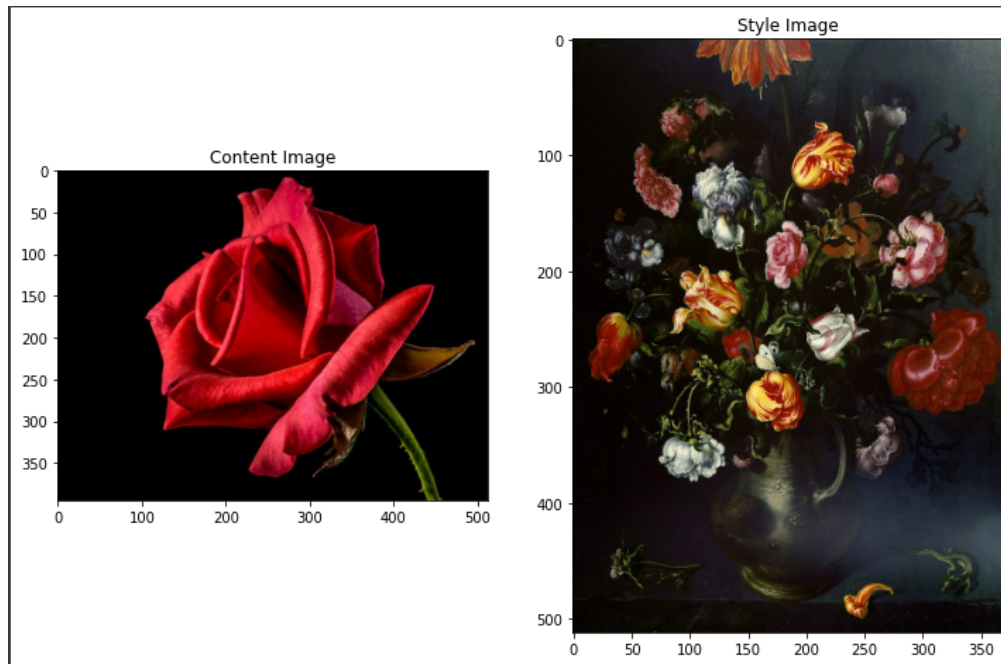


Style transfer augmentation

Style transfer is een deep learning techniek om de stijl van afbeelding 1 te mengen in een andere afbeelding maar de inhoud van afbeelding 2 te behouden.

Afbeelding 2 behoudt de context door zijn randen te behouden, afbeelding 1 geeft het kleur en de stijl. Dit gebeurt door een getraind convolutional netwerk dat al bestaat.



We gebruiken verschillende lagen om de inhoud en de stijl van de afbeelding te verkrijgen. In de input laag zien we voornamelijk low-level eigenschappen zoals randen en textuur. Naarmate we hoger gaan in lagen zien we meer high-level eigenschappen zoals bijvoorbeeld ogen.

Voor een netwerk om image classificatie te doen moet het de afbeelding kunnen begrijpen, daarom moet het de afbeelding opsplitsen in pixels. Het krijgt de input image dus binnen als pixels en maakt een interne representatie dat de pixels omzet naar een manier om de eigenschappen die aanwezig zijn in de afbeelding te begrijpen. Tussen de input image en het uiteindelijk resultaat die het model een soort zeer complexe eigenschap extractie.

De inhoud van een afbeelding wordt weergegeven door een feature map. De stijl daarentegen kan worden beschreven als de middelen en correlaties tussen verschillende feature maps.

Een gram matrix bevat deze info door op elke locatie het resultaat te nemen met zichzelf en daar het gemiddelde mee te nemen over het resultaat over alle locaties.

```
def gram_matrix(input_tensor):  
    result = tf.linalg.einsum('bijc,bijd->bcd', input_tensor, input_tensor)  
    input_shape = tf.shape(input_tensor)  
    num_locations = tf.cast(input_shape[1]*input_shape[2], tf.float32)  
    return result/(num_locations)
```

Vervolgens bouwen we een model dat stijl en inhoud tensors (een generalisatie van vectoren en matrices, makkelijkst vergeleken met een multidimensionale array) teruggeeft.

Wanneer we dit model op een afbeelding gebruiken krijgen we een gram matrix van de stijl lagen en de inhoud van de inhoud lagen.

Nu kunnen we het stijl transfer algoritme toepassen, dit doen we door de MSE te berekenen voor de afbeelding relatief naar elk doel.

Het resultaat hiervan is als volgt.



Dit is een veelbelovend resultaat, alhoewel de stijl hierin niet direct terug te vinden is (maar daar keren we dadelijk op terug) zien we duidelijk een roos en lijkt wel alsof deze geschilderd is. Dit zou dus al een mogelijke manier zijn om meer afbeeldingen te maken. Maar laten we eerst het model nog wat optimaliseren om de stijl er duidelijker in te verwerken.

Na optimalisatie met behulp van variation loss verkrijgen we dit resultaat.



We zien dat de stijl hier veel duidelijker in te herkennen valt en dat de roos nog duidelijk zichtbaar is. De mogelijkheid om hiermee extra trainingsdata te maken is dus zeker niet uitgesloten.