ART CLASSIFICATION WITH DEEP LEARNING

JUSTINE PICAR, JULY 2021

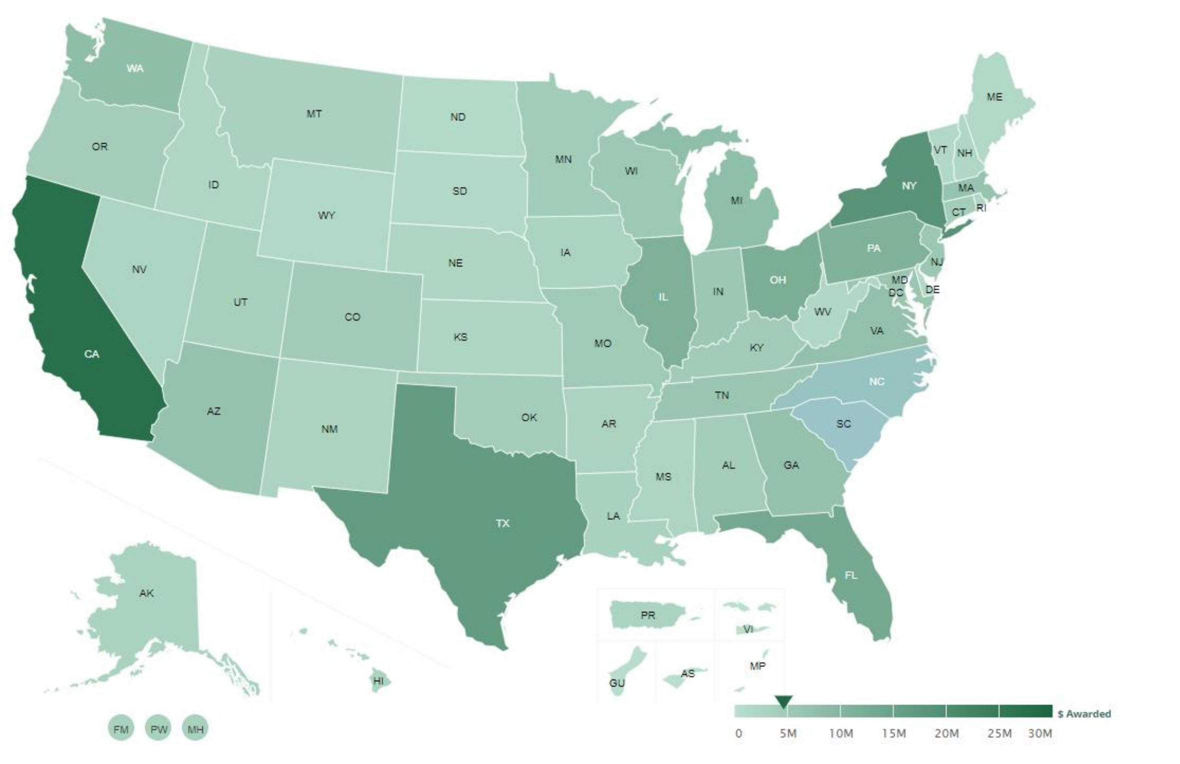
ART MUSEUMS: A BACKGROUND

- Important public institutions
- Hold a lot of cultural history
- A popular experience shared by many people with the help of the internet
- Opportunities exist to reach a broader, diverse group of people

PROBLEM

- Rely on increased audience engagement and patrons
- Lesser known museums cannot afford engaging programs and activities
- Museums are must index the genre of a painting before displaying it, which can be time consuming (source)
- New art enthusiasts can be intimidated by art without background knowledge
- An online poll by art start-up, Meural, shows that from a group of 1,000 people, only 18% of all ages in the US could identify who painted *The Girl with the Pearl Earring* (source)
- However, 20% of people would appreciate art if there was more accessibility

