

Métodos em Aprendizado Não Supervisionado de Máquina - Trabalho I

Daniel Lopes Toso

Abril 2024

1 Introdução

O presente trabalho tem como objetivo explorar técnicas para a extração de atributos quantitativos de imagens das pinturas do artista neerlandês Piet Mondrian, a fim de aplicar métodos para a clusterização para variáveis reais a esses quadros.

Os dados foram obtidos através do webscraping do site [WikiArt](#). Foram salvas as imagens, bem como o nome e data, se disponível, de criação das obras. Os códigos para o webscraping em [Python](#) e para a análise e clusterização em [R](#) estão disponíveis no [Github](#).

As imagens foram convertidas em arrays tridimensionais com o pacote "png" do R. Para uma imagem $N \times M$, os arrays consistem em 3 matrizes $N \times M$, uma para cada canal de cor - Vermelho, Azul ou Verde, respectivamente. Os números encontrados nas matrizes representam a intensidade de cada canal de cor para cada pixel, em uma escala de 0 a 1.

Para algumas das métricas, foi realizada uma conversão das pinturas do colorido para o preto e branco. Isso foi feito calculando a média aritmética de cada pixel sobre os 3 canais de cores, resultando em um array unidimensional.

2 Materiais e Métodos

2.1 Métricas

Fontes contribuidoras da inspiração para a adoção e aplicação das métricas são:

- Contraste - [Information Theory](#)
- Entropia - [Entropy of an Image](#)
- Tons - [Identify Redness of an Image](#)

2.1.1 Contraste

O contraste foi calculado da seguinte maneira:

- O array tridimensional da imagem é transformado em uma única matriz referente à versão preto e branca do quadro.
- Para cada pixel da matriz, é calculada a soma dos desvios ao quadrado entre o valor do pixel e todas as células em um quadrado 10×10 em seu entorno.
- O contraste é definido como a soma de todos os desvios calculados na matriz

A escolha desse procedimento para a métrica do contraste se baseia na ideia de que imagens com mais contraste possuiriam áreas onde os valores dos pixels mudariam bruscamente (como o contraste do preto com o branco, ou do claro pro escuro), o que aumentaria o desvio dos pontos para seus vizinhos.

Devido ao custo computacional de calcular os desvios para cada uma das células da matriz, o contraste foi estimado por amostragem aleatória simples de células na matriz. Para cada imagem, foram utilizados 500 pontos com coordenadas escolhidas por uma distribuição uniforme para as linhas e colunas. A variância do estimador depende do nível de complexidade da imagem, abaixo foram realizadas estimativas dessa variância para as imagens.

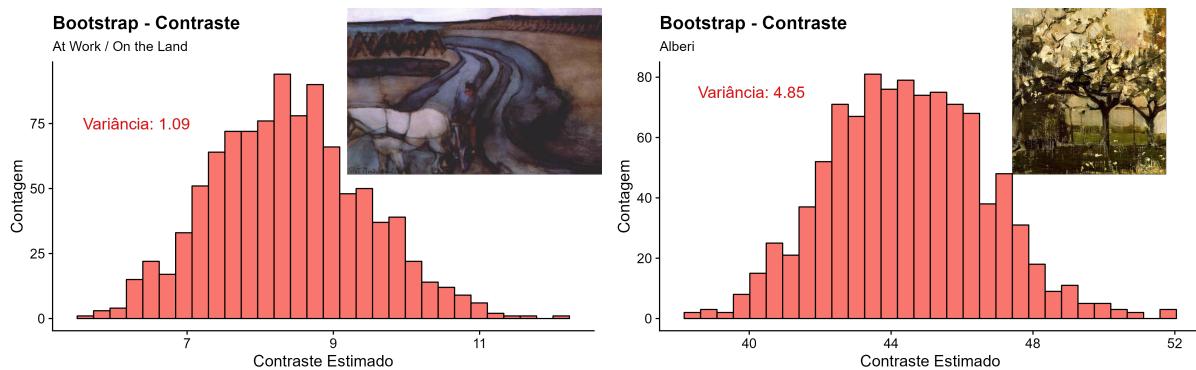


Figura 1: Comparação da variância do estimador do contraste para dois quadros com níveis diferentes de complexidade

