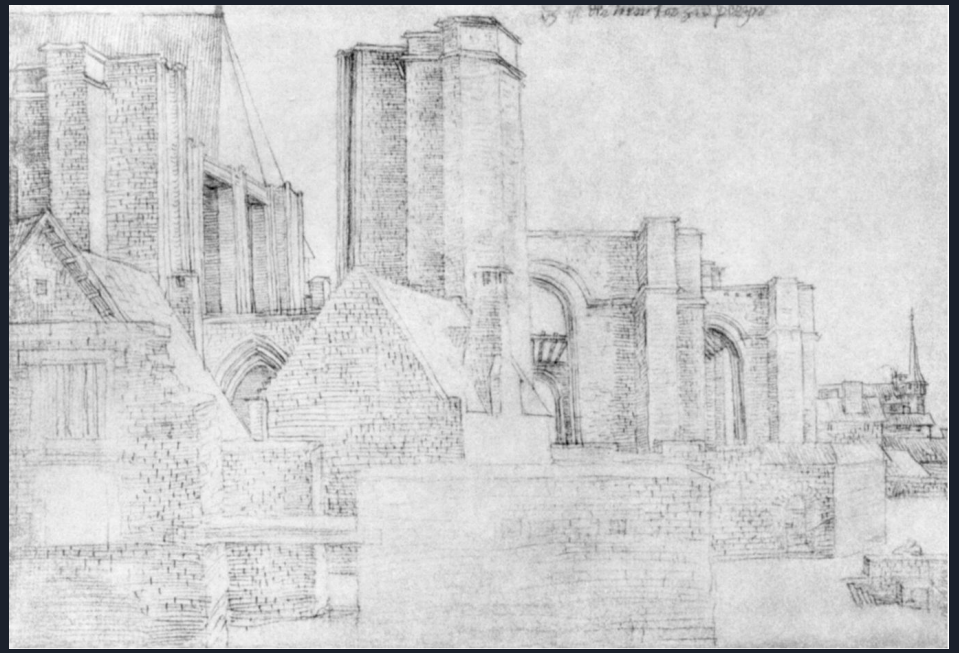
Artist Recognition via Keras

Haoyu Zhang, Patrick Kim, Yixun Zhang

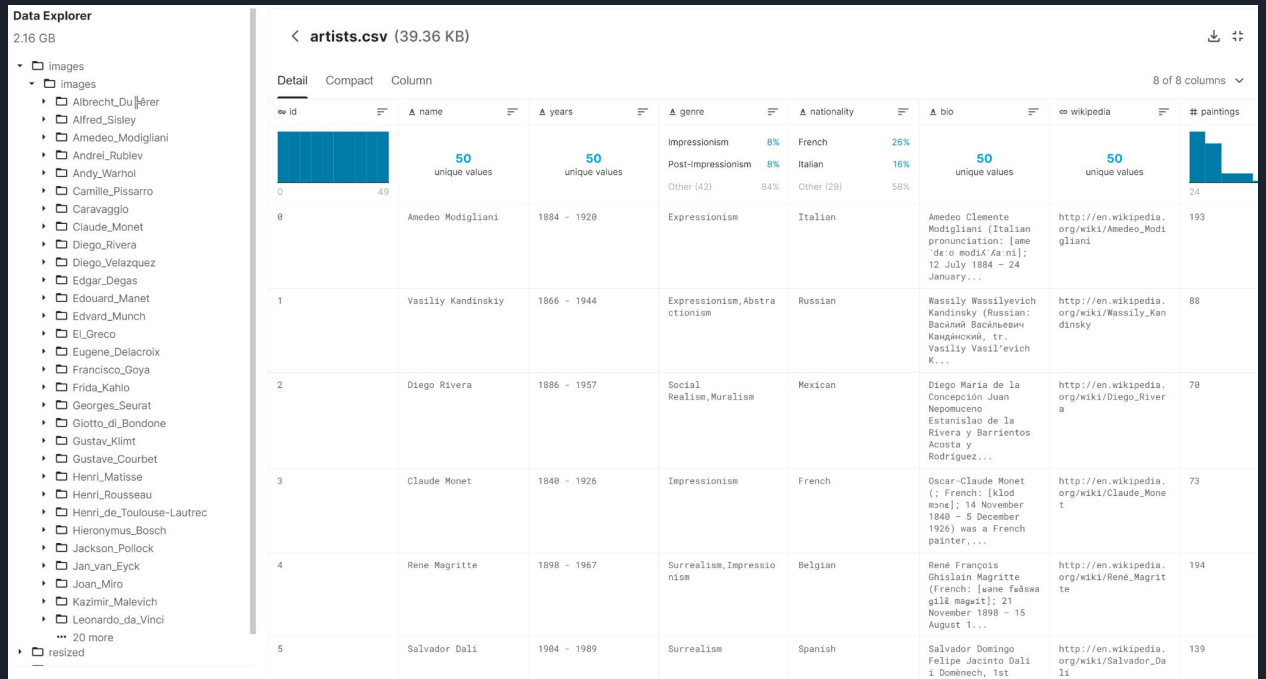
Goal

- Identify the painter of the images based on their painting styles.



