoughts in an otherwise beautiful scene.

A deeply religious man, van Gogh would spend time as a missionary in Belgium. The only light in this painting comes from the heavens – possibly a hint of what he focused on during this dark period of his life.

2. London, but its grey

Input Content:



My first trip to London

Input Style:

"Seated Nude"
By Pablo Picasso

Output:

I find it captivating how art changes to match the times and the people – art movements are triggered both by the changing world around them, and by pioneers who discover new ways to create. In the case of Cubism, the changing world was a rapidly modernizing 1910s Paris, and the pioneer was Pablo Picasso.

Picasso believed art should not try to imitate nature – it did not matter *what* the artist saw, but instead *how* they perceived it. Cubism was therefore a challenge to the conventional, realistic forms of his time. He comments on this breaking of norms in a "Seated Nude" - the pose and lighting are that of traditional portraiture, while the mechanization and stripping down of his subject's form is not.

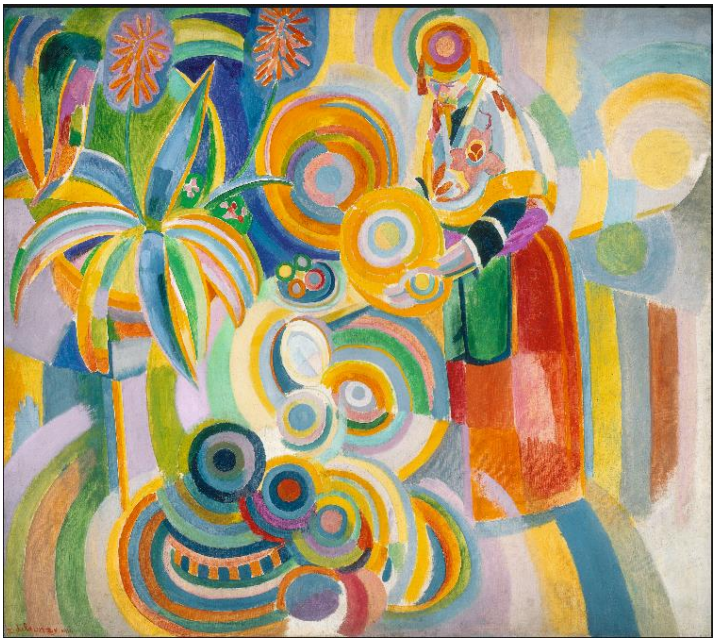
Picasso liked painting an object from many different viewpoints at once, and especially in his earlier experimentation (dubbed 'Analytic Cubism'), his subjects would be made up of multi-faceted geometric shapes. The drab greys of this painting are interesting and convey an almost oppressive tone; the critics of his time would often classify Picasso as portraying the world through an analytical, scientific lens. This painting proves the type of expressiveness to which Picasso's style could lend itself.

3. Tall English College

Content Input:

Matriculation photos at Jesus College, Cambridge

Style Input:



Tall Portuguese Woman
By Robert Delaunay

Output image:

Abstraction is a powerful tool, and this fact was not lost on artists at the turn of the last century – many art movements of this time sought to move away from depicting scenes of the physical world, and create more 'pure' art which would capture some concept or idea without the shackles of reality. Orphism, spearheaded by Robert Delaunay, aimed to capture and give priority to light and color.

Although painted without a subject, 'Tall Portuguese Woman' still has some grounding in reality when compared to other Orphist paintings - most of which are colorful shapes on a canvas and bear no resemblance to the world we perceive. The reality in this case is the one of Delaunay, who found himself stuck in Portugal on the outbreak of World War I. What had been planned as a family summer retreat turned into a 3-year hideout.

It is perhaps a shame that Ophism was not more often applied to 'the real world', as this painting demonstrates the style's ability to create a compelling scene. The vibrant colours and simple shapes give the painting an upbeat, yet slightly surreal energy. Perhaps it captures what one might feel on a hot summer's day in Portugal, sheltering away from the world's settling of differences.

Boring technical details

Neural style transfer takes a base input image, a style image, and a content image and then tries to morph the input image into a combination of the content and style of both.

This is achieved by leveraging the intermediate layer activations of images within a pretrained network. Neural networks build up more and more complex and abstract (although often opaque) representations of relevant input features as you go deeper in a network. These can then be used to try and create loss functions describing the difference in both style and content between two images. Starting with a white noise image, gradient descent is used to minimize this loss to both the style and the content images, hopefully ending up with the desired combination.

In this case, we use the pretrained VGG19 image classifier network to get our activations (as suggested in the original paper) and define our loss functions as follows:

Given an input image $\vec{x}$ and a target image $\vec{p}$ we, look at the activations of feeding each image through our network. Say the L^{th} layer of our network is a N^L by M^L matrix, we can write F^L and P^L to denote the L^{th} layer activation of the input and target image respectively. We then define:

1) Content loss:

$$L_{loss}(\vec{x}, \vec{p}, l) = \frac{1}{2}\sum(F_{ij}^{il} - P_{ij}^{il})^2$$

In other words, content loss at layer l is simply the 'classic' squared difference between the two layers. Total content loss is a weighted sum across layers.

2) Style loss:

In our representation, each row of a matrix describes a feature map (a single filter being applied across the whole image). From this, we define the 'Gram matrix' which is made up of the inner product of these row vectors:

$$G_{ij}^L = \sum_k F_{ik}^L - F_{jk}^L$$

Our content loss for a layer is then the squared difference between the gram matrices of the input image and the target images at that layer. Total style loss is then a weighted sum across layers.

Our total loss is a weighted sum of the total content and total style losses between our input image and the content and style images. Full details can be found in Leon A. Gatys' paper "A Neural Algorithm of Artistic Style".

Git repository for my implementation: https://github.com/max-kaufmann/Style-Transfer