As imagens com mais e menos contrastes foram:

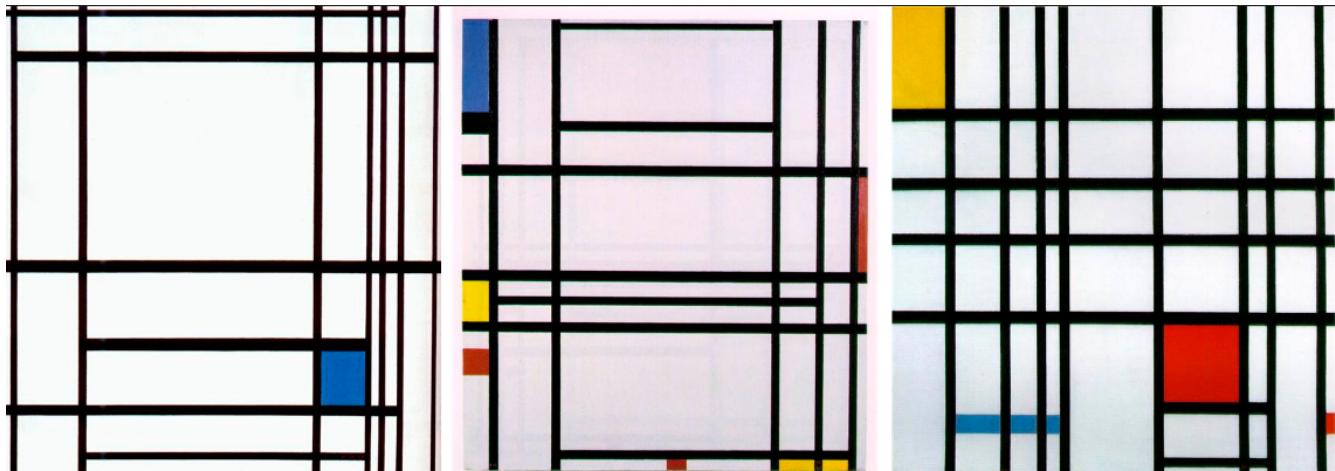


Figura 2: As 3 imagens com maior contraste: *Composition with Blue*, *Composition No.10* e *Composition with Red, Yellow and Blue*



Figura 3: As 3 imagens com menor contraste: *The Red Mill*, *View from the Dunes with Beach and Piers* e *Portrait of a Girl with Flowers*

2.1.2 Entropia

A entropia foi calculada como sendo a entropia de Shannon da matriz de valores da imagem em sua versão preto

e branca.

Os elementos a_{ij} da Matriz A foram discretizados em intervalos $C_i, i = \{1, 2, \dots, 256\}$ igualmente espaçados e a entropia de Shannon da matriz foi definida como:

$$H = - \sum_{i=1}^{256} \hat{P}_i \ln(\hat{P}_i)$$

Onde:

$$\hat{P}_i = \frac{1}{\#\{a_{ij} \in C_i\}}$$

A escolha dessa métrica foi baseada na intuição de criar uma variável proxy para a "desordem" da matriz. Imagens mais ordenadas devem possuir mais canais de cores iguais, o que leva a \hat{P}_i maiores e diminui a estatística H, enquanto imagens mais desordenadas possuem níveis de ativação dos pixels mais uniformemente distribuídos entre os possíveis valores. Abaixo, exemplos das entropias obtidas para certas imagens.

