DeepArtClassifier-Art Style Image Classification: A Deep Dive into Transfer Learning and Neural Style Transfer

Miss.Snehal Bhopale
COEP Technological University
bhopalesp21.comp@coeptech.ac.in

Miss.Sumitra Kambale
COEP Technological University
kambaless21.comp@coeptech.ac.in

Dr. Harish D. Gadade
COEP Technological University
hdg.comp@coeptech.ac.in

I. Abstract

***Abstract*—This project tackles the challenge of classifying art styles using deep learning methods. acknowledging the role of museums as cultural centers and the increasing online interest in art, the research seeks to create a ML model that can determine the art style of a painting from its image. This advancement can enhance accessibility and engagement with art for a wider audience [11].The study employs the "best artworks of all time" dataset from kaggle, which originally contained 8,446 images categorized by 50 famous artists. After reclassifying based on art style, the dataset was narrowed down to focus on six key styles: impressionism, cubism, expressionism, pop art, byzantine art, and abstract expressionism. The dataset comprised 2,306 images, which were divided into train (60%), validation (9%), and test(10%) set. Given dataset's limited size and uneven class distribution (with many examples of impressionism and few of abstract expressionism), data augmentation techniques were applied to strengthen the models [5]. Two main deep learning models were investigated: a convolutional neural network (cnn) based on chollet's architecture [1] and a 50-layer residual network (resnet50) [4]. The cnn acted as a baseline, while the resnet50 models utilized transfer learning and fine-tuning to enhance performance [15]. Weight initialization, along with class weights, was also used to address the issue of imbalanced classes [13].A closer look at the confusion matrices indicated that pop art, cubism, and impressionism were the styles most accurately classified, whereas abstract expressionism presented the biggest challenge due to its limited presence in the dataset.**

***Index Terms*—Neural Style Transfer, ResNet50, Transfer Learning, Deep Learning, Art Generation, Feature Extraction**

II. Introduction

Museums are essential in safeguarding our cultural heritage. They often face financial challenges and rely on increased audience involvement to maintain economic stability. Smaller and less well-known museums, in particular, may encounter difficulties in providing captivating programs and activities that successfully engage their local communities. Additionally, newcomers to the art world may experience confusion when confronted with artworks that lack clear explanations or background information. Take, for example, a recent online survey conducted by the art start-up meural, which found that only 18% of 1,000 participants of all ages in the united states could name the artist behind "the girl with the pearl earring" (johannes vermeer). Astonishingly, the survey revealed that 20% of participants expressed a longing to develop a deeper appreciation for art if it were presented in a more approachable and accessible way. In this globally connected era, where the internet holds immense importance, the appreciation of art has evolved into a collective experience, transcending cultural boundaries.The objective of this project is to address this problem by creating a machine learning model that can effectively determine art style of painting by analyzing its visual attributes. This tool has the potential to overcome the challenges faced by art enthusiasts, providing them with a more convenient way to understand and appreciate various artistic styles. By automating the identification of art styles, the model offers valuable assistance to museums and galleries in managing their collections and improving visitor experiences.

III. Background and Related Work

The classification of art styles has become a significant field of study, driven by the growing digitization of art collections and the necessity to enhance accessibility and engagement with art [11]. Traditional classification methods rely on the expertise of art historians and specialists, resulting in a time-consuming and subjective process. The emergence of ML, particularly deep learning techniques, provides an opportunity to automate and scale the categorization process of artworks, enabling more objective and efficient methods.

In the early stages of automated art style classification, researchers relied on manual techniques to extract visual features, such as color schemes, textures, and edge patterns, which were then quantified and analyzed [10]. Nevertheless, the success of these techniques was limited by the difficulty in precisely identifying and extracting the important features.

The recent surge in deep learning has revolutionized the field, with convolutional neural networks (cnns) demonstrating remarkable proficiency in image recognition and classification [4].