PROBLEM

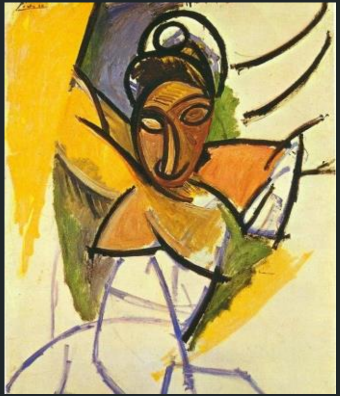


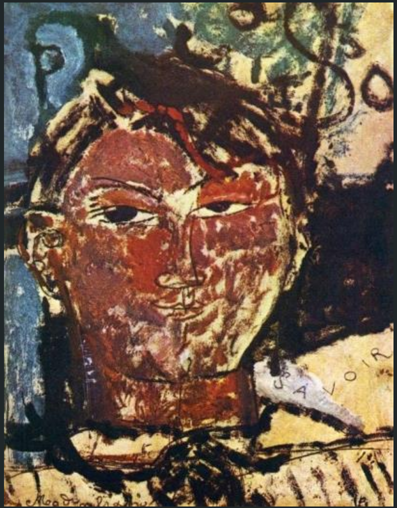
A map of funding for the 9k libraries and 35k museums in the U.S. during the fiscal year 2020 from the [Institute of Museum and Library Services (IMLS)]

PROJECT OBJECTIVE:

Classify the art style of a painting to use as an interactive tool for patrons and for museum curators to label new paintings before an exhibition

DATA: KAGGLE'S BEST ARTWORKS OF ALL TIME



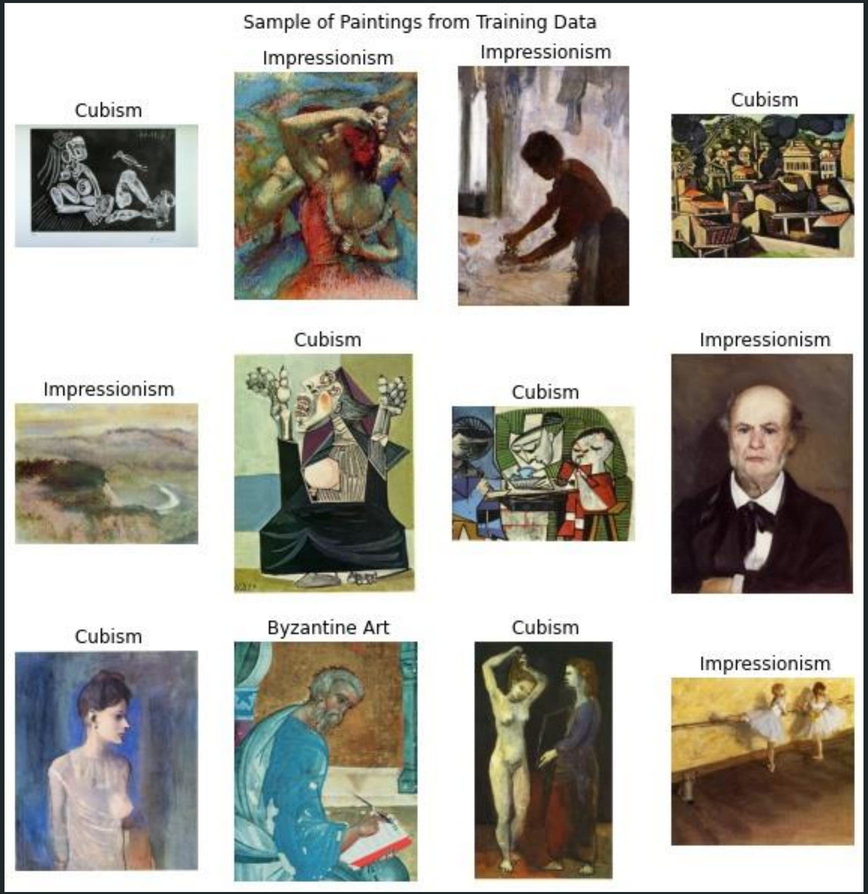


- 8,446 PAINTINGS
- FROM TOP 50 RENOWNED ARTISTS (Warhol, Monet, Degas, etc)
- 20 UNIQUE ART STYLES