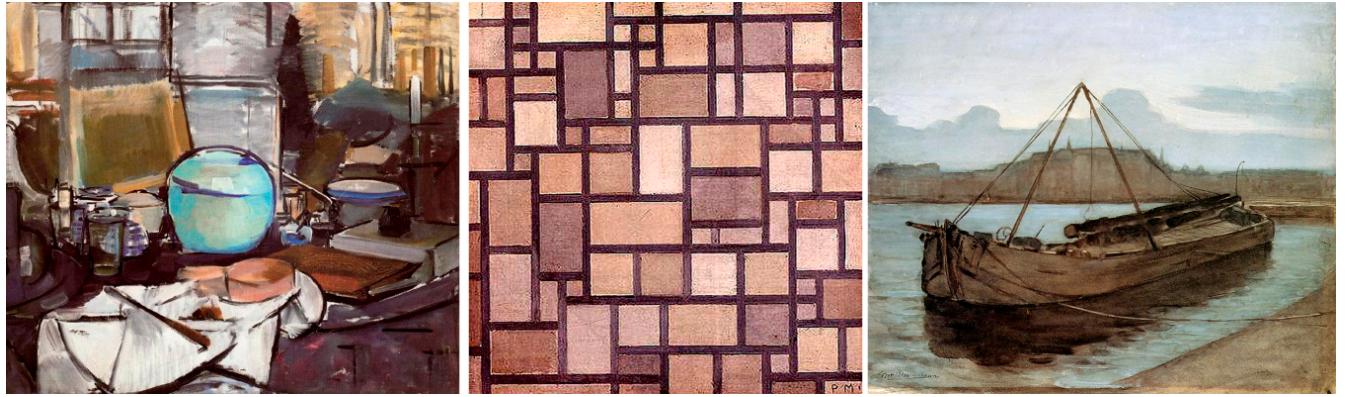


Figura 4: As 3 imagens com maior entropia: *Still Life with Gingerpot 1*, *Composition: Light Color Planes with Grey Contours* e *Evening at Weesperzijde Sun*

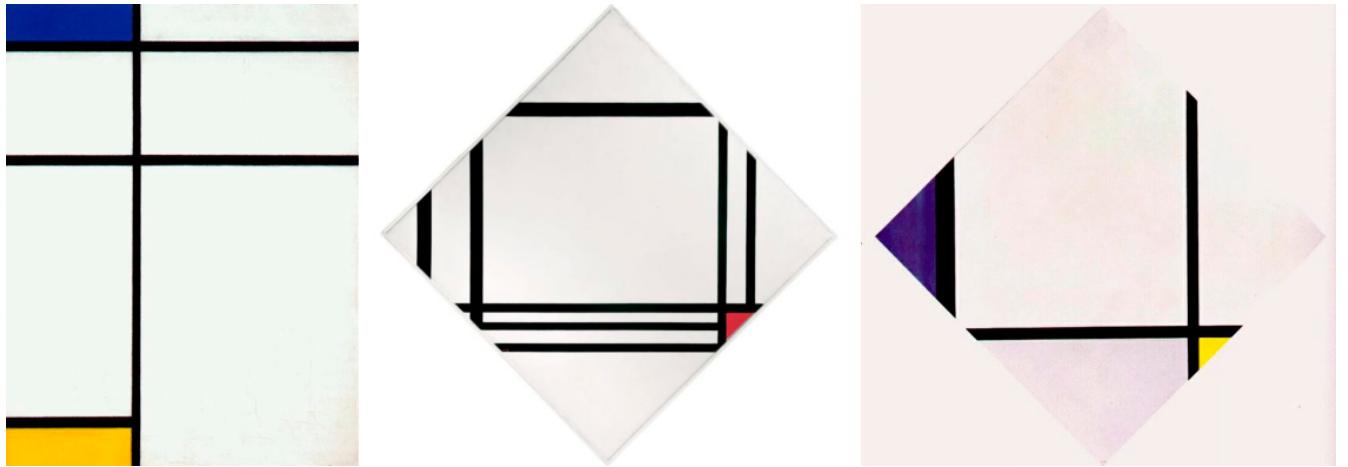


Figura 5: As 3 imagens com menor entropia: *Composition III with Blue, Yellow and White*, *PICTURE NO. III, Lozenge Composition 3 Lines Blue Gray Yellow*

2.1.3 Tons

Os tons, nomeados de *Redness*, *Blueness* e *Greeness*, foram dados da seguinte forma:

$$Tom^{Color} = \sum_{i=1}^{n_{row}} \sum_{j=1}^{n_{column}} \max(a_{ij}^{Color} - \frac{a_{ij}^{Color^*} + a_{ij}^{Color^{**}}}{2}, 0)$$