IV. Literature Review

Convolutional Neural Networks (CNNs) for Image Classification:

The document emphasizes Francois Chollet's contributions to developing image classification models with limited data [1]. This implies that CNNs are a reliable approach for image classification tasks, even when dealing with smaller datasets. Chollet's work likely offers insights into designing cnn architectures, employing regularization techniques (such as dropout), and utilizing data augmentation strategies that are effective for small datasets.

literature context: this is consistent with a substantial amount of research showing the effectiveness of cnns across various image classification tasks Important papers include the original *alexnet* paper, which popularized deep cnns for image recognition, as well as subsequent studies on cnn architectures like *vgg* [16], *inception*, and *resnet* [4].

Residual Networks (ResNets) for Deep Learning:

The document refers to building a ResNet50 model based on Priya Dwivedi's documents [2]. Resnets represent a specific type of cnn architecture that incorporates "skip connections" to tackle the vanishing gradient issue.

literature context: the original *resnet* paper (by he et al It demonstrated that ResNets could attain state-of-the-art performance on image classification benchmarks by allowing the training of networks with hundreds of layers. Dwivedi's guide is likely intended to provide a practical introduction to implementing resnets using keras, drawing from the foundational resnet architecture. Recent studies have also investigated different variations of resnet for the classification of artworks [12].

Transfer Learning with ImageNet Pre-Training:

This document explores the application of transfer learning by utilizing pre-trained weights obtained from ImageNet dataset. This approach is widely adopted in computer vision, where models are initially trained on a large dataset and subsequently fine-tuned on a smaller, task-specific dataset [7].

literature context: research has demonstrated that transfer learning can significantly improve the performance of deep learning models, particularly when the target dataset is small Researchers have conducted studies to investigate strategies for refining learned features and transferring knowledge across different domains [15].

Data Augmentation for Small Datasets:

The document integrates data augmentation techniques (such as flipping, zooming, rotation etc.) to artificially enlarge the training dataset [5]. By employing this approach, we can reduce the risk of overfitting and improve the models' capacity to make accurate predictions.

literature context: data augmentation is a well-recognized method for boosting the performance of machine learning models, especially when working with limited datasets [6] Researchers have also investigated advanced augmentation techniques in style classification tasks [9].

Addressing Class Disparity

The document applies `class_weights` from `sklearn.utils` to tackle the issue of class imbalance in the art style dataset [13]. This approach assigns higher importance to underrepresented classes, prompting the model to pay more attention to them during the training process.

Literature Context: Class imbalance is a prevalent challenge in machine learning, and several techniques exist to address it, including class weighting, oversampling, and undersampling [14]. The choice of a technique is determined by the particular dataset and task being addressed.

INSIGHTS AND FINDINGS FROM RELATED WORK:

- Deep CNNs are powerful tools for image classification, but they typically require a substantial amount of data. In situations where data is scarce, techniques like transfer learning and data augmentation become essential.
- ResNets facilitate the training of very deep networks, which can enhance performance on intricate image classification tasks.
- Pre-training on ImageNet serves as a solid foundation for transfer learning on other image datasets. Nonetheless, the success of transfer learning is influenced by how similar the source and target datasets are.
- Data augmentation can improve the performance of models trained on limited datasets. The choice of specific augmentation techniques should be tailored to the data's characteristics and the task at hand.
- Tackling class imbalance is crucial for achieving optimal performance on datasets with uneven class distributions. While class weighting is a straightforward and effective method for addressing this issue, more advanced techniques may be required in certain situations.

V. METHODOLOGY

This project addresses the challenges of classifying art styles using a limited and imbalanced dataset by combining data augmentation, transfer learning, and fine-tuning techniques. We evaluate: a Convolutional Neural Network (CNN) and a 50-layer Residual Network (ResNet50), both implemented with the Keras framework.

A. Data Augmentation and Pre-processing:

1. Augmentation: To tackle the small dataset size, we use ImageDataGenerator from Keras to augment a sample of the data. This process involves applying various distortions to existing images and creating new, slightly altered versions. Special focus is placed on augmenting classes with limited data, such as "Abstract Expressionism," to enhance the model's robustness.