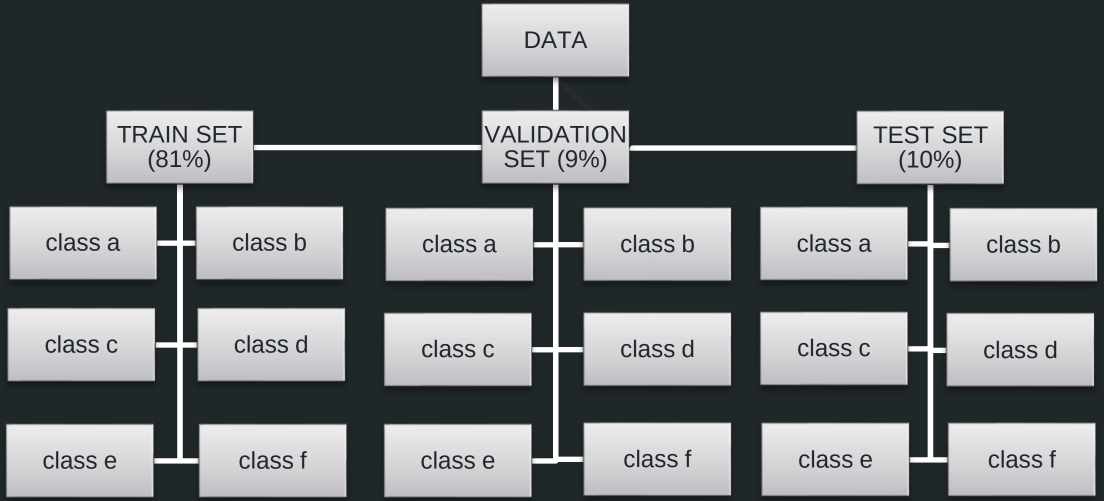
DATA WRANGLING

- Classified by artists; relabeled by genre
- 20/31 art styles are unique; drop columns with multiple genres
- Focus on 6 particular classes:
  - Impressionism
  - Cubism
  - Expressionism
  - Pop Art
  - Byzantine Art
  - Abstract Expressionism
- Total of 2,206 images
- 81% training, 9% validation, 10% testing*

* There is a separate testing set independent of the training and validation set so the model will not “cheat” or “learn”