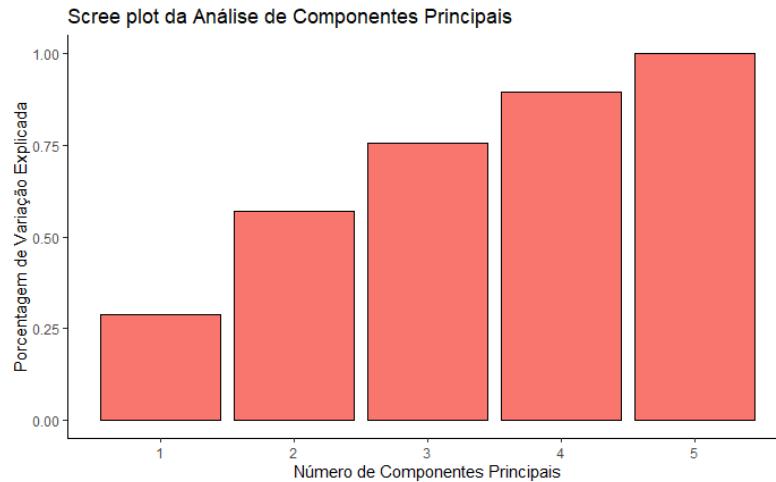
$$Color \neq Color^* \neq Color^{**} \in \{Red, Blue, Green\}$$

A intuição por trás dessa definição de tom se dá pelo desejo de uma medida que leva em consideração o destaque de uma das cores sobre as outras, e não da sua influência geral. Por exemplo, se um pixel roxo é formado por 1 vermelho, 1 azul e 0 verde, Redness será dado por 0.5, Blueness será dado por 0.5 e Greeness por 0, enquanto um pixel branco formado por 1 vermelho, 1 azul e 1 verde terá as 3 medidas como 0 (pois nenhuma se destaca sobre as outras).

2.2 Clusterização

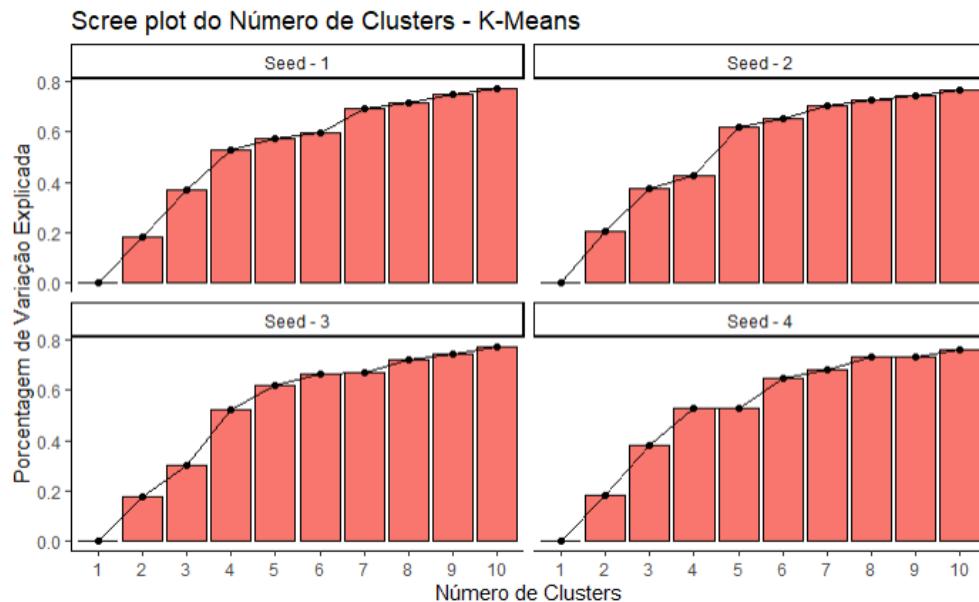
Para a clusterização, foi utilizada a técnica de k-means, baseado na intuição de que os grupos latentes se dividem em clusters relativamente esferoidais em 5 dimensões (Entropia, Constraste, Redness, Blueness e Greeness), devido ao caráter contínuo e consistente das variáveis, i.e. duas imagens são parecidas se seus valores de entropia, contraste e tons são parecidos.

Depois de obter as variáveis, elas foram tratadas para melhor se adequarem ao método de k-means. Para isso, primeiramente aplicou-se a operação de raiz quadrada à variável de contraste, para tornar a relação entre todas as variáveis linear, após isso, todas as variáveis foram escalonadas com a função `scale()` e, por fim, foi realizada uma análise de componentes principais para reduzir a dimensionalidade.



