Painter Classification using Deep learning

Introduction:

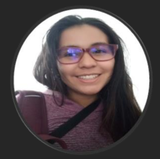
This machine learning project seeks to develop a **neural network** model that can **distinguish the author of a given painting** from a defined list of painters. For this task, the Best Artworks of all time Dataset will be used, which consists of a set of artworks labeled with their respective author. The aim is to carry out several tests with different architectures in order to establish a model that can correctly perform the classification.

Example of a painting classification:

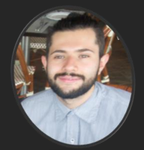
Actual artist = Paul Gauguin
Predicted artist = Paul Gauguin
Prediction probability = 92.08 %

Developers:

Jimmy Pulido
jiapulidoar@unal.edu.co



Laura Santos
lbsantosg@unal.edu.co



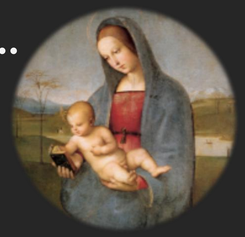
Juan Chaves
jschavesr@unal.edu.co

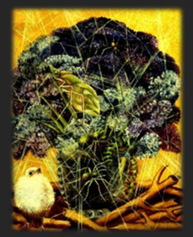
For more information:

Video

Paper

Guess the artist...

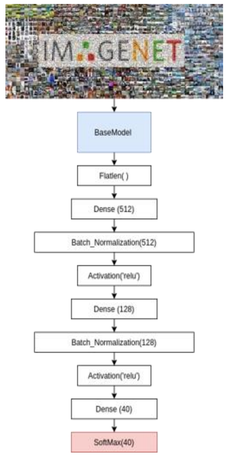




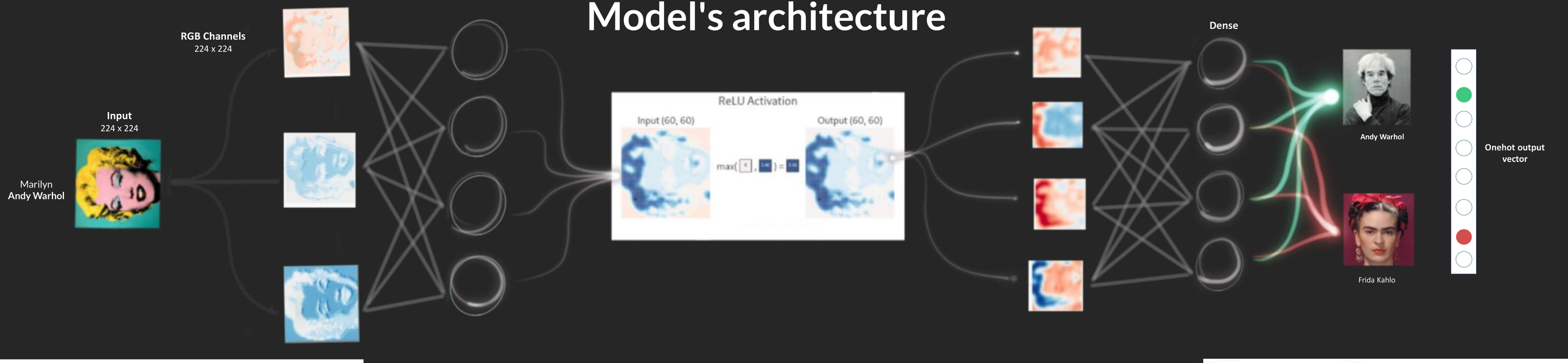
Method and Setup

The approach is to use transfer learning technic, we use a base model which was pre-trained with a specific architecture. As base model three different pre-trained models were used, Resnet50, InceptionV3 and InceptionResnetV2, all of them are convolutional image recognition neural networks model, trained using the ImageNet dataset. The output layer is a vector of 40 positions, each one represents the probability of the painters to be the author of a given image.

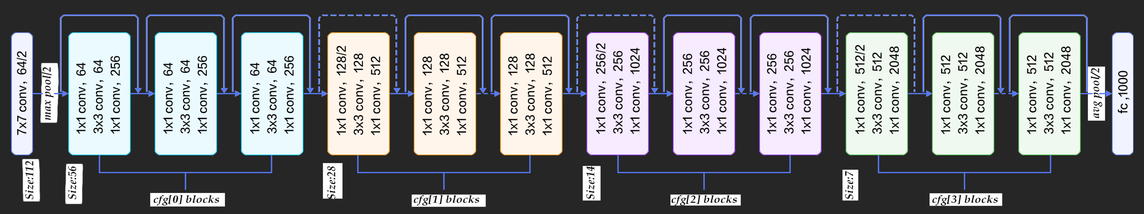
The dataset was divided into test and training set using a proportion of 70% - 30% respectively. From the training partition, is taken 10% of the data as the validation set. This partition seeks to been able to test the model with unseen data, and so to generate unbiased results while testing the test set



Model's architecture



Resnet50's architecture



Results

All the three models were tested using the same training/test set. To test the model tree different approaches were used:

•Top 1 : A result is consider to be correct just if the real author of the painting corresponds to the artist with the highest probability in the output vector of the model.

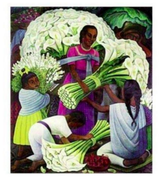
•Top 3 and Top 5: A result is consider to be correct if the real author of the painting is found in the artists with three or five highest probabilities of the output vector of the model, respectively.

Resnet50			
Description	Precision	Recall	F1-score
Top 1	0.66	0.70	0.68
Top 3	0.85	0.86	0.85
Top 5	0.90	0.91	0.91

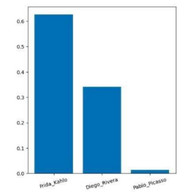
InceptionV3			
Description	Precision	Recall	F1-score
Top 1	0.60	0.65	0.61
Top 3	0.80	0.83	0.81
Top 5	0.87	0.89	0.88

InceptionResNetV2			
Description	Precision	Recall	F1-score
Top 1	0.67	0.70	0.68
Top 3	0.84	0.86	0.85
Top 5	0.90	0.90	0.90

Results



Actual artist = Diego Rivera
Predicted artist = Frida Kahlo
Prediction probability = 62.58 %



Actual artist = Diego Rivera
Predicted artist = Diego Rivera
Prediction probability = 39.98 %