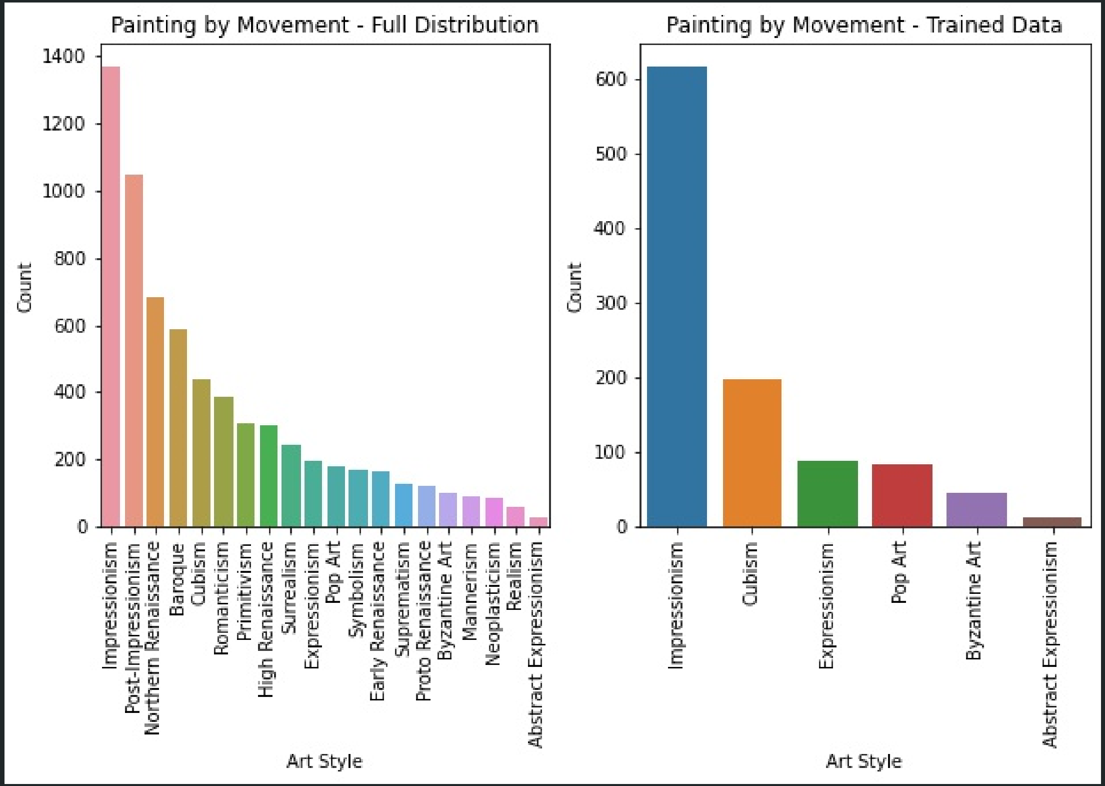
BREAKDOWN OF DATA



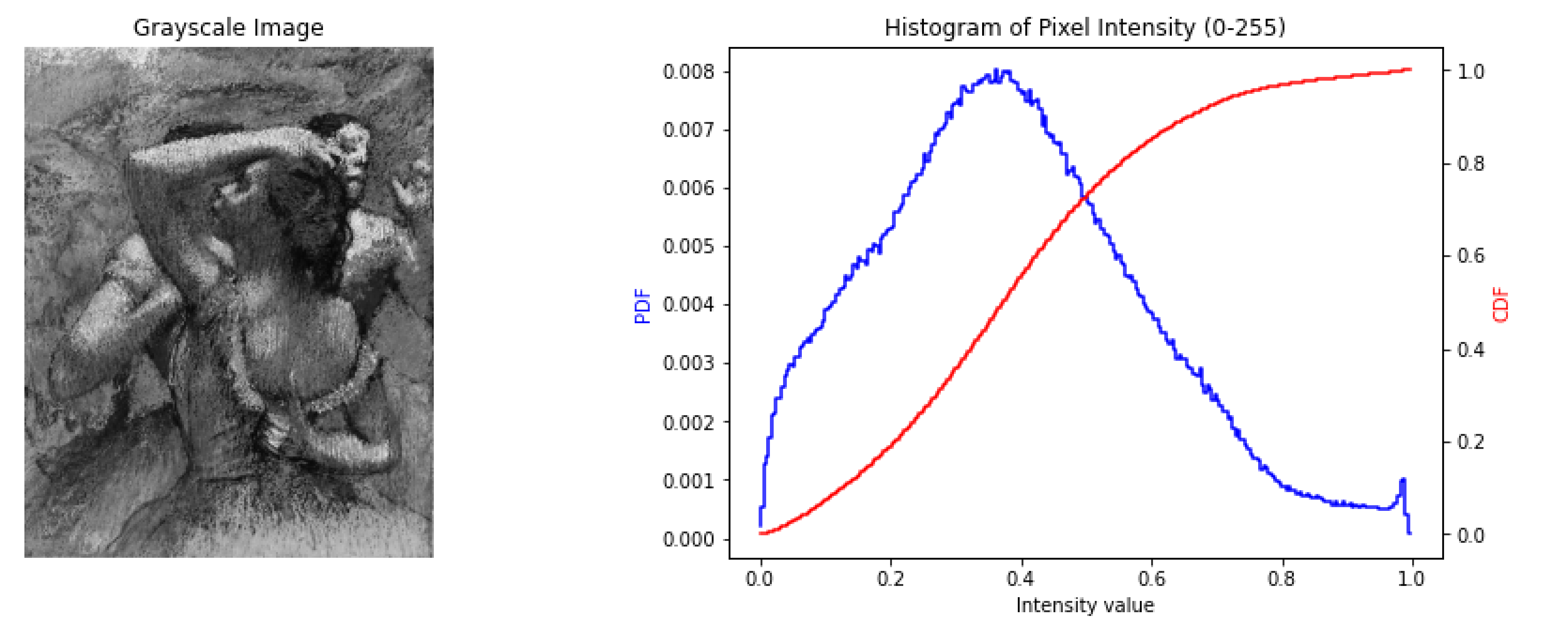
PACKAGES

- For training and modeling the data
  - keras
- For visualizing the raw data, the pre-processed images, and metrics
  - skimage
  - matplotlib.pyplot
  - seaborn
- For visualizing neural network architectures
  - graphviz
  - pydot