Raw Dataset

- Kaggle dataset (https://www.kaggle.com/ikarus777/best-artworks-of-all-time)
- 8000+ labeled images from 50 prominent artists
- a CSV file with artist information



Data Explorer

2.16 GB

- images
  - images
    - Albrecht_Dürer
    - Alfred_Sisley
    - Amedeo_Modigliani
    - Andrei_Rublev
    - Andy_Warhol
    - Camille_Pissarro
    - Caravaggio
    - Claude_Monet
    - Diego_Rivera
    - Diego_Velazquez
    - Edgar_Degas
    - Edouard_Manet
    - Edvard_Munch
    - El_Greco
    - Eugene_Delacroix
    - Francisco_Goya
    - Frida_Kahlo
    - Georges_Seurat
    - Giotto_di_Bondone
    - Gustav_Klimt
    - Gustave_Courbet
    - Henri_Matisse
    - Henri_Rousseau
    - Henri_de_Toulouse-Lautrec
    - Hieronymus_Bosch
    - Jackson_Pollock
    - Jan_van_Eyck
    - Joan_Miro
    - Kazimir_Malevich
    - Leonardo_da_Vinci
    - ••• 20 more
- resized

artists.csv (39.36 KB)

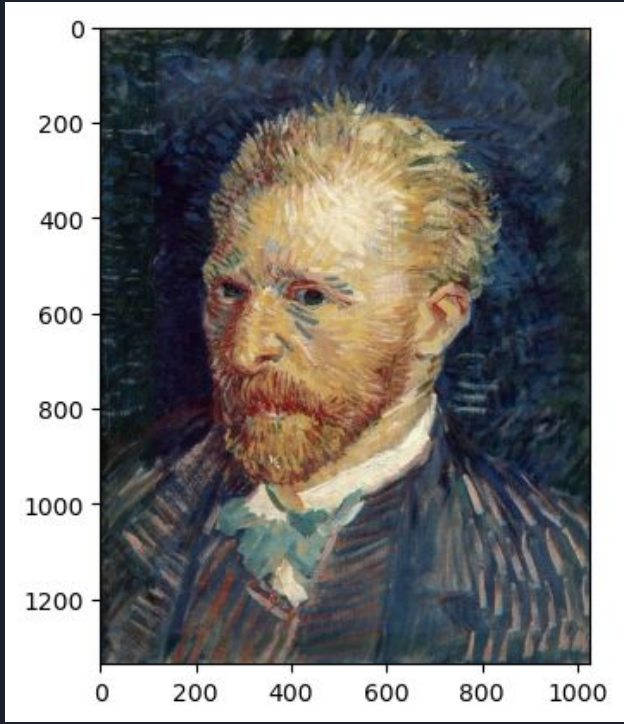
Detail | Compact | Column — 8 of 8 columns

