Enhancing Art Classification: A Comparative Study of CNN, Transfer Learning, and SVM Models

A Comprehensive Approach to Identifying Art Styles

Aditya Prakash (22201796), Shreya Grover (22200383)

School of Mathematics and Statistics, University College Dublin

Contact Information:
Email: aditya.prakash@ucdconnect.ie
shreya.grover@ucdconnect.ie

Introduction

Art classification and authentication have been subjects of increasing interest in both the art and scientific communities. Techniques such as automated analysis of drawings at the stroke level have been applied for attribution and authentication [2]. Additionally, machine learning has been utilized for the identification of art paintings [1].

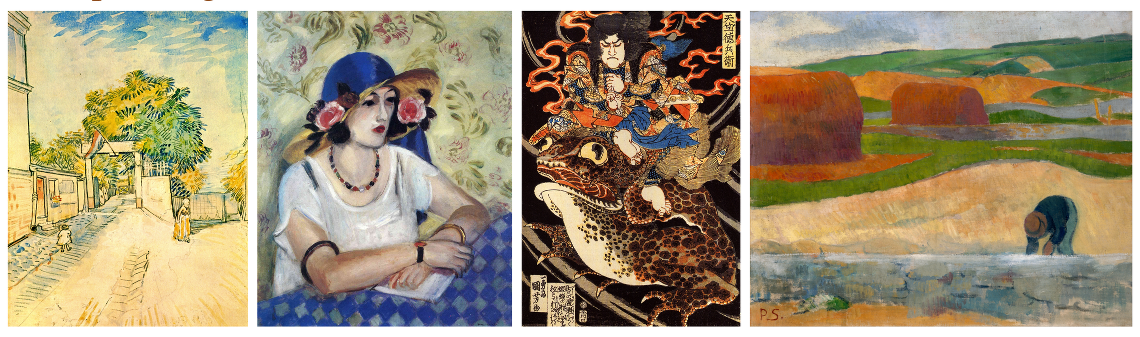

Figure 1: Some of the paintings used in the dataset. Works by Van Gogh, Matisse, Toyokuni, and Aivazovsky.

Dataset Description

The dataset primarily sourced from WikiArt.org encompasses a total of 2,319 different artists and 43 unique painting genres. For practical classification, a threshold of 500 paintings per artist was set to filter the dataset, resulting in a balanced selection that includes 12 different artists and 30 unique painting genres.

- **Total Artists:** 12
- **Total Genres:** 30
- **Artists Included:** Ivan Aivazovsky, Gustave Dore, Rembrandt, Pierre-Auguste Renoir, Albrecht Durer, Ivan Shishkin, Giovanni Battista Piranesi, John Singer Sargent, Zdislav Beksinski, Ilya Repin, Pablo Picasso, Marc Chagall

This approach ensures a robust dataset for classification by excluding artists with insufficient samples and retaining those with substantial numbers. The images in the dataset are utilized exclusively for data mining, adhering to fair use principles, and are assumed to be protected by copyright.

Main Objectives

1. Develop a custom CNN for art classification.
2. Utilize transfer learning with ResNet50 for enhanced performance.
3. Implement a multi-class SVM using various feature extraction methods.
4. Analyze and compare the performance of the three models.

Materials and Methods

A tailor-made Convolutional Neural Network (CNN) is developed for art classification, comprising multiple convolutional layers, activation functions, and a fully connected layer. Leveraging the pre-trained ResNet50 model, fine-tuning is performed to adapt to the specific art classification task. Additionally, a Support Vector Machine (SVM) classifier is implemented, utilizing various feature extraction techniques such as SIFT, HOG, ORB, and Gabor filters. The dataset is preprocessed to include only artists with at least 500 paintings and the artist names are encoded. The data is then split into training and validation sets for model training.

Mathematical Formulation of Models

Custom CNN

A Convolutional Neural Network (CNN) consists of convolutional layers, activation functions, pooling layers, and fully connected layers.