Percebe-se que as variáveis parecem ter uma correlação não tão alta entre si, pois as primeiras componentes principais não são suficientes para explicar grande parte da variação. Devido a isso, as análises que se seguiram se basearam nas 3 primeiras componentes principais (que representam 75% da variação), porém os gráficos só utilizarão 2 componentes principais, por simplicidade.

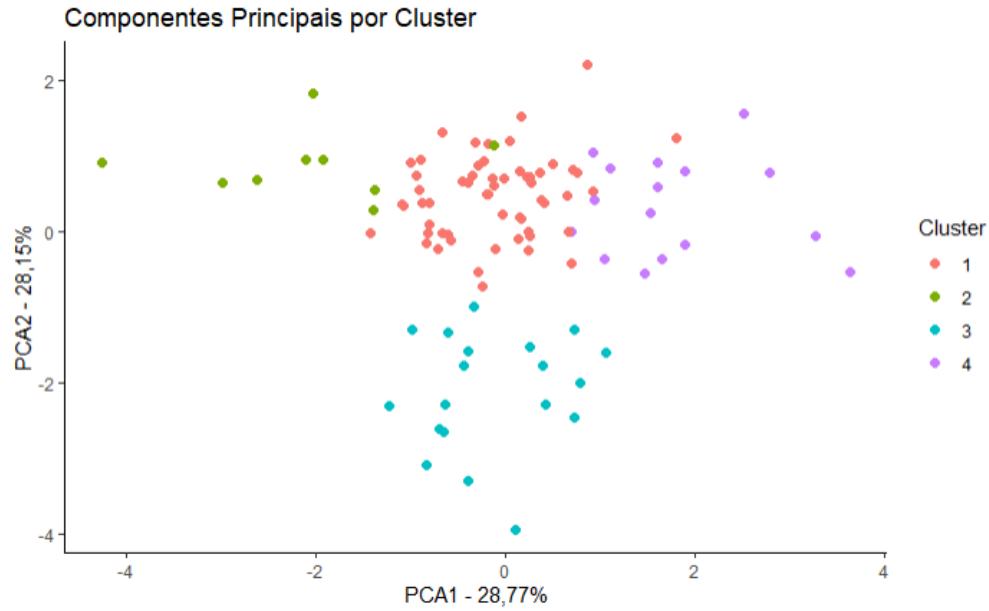
Para 4 seeds diferentes, o Scree plot relativo ao número de clusters utilizando o método do k-means é dado por:



Deste modo, entende-se que o ponto de inflexão se dá a partir de 4 clusters. Esse será o número de clusters calculados para o presente trabalho.