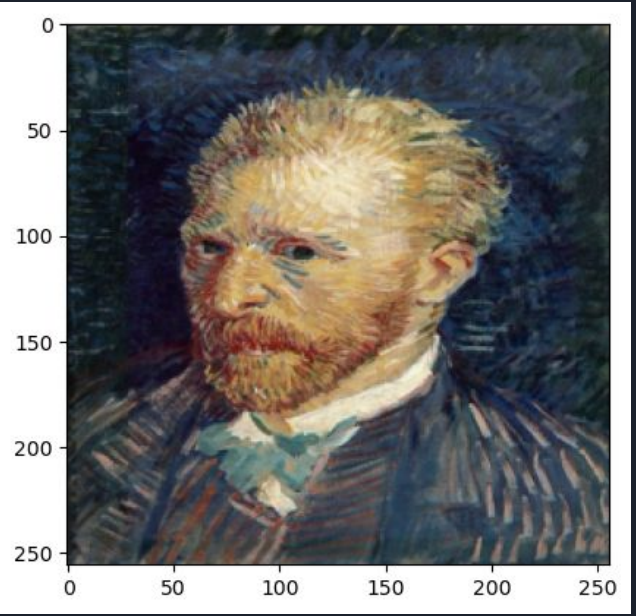
id	name	years	genre	nationality	bio	wikipedia	paintings
0 – 49	50 unique values	50 unique values	Impressionism 8%, Post-Impressionism 8%, Other (42) 84%	French 26%, Italian 16%, Other (29) 58%	50 unique values	50 unique values	24
0	Amedeo Modigliani	1884 - 1920	Expressionism	Italian	Amedeo Clemente Modigliani (Italian pronunciation: [ame ˈdeːo modiʎˈʎaːni]; 12 July 1884 – 24 January...	http://en.wikipedia.org/wiki/Amedeo_Modigliani	193
1	Vasiliy Kandinskiy	1866 - 1944	Expressionism,Abstractionism	Russian	Wassily Wassilyevich Kandinsky (Russian: Васи́лий Васи́льевич Канди́нский, tr. Vasíliy Vasíl'evich K...	http://en.wikipedia.org/wiki/Wassily_Kandinsky	88
2	Diego Rivera	1886 - 1957	Social Realism,Muralism	Mexican	Diego María de la Concepción Juan Nepomuceno Estanislao de la Rivera y Barrientos Acosta y Rodríguez...	http://en.wikipedia.org/wiki/Diego_Rivera	70
3	Claude Monet	1840 - 1926	Impressionism	French	Oscar-Claude Monet (; French: [klod mɔnɛ]; 14 November 1840 – 5 December 1926) was a French painter,...	http://en.wikipedia.org/wiki/Claude_Monet	73
4	Rene Magritte	1898 - 1967	Surrealism,Impressionism	Belgian	René François Ghislain Magritte (French: [ʁəne fʁɑswa gilɛ magʁit]; 21 November 1898 – 15 August 1...	http://en.wikipedia.org/wiki/René_Magritte	194
5	Salvador Dali	1904 - 1989	Surrealism	Spanish	Salvador Domingo Felipe Jacinto Dalí i Domènech, 1st	http://en.wikipedia.org/wiki/Salvador_Dalí	139

Preprocess Dataset

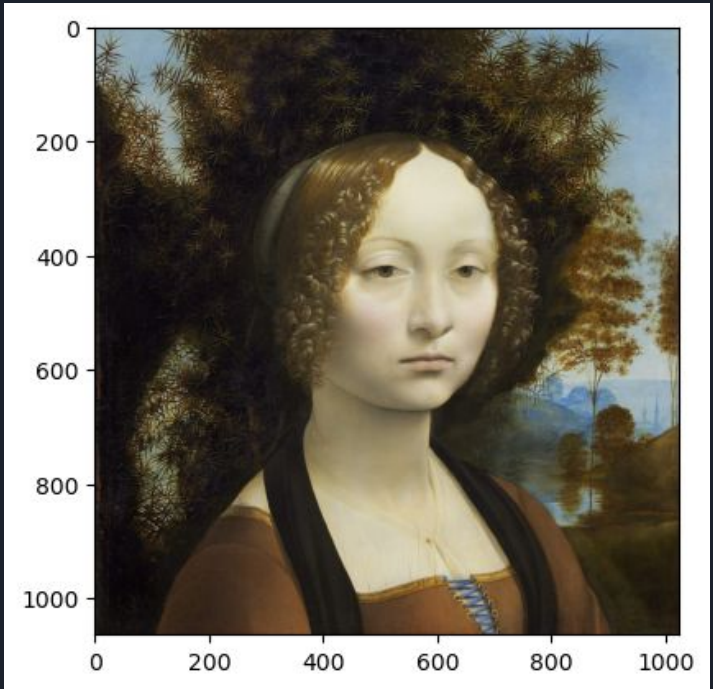
- Correct special characters in artist names
- Convert to 3 RGB channels
- Split data into 80% training data, 10% dev data and 10% test data.
- Create two new datasets from the raw datasets
  - Reshaped images
  - Sampling cropped images