EDA: PAINTING DISTRIBUTION

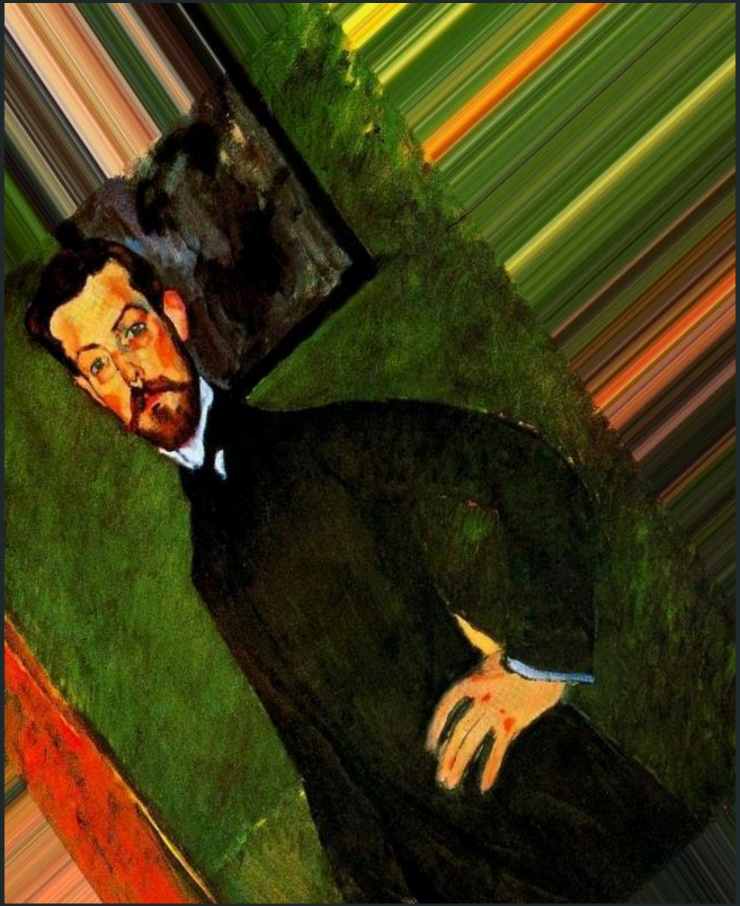


EDA: PIXEL INTENSITY



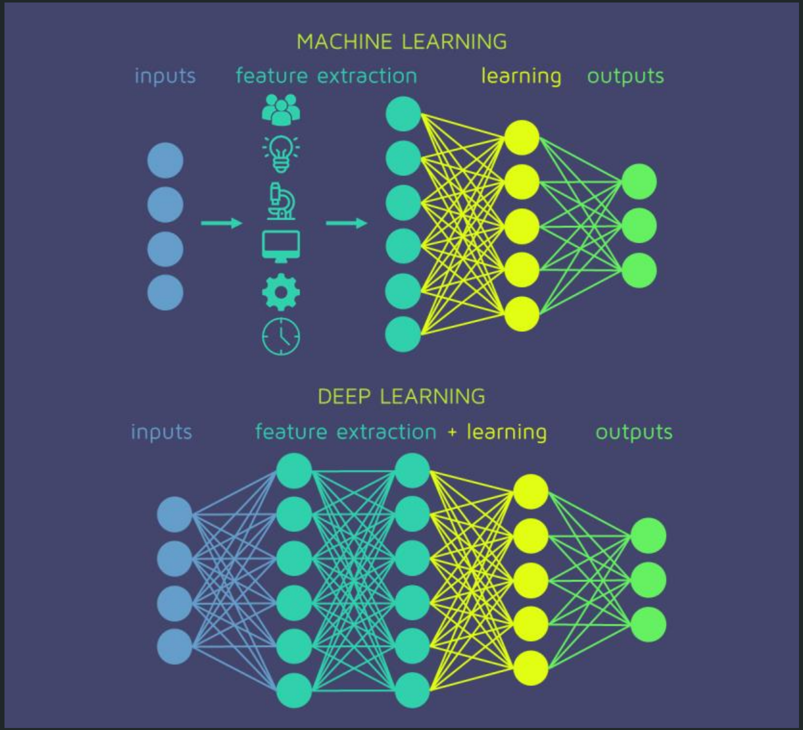
METHODOLOGY: DATA AUGMENTATION & PRE-PROCESSING

- Increase robustness of models and ability to generalize data
- Maintain distribution
- Apply distortions randomly
  - Rotations
  - Nearest ‘fill’ method
  - Horizontal flip
  - Zoom
  - Stretch (width and height)



METHODOLOGY: USING DEEP LEARNING

- DEEP LEARNING: a subset of machine learning that is more complex and uses neural networks
- NEURAL NETWORKS: computing system inspired by the brain; receives, processes and “signals” other neurons to output a sum of its inputs with input, hidden, and output layers
- WEIGHT INITIALIZATION: sets the weights of a neural network to adjust the learning of the model
- PRE-TRAINING/TRANSFER LEARNING: Uses a pre-trained model trained on a large benchmark dataset (ex. ImageNet) to solve a similar problem to the one we want to solve
- FINE-TUNING: process of adjusting or tuning parameters in a model to perform a similar task or to perform better



METHODOLOGY

WEIGHT INITIALIZATION

- class_weights from sklearn.utils
- balance the distribution between art styles in the training data