- **Convolutional Layer**: Applies a convolution operation to the input, passing the result to the next layer. Mathematically, the convolution is expressed as:

$$f * g(x, y) = \sum_{m} \sum_{n} f(m, n) \cdot g(x \cdot s - m + p, y \cdot s - n + p)$$

- **Activation Function**: Introduces non-linear properties to the system. Commonly used is the ReLU (Rectified Linear Unit) function:

$$f(x) = \max(0, x)$$

- **Pooling Layer**: Reduces the spatial size of the representation, commonly using max pooling:

$$f(x, y) = \max_{(m,n)\in W} g(x + m, y + n)$$

where W is the window size.

- **Fully Connected Layer**: Connects every neuron in one layer to every neuron in the next layer, typically used in the final classification step.

Transfer Learning using ResNet50

ResNet (Residual Network) leverages residual connections or "skip connections" that bypass one or more layers. The residual block can be expressed as:

$$\mathbf{F}(\mathbf{x}) = \mathbf{H}(\mathbf{x}) - \mathbf{x}$$

where $\mathbf{F}(\mathbf{x})$ is the residual mapping to be learned, and $\mathbf{H}(\mathbf{x})$ is the desired mapping.

Transfer learning involves utilizing a pre-trained model (e.g., trained on ImageNet) and fine-tuning it for a specific task.

Multi-Class SVM with SIFT, HOG, ORB, and Gabor

A Support Vector Machine (SVM) finds the hyperplane that best separates the classes.

- **Hyperplane Equation**:

$$\mathbf{w} \cdot \mathbf{x} + b = 0$$

where $\mathbf{w}$ is the weight vector, $\mathbf{x}$ is the input vector, and b is the bias.

- **Feature Extraction**:
  - **SIFT** (Scale-Invariant Feature Transform): Detects and describes local features.
  - **HOG** (Histogram of Oriented Gradients): Captures the gradient orientation in localized portions.
  - **ORB** (Oriented FAST and Rotated BRIEF): A fast binary descriptor.
  - **Gabor Filters**: Used for texture representation and discrimination.

Results

The test results of the experiments with the custom CNN, ResNet50, and SVM models are summarized in the table below:

Model	Test Loss	Test Accuracy
CNN	1.2615	59.70%
ResNet	0.9001	72.37%
SVM	N/A	55.67%

Table 1: Comparison of test loss and accuracy for the custom CNN, ResNet50, and SVM models.