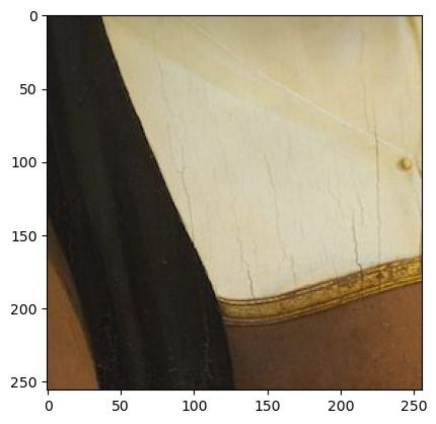
Image Reshape

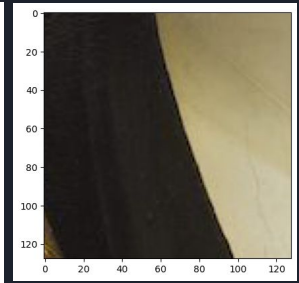




Random Sampling





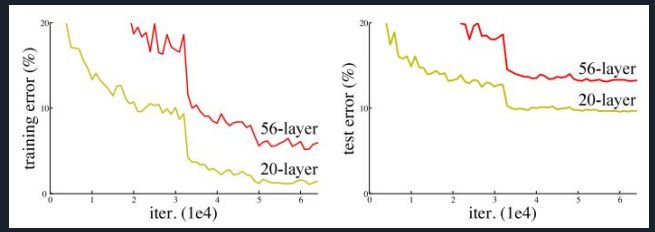


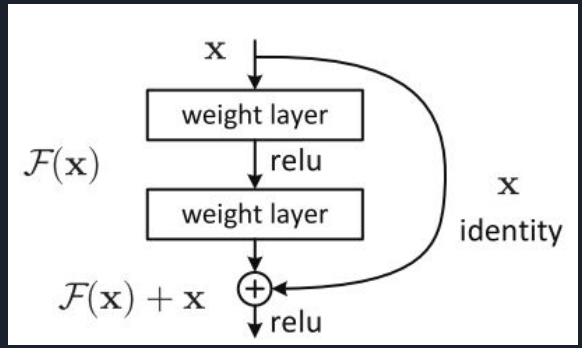
Infrastructure

- g4dn.2xlarge instance with an AMI for Nvidia deep learning
  - 8 vCPUs and 32 GB memory
- We exclusively use Keras for our model
- Jupyter notebook is run in the background for us to code our model

Model Frameworks

- Resnet-50 was chosen as our first model
  - This model is available in a single line in Keras
  - We added a dropout layer trying values ranging from 0.2 to 0.7
- We also experimented with MobileNet and other models available preconfigured in Keras
- We also experimented with various parameters and data resolutions





https://towardsdatascience.com/an-overview-of-resnet-and-its-variants-5281e2f56035

Model Results

	Image Size	Model	Preprocess	Top 1 Test Accuracy	Top 5 Test Accuracy
model_2c	256x256	MobileNet	Image Reshape	0.3597	0.652
model_2d	256x256	ResNet50	Image Reshape	0.3431	0.6674
model_4b	64x64	ResNet50	Random Sampling	0.3157	0.6402

Model Results - Model_2c

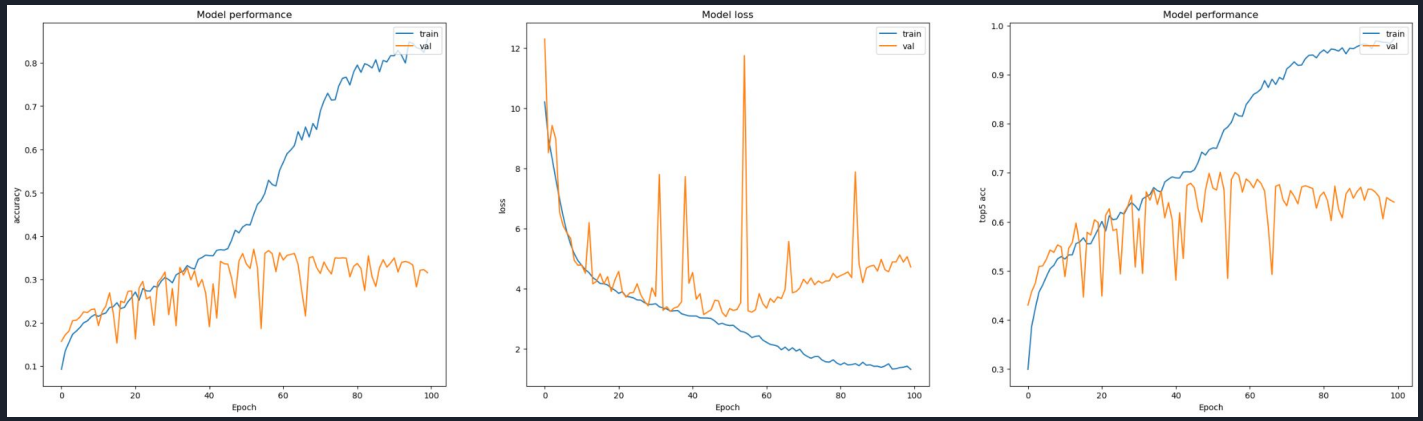
Model Results - Model_2d

Model Results - Model4b

Difficulties

Difficulties

- Overfitting
- Train data size
- Uneven distribution of artworks per artist



Thank You!
Questions?