Intro to ML - Final Project Report

Jake Blackburn
College of William and Mary
Williamsburg, VA, USA
jablackburn@wm.edu

***Abstract*—This project seeks to explore whether machines can appreciate aesthetics by testing whether a model designed with modern ML techniques will be better able to predict the price of an artwork at varying levels of complexity. The results show that the model tested did not benefit from additional complexity. This supports the idea that ML techniques will fail to model artistic beauty, even at large scales, however the project suffers from various limitations and therefore does not assert this conclusively.**

I. Introduction

Can modern ML techniques appreciate and understand (or at least model an understanding of) aesthetic beauty, especially in the highly subjective realm of visual art? This topic has deep (and deeply confusing) implications on the limitations of modern ML technology, how that technology compares to human intelligence, and the future of AI, not to mention many deeper philosophical questions about the human brain (e.g., what is meant by ”aesthetic beauty”? what about the human brain makes us appreciate art?).

Previous work using the same dataset [2] as used this project has found success in predicting the sale price of a work of art based on an image of the artwork as well as metadata about that artwork such as the level of fame of the artist [3]. In contrast, this project focuses on how ML models can learn how aesthetically beautiful an artwork is (crudely approximated by the sale price) based on an image of the artwork alone.

This is tested with a transfer-learning-based model, which leverages a pre-trained image classification model to extract features of an image, and then pass those features to a simple Multi-Layer Perceptron which predicts the price of the artwork in the image.

The initial prediction of this experimental project is that simple applications of modern ML technology will fail to capture the essence of beauty by predicting the price of artworks. Specifically, this project aims to show that increasing the size and complexity of the model will not meaningfully improve convergence. To prove / disprove this, the model is evaluated at two different sizes, showing how the scale of the model affects convergence.

II. Approaches

A. Pre-Trained CNN

Rather than implementing an image processing model from scratch, the first step in this project applies transfer learning techniques, using a pre-trained model (ResNet101) [1] to capture high level features of a given image for later processing. ResNet101 is a CNN which uses residual learning for computer vision and pattern recognition. This model can be leveraged to extract higher level features of an image which will be used to predict the price in the next step. This means the next layer of the prediction mechanism doesn't need to use all the image data, just the abstracted and compressed representation created by ResNet. To achieve this, the final classification output layer of the ResNet model is dropped; the output of the second-to-final layer is then fed directly to the next stage of the model.

B. Multi-Layer Perceptron

The features extracted by ResNet101 are then passed to a simple Multi-Layer Perceptron which predicts the price of the artwork. The MLP uses SGD optimization and mean squared error/loss. It has 2048 inputs, corresponding to the final non-output layer of ResNet101, and each hidden layer is fully connected and has 512 nodes. The MLP uses the sigmoid activation function to produce outputs in the range between zero and one. All of the prices in the dataset are also scaled into the same range to improve convergence and such that the output of MLP can compare directly to these normalized prices, rather than requiring the output to be scaled in reverse to a true price.

C. Training & Validation

The model was trained iteratively in batches. After each round of training, the model was validated iteratively against batches of the same size and the average loss across these batches was recorded to show how the model converges across the rounds / epochs.

D. Dataset

The project uses ”Art Price Dataset” [2], a set of 754 images of artworks that have been offered for sale through Sotheby's. The images are pre-processed to fit the dimensions required by the model (244x244).

III. Evaluation

To show that scaling the model does not meaningfully improve convergence, the model was evaluated in two sets of trials, one where the MLP had two hidden layers, the other with three. At both of the sizes, the model was tested with two different learning rates (0.0001 and 0.001). Each trial was run with 12 epochs and a training/validation batch size of 16.

A. *Learning Rate Comparison*

Learning rates of 0.001 showed no pattern of convergence, instead presenting oscillating patterns of loss, (see Fig. 1) while learning rates at 0.0001 showed a consistent pattern of convergence, however, the average loss was still extremely high.

B. *Size Comparison*

Trials with two and three hidden layers in the MLP did not show significant differences in the average loss, both converging to very high values (see Fig. 2).

IV. Discussion

The trials show that in this constrained scenario, increasing model size does not meaningfully improve performance in predicting the price of an artwork based on the image alone. These results do not completely answer the question of how size improves or fails to improve price prediction. For instance, the model is tested at only two, relatively small scales, the dataset is very limited which may play a role in poor performance, and the choice of learning rate, activation function, scaling and loss function was not fine-tuned in any significant way. Future projects investigating this problem might try experimenting with more hidden layers / nodes, using different functions and hyper-parameters for the MLP, or try a different approach entirely such as a CNN in place of the MLP or fine tuning the pretrained model along with the MLP.