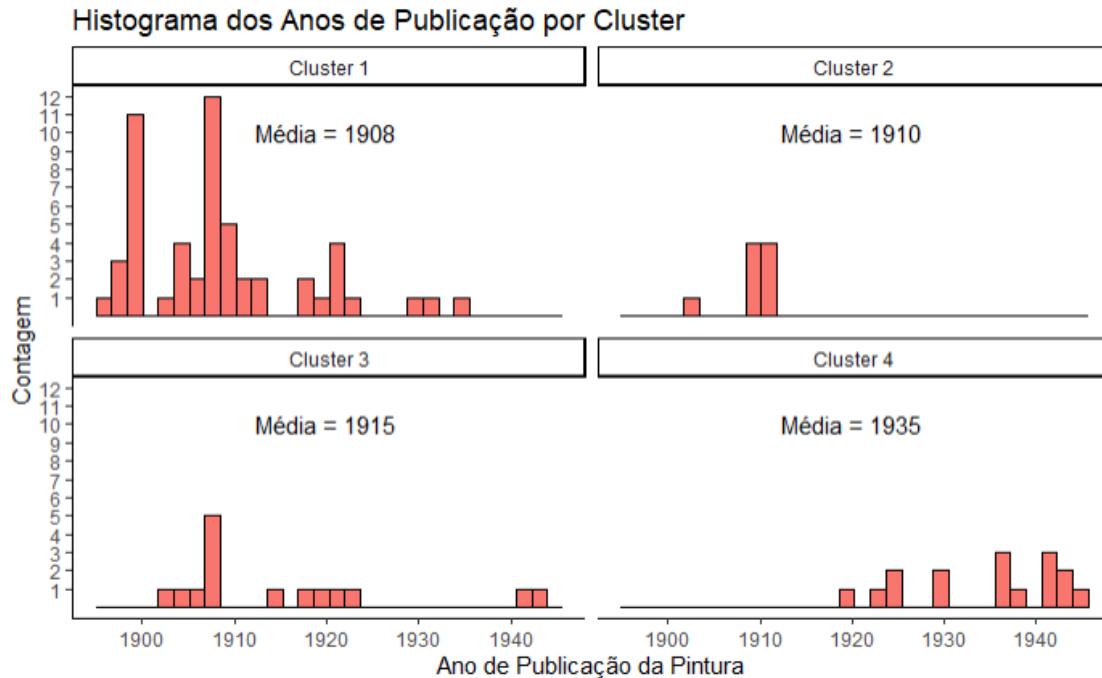
3 Resultados

Abaixo, apresenta-se o gráfico das posição das pinturas sobre suas duas primeiras componentes principais, coloridas por clusters.



Vale lembrar que a sobreposição de clusters se dá pois o método do k-means foi aplicado sobre 3 dimensões de componentes principais.

Também é possível analisar como a data de publicação dos quadros se comporta em cada cluster.



Abaixo, imagens-móvel de cada cluster.



Figura 6: 3 Pinturas escolhidas aleatoriamente do Cluster 1: *Study for Blue Apple Tree Series*, *Composición en Gris - Azul* e *Composition with Red, Yellow and Blue*



Figura 7: 3 Pinturas escolhidas aleatoriamente do Cluster 2: *Church tower at Domburg*, *Evolution* e *Evening landscape*



Figura 8: 3 Pinturas escolhidas aleatoriamente do Cluster 3: *Tableau 3 with Orange -Red, Yellow, Black, Blue and Gray*, *Windmill in the Gein* e *Along the Amstel*



Figura 9: 3 Pinturas escolhidas aleatoriamente do Cluster 4: *Composition No. IV, with Red, Blue and Yellow*, *Composition with Grid VII* e *Composition III with Blue, Yellow and White*

4 Discussão

Os clusteres construídos conseguiram dividir bem pelo menos 2 períodos da arte de Mondrian: o Cluster 2 (figura 7) engloba o período **expressionista** do pintor com destaque às obras com forte tom de azul, muito prevalente nas suas composições de 1907 a 1911. Já o cluster 4 (figura 9) se refere inteiramente às obras **neoplásticas**, movimento do qual Mondrian foi pioneiro, com suas composições características de traços puros, geométricos e abstratos.

O cluster 1 se concentra principalmente nos anos iniciais de Mondrian como artista, mas também possui obras de outros períodos. É caracterizado principalmente por pinturas ricas em detalhes e com cores pouco vivas. Já o Cluster 3 engloba um vários anos distintos da vida de Mondrian, contendo pinturas com cores fortes e que, por vezes, contém elementos reais.

O exercício da clusterização se provou eficiente não só para dividir as pinturas em períodos artísticos latentes, mas também para apontar as semelhanças entre os tipos de pintura diferentes com base nas suas composições de cores, traços e organização de elementos. Por exemplo, as poucas obras neoplásticas fora do cluster 4 podem indicar que a influência da inspiração em elementos externos durante sua composição teve um peso maior do que a busca pelo ideal de traços puros, indicando que essas pinturas podem despertar interpretações mais vinculadas à elementos reais, como é o caso de *Broadway Boogie Woogie*, uma pintura neoplástica alocada ao cluster 3, que tem forte conotação de elementos reais de trânsito e vida urbana na cidade de Nova York.

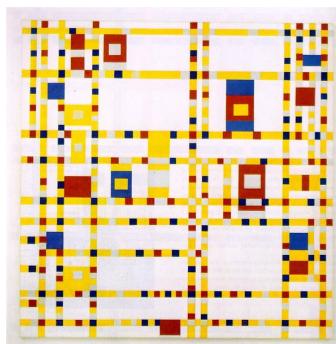


Figura 10: *Broadway Boogie Woogie* - 1943