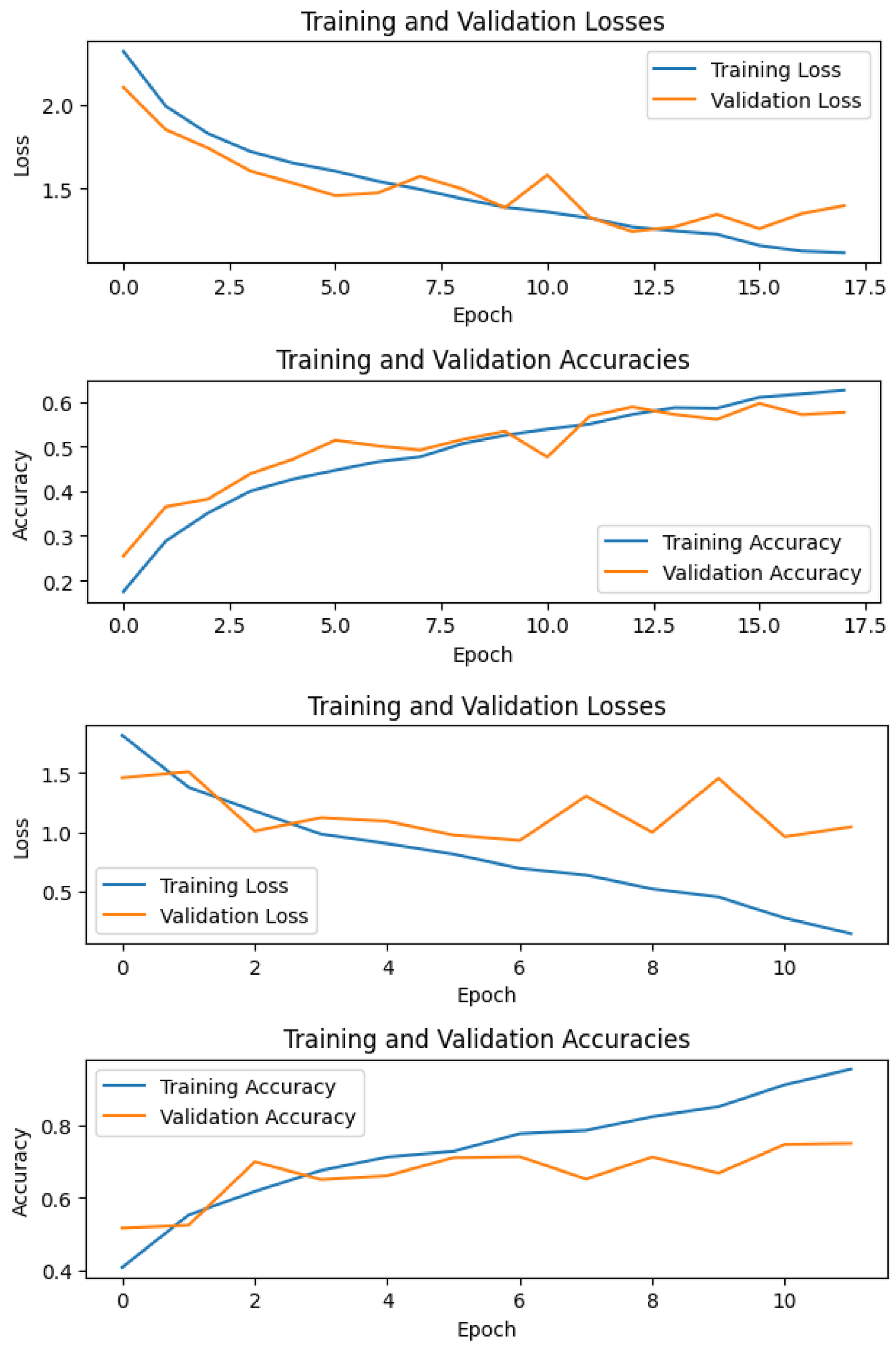

Figure 2: The first two images show the Training and Validation Losses and Accuracies for the custom CNN. The latter two images show the Training and Validation Losses and Accuracies for the ResNet model.

Conclusions

- The study explored the task of art classification through three different approaches: a custom CNN, ResNet50, and a multi-class SVM. Each method provided unique insights into the complexities of classifying artworks based on artists.
- ResNet50, leveraging transfer learning, stood out as the top-performing model, highlighting the effectiveness of pretrained deep learning models in handling complex visual patterns.
- The custom CNN offered a balance between model complexity and performance, showcasing the potential of tailored convolutional networks in art classification.
- The multi-class SVM with multiple feature extractions demonstrated the versatility of traditional machine learning techniques when combined with various image feature extractions.
- The comparative analysis not only underscores the strengths and weaknesses of each approach but also opens avenues for future research. The exploration of Generative Adversarial Networks (GANs) for art forgery detection is particularly promising, which can further broaden the application of machine learning in art authentication and analysis.

Further Research

The future research direction includes exploring the application of Generative Adversarial Networks (GANs) for art forgery authentication. GANs, composed of a generator and a discriminator, could be employed to recognize nuanced characteristics of an artist's style, such as brush strokes and colour palettes. By training on authentic art pieces, the network might discern genuine works from forgeries. Though promising, this approach presents challenges, such as the potential misuse of the GAN to create convincing fakes. Integrating GANs with existing models like CNN, ResNet50, and SVM could result in a comprehensive and more robust art authentication system.

References

[1] Alexander Blessing. Using machine learning for identification of art paintings. 2010.

[2] Ahmed Elgammal, Yan Kang, and Milko Den Leeuw. Picasso, matisse, or a fake? automated analysis of drawings at the stroke level for attribution and authentication. *arXiv preprint*, arXiv:1711.03536, 2017. Subjects: Image and Video Processing (eess.IV); Artificial Intelligence (cs.AI); Computer Vision and Pattern Recognition (cs.CV).