Furthermore, In the context of the overall question of whether ML models can ”understand” beauty, this project falls short of conclusive results. Price as a measure of aesthetics is not perfect; as we have seen from previous projects, other factors such as the artist's fame play an important role in determining the sale price. To explore this question further, future research could involve aggregating more subjective measures of an artwork's beauty, for example taking surveys of how beautiful a work is on a scale of 1-10, or asking how much people would pay for a work of art irregardless of other factors like the fame of the artist or how much it could be resold for.

V. Conclusion

This project attempted to answer the broad question of whether modern ML techniques can model aesthetic beauty. Specifically, we looked at how a transfer learning model composed of an advanced image classification CNN in concert with a simple MLP could learn to predict how beautiful a work of art was, measured by the price of that artwork. The project showed that in this context, increasing the complexity of the model did not improve convergence, thus loosely supporting the idea that this kind of model cannot meaningfully approximate an appreciation of beauty. While the project suffered from many limitations that kept it from answering the question fully, it also opens the door to more research that could test a wider range of ML techniques at much larger scales, and with more data and better metrics for aesthetic beauty.

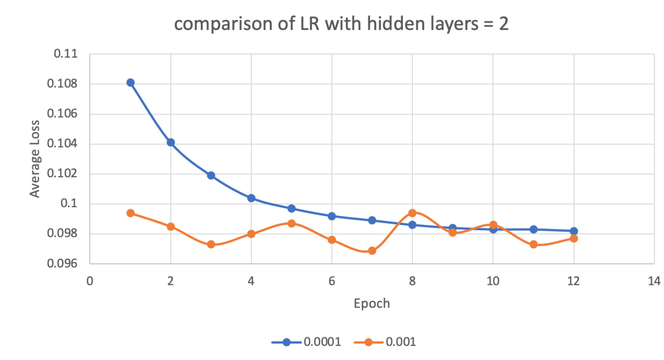

Fig. 1. Comparison of trials 1 and 2

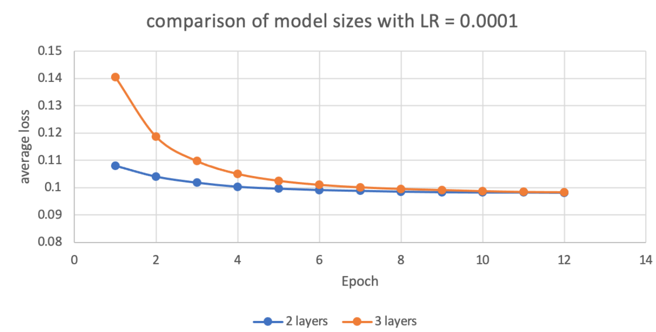

Fig. 2. Comparison of trials 1 and 3

References

[1] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition. arXiv. https://arxiv.org/abs/1512.03385

[2] Kuhm, F. (n.d.). Art Price Dataset. Kaggle. Retrieved from https://www.kaggle.com/datasets/flkuhm/art-price-dataset/data?select=artDataset

[3] Lau, I. (n.d.). Art Price Prediction Model. GitHub. Retrieved from https://github.com/Idanlau/Art-Price-Prediction-Model/blob/main/README.md

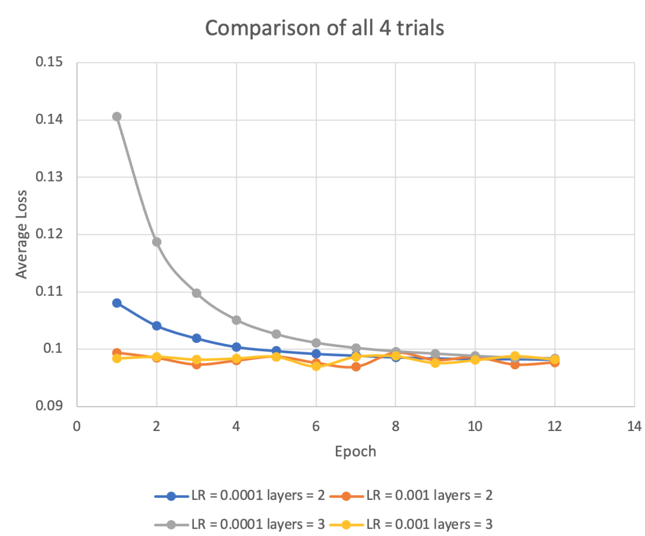

Fig. 3. Comparison of all trials