PRE-TRAINING

- Transfer learning with the ImageNet database
- Train on top of a RESNET50 model

FINE TUNING

- Re-train last 20% of layers
- Freeze all other layers
- Slower learning rate
- Use Stochastic Gradient Descent (SGD) Optimizer

SIMPLE CNN VS. RESNET50

Simple Convolutional Neural Network (CNN)

- Specializes in processing data that has a grid-like topology, such as an image
- Each layer is fully connected to all neurons in the layer
- As more layers are added, the performance of the model degrades rapidly due to vanishing gradients

Residual Neural Network (RESNET50)

- A type of convolutional neural network
- Uses skip connections to train the model at deeper layers and mitigates the issue of gradient descent
- Skip connections ensure the higher layers will perform at least as good the lower layers
- Skip connection adds the original input to the output of a convolutional block

NEURAL NETWORK ARCHITECTURES

Simple Convolutional Neural Network (CNN)

- Input shape: 200 x 200 x 3
- 3 stages with structure:
  - Conv2D
  - RELU Activation
  - MaxPooling2D
  - Filters at each stage: 32, 64,128 respectively
- Final Stage:
  - Flatten
  - Dense Layer with 128 feature maps
  - Dropout ~40% of the parameters to prevent overfitting
  - Dense layer = the number of classes (6)
  - Softmax activation for multi-classification

Click Here for Visual

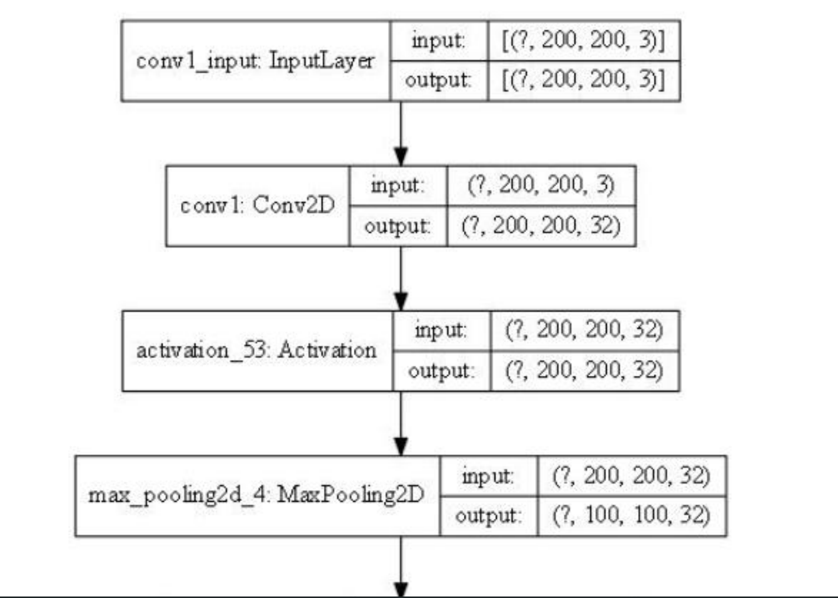
Residual Neural Network (RESNET50)

- Input shape: 200 x 200 x 3
- Identity Block with skip connection
  - Conv2D -> BatchNormalization
- Convolutional Block with skip connection
  - Conv2D -> BatchNormalization
- Residual Neural Network
  - Conv2D
  - BatchNormalization
  - MaxPooling2D
  - 4 stages of convolutional and identity block
  - AveragePooling
  - Dense layer = the number of classes (6)
  - Softmax activation for multi-classification
- Total: 53 layers

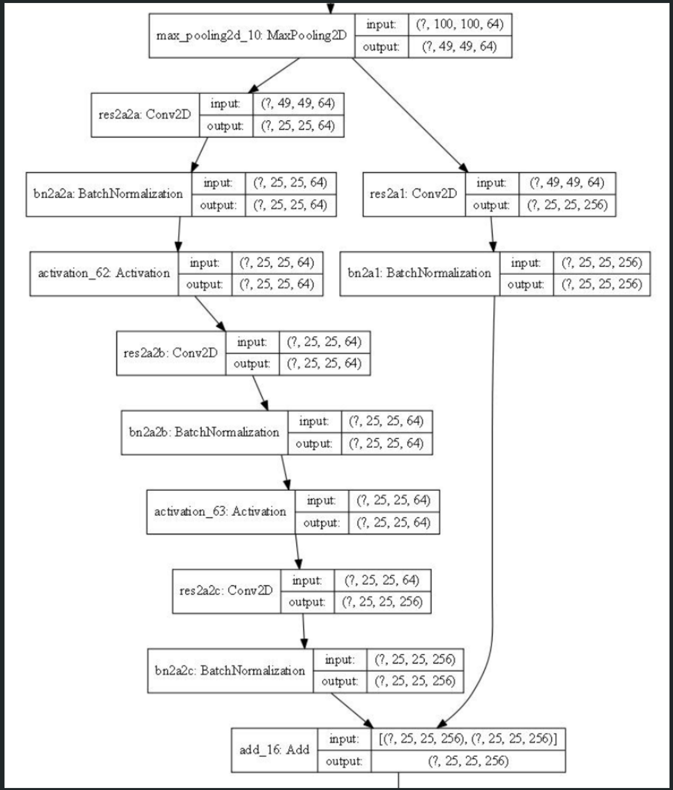
Click Here for Visual

NEURAL NETWORK ARCHITECTURES

Convolutional Neural Network (CNN)

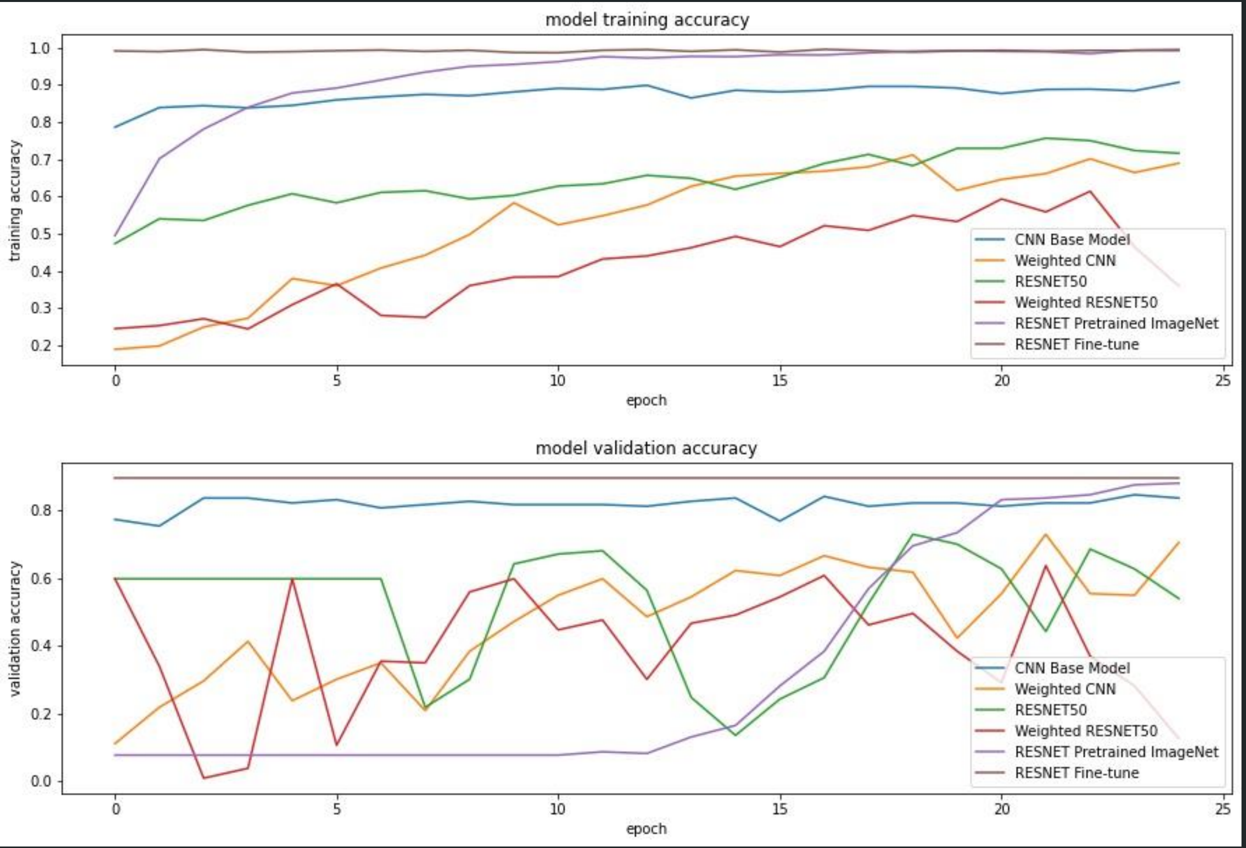


Residual Neural Network (RESNET50)



MODEL RESULTS AND ANALYSIS

We use accuracy as the main metric when comparing the models because the goal is to predict a specific label for each class



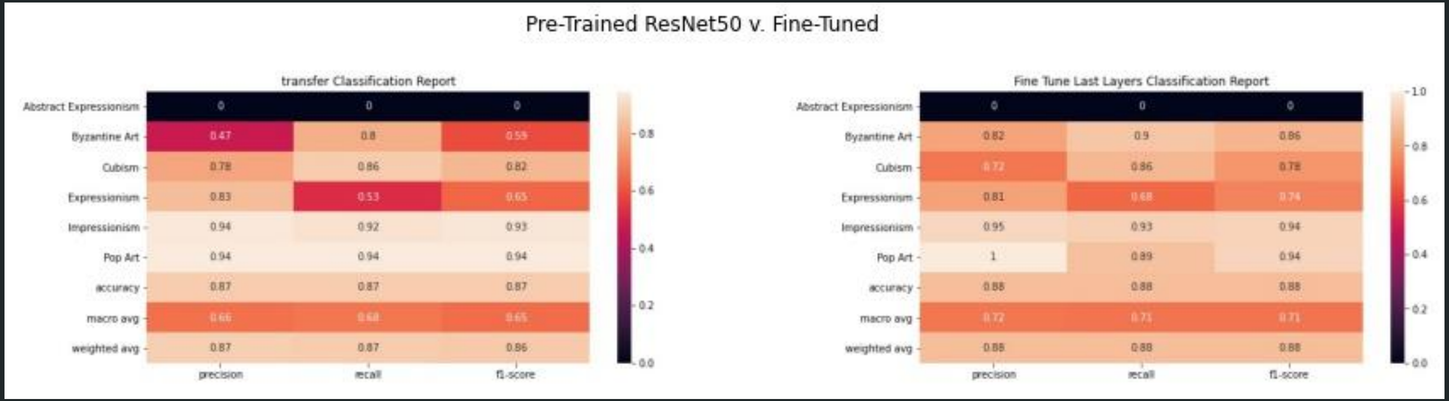
MODEL RESULTS AND ANALYSIS

	CNN	CNN_weighted	ResNet50	ResNet50_Weighted	ResNet50_Pre-Trained	ResNet50_Fine_Tune
Pop Art	0.888889	0.500000	0.888889	0.500000	0.944444	0.888889
Impressionism	0.824818	0.642336	0.547445	0.642336	0.919708	0.927007
Cubism	0.636364	0.568182	0.568182	0.568182	0.863636	0.863636
Expressionism	0.631579	0.578947	0.421053	0.578947	0.526316	0.684211
Abstract Expressionism	0.500000	0.500000	NaN	0.500000	NaN	NaN
Byzantine Art	0.400000	0.500000	NaN	0.500000	0.800000	0.900000

PRE-TRAINED RESNET50 WAS THE BEST MODEL

WEIGHTED RESNET50 WAS THE WORST MODEL

MODEL RESULTS AND ANALYSIS



Key Findings

- Abstract Expressionism labels misclassified as Cubism and Impressionism
- Impressionism has a very open composition that appears abstract up close
- Cubism is specific, yet abstract form of art

Key Findings

- Pop Art highly recognizable due to unique style (repeating patterns)
- Model detects styles with distinct repeating patterns
- Struggles with styles closer to one another, much like a person

CONCLUSION & LESSONS LEARNED

- 87% ACCURACY ON PRE-TRAINED RESNET50
- PRE-TRAIN OTHER MODELS:
  - VGG-16
  - INCEPTIONV3
  - EFFICIENT NET
- METHODOLOGY:
  - COMBINING CLASSES
  - BAGGING
  - CROSS-VALIDATION
- DIVERSE DATASET

THANK YOU!

Thank you to my patient and wonderful mentor Nik Skhirtladze, Francois Chollet and Priya Dwivedi on their helpful tutorials, and Stack Overflow for all my troubleshooting needs.

SOURCES

- Image Classification Using Very Little Data by Francois Chollet
- Kaggle Best Artworks of All Time
- Keras API
- Map of All 35,000 Museums in the United States Reveals Empty Region
- MOMA Contemporary Challenges
- Stack Overflow
- Understanding ResNet in Keras by Priya Dwivedi