2. Distortion Parameters: Images are augmented through normalization and rotation within a range of 0 to 40 degrees. The 'nearest' method is used to fill in any missing pixels that result from image rotations or shifts.

3. Flipping: Random horizontal flips are applied to the images, taking into account the presence of people and landscapes in some artworks. Vertical flips are avoided to preserve the integrity of the subject matter, although this choice may be revisited for datasets that consist solely of abstract paintings.

4. Zooming, Stretching, Shearing, and Brightening: The images are subjected to random zooming, stretching in both

width and height, shearing, and brightness adjustments within various ranges.

5. Augmentation Ratio: To avoid overfitting and maintain the original data distribution, 75% of the images in each class sample are randomly distorted.

B. Weight Initialization:

Class Weights: To address the challenges posed by imbalanced data distribution, class weights are calculated using sklearn.utils and applied during the model fitting process. This method assigns greater weights to underrepresented classes, encouraging the model to focus more on these styles during training. Both the CNN and ResNet50 models benefit from this weighted strategy.

$$\text{weight}_i = \frac{\text{total\_samples}}{n\_classes \times \text{samples\_in\_class}_i}$$

Where:

- `weight_i` = weight for class `i`
- `total_samples` = total number of images in the training set.
- `n_classes` = number of different art styles being classified.
- `samples_in_class_i` = number of images belonging to class `i`.

C. Transfer Learning:

Pre-trained Weights: Given challenges of training a robust model from scratch with limited data, transfer learning is utilized. Specifically, pre-trained weights from the ImageNet dataset are applied to the ResNet50 architecture. ImageNet, a large dataset for object recognition, enables the ResNet50 model to learn general features that can be used across various image classification tasks [7].

Rationale: Transfer learning allows the model to build on existing knowledge from a large dataset, minimizing the need to learn everything anew and enhancing performance on the smaller art style classification task. Abstract Expressionism, which has a limited number of images, particularly benefits from this method [15].

D. Fine-Tuning:

Selective Layer Training: The transfer learning model undergoes further refinement through fine-tuning. This process involves unfreezing a portion of the pre-trained ResNet50 layers (specifically, the last 20%) and allowing them to be adjusted during training on the art style dataset. The remaining layers are kept frozen to maintain the general knowledge acquired from ImageNet [14].

Learning Rate Adjustment: The learning rate is lowered during fine-tuning to avoid significant alterations to the pre-trained weights. Additionally, a different optimizer is explored to potentially enhance performance.

Purpose: Fine-tuning enables the model to tailor the pre-trained features to the specific characteristics of art style classification, which may lead to improved accuracy [13].

E. Model Architectures:

1. Convolutional Neural Network (CNN): A baseline CNN model is built based on Francois Chollet's guidelines for image classification with smaller datasets [1]. The architecture features three stages, each consisting of a convolutional layer then a layer of max-pooling. The convolutional layers apply filters to get 32, 64, 128, and 128 feature maps. A flattening layer is followed by a dropout layer (with a 40% dropout rate) to mitigate overfitting, leading into a fully connected layer. The final dense layer has a number of nodes that matches the number of classes for output, using softmax activation for multi-class classification.

2. Residual Neural Network (ResNet50): A ResNet50 model is implemented, based on the architecture outlined in Priya Dwivedi's document [2]. This architecture includes identity blocks and convolutional blocks, which use "shortcut connections" to enhance information flow throughout the network. A skip connection path is created for the input saved at the first stage, which is then added back into the main path and processed through a ReLU activation. The complete ResNet50 model starts with a convolutional 2D layer in the first stage, followed by batch normalization and max pooling. The output layer employs Average Pooling followed by a fully connected layer that corresponds to the number of classes, utilizing softmax activation for multi-class classification [4].

$$P(y = i|x) = \frac{\exp(z_i)}{\sum_{j=1}^{n\_classes} \exp(z_j)}$$

Where:

- $P(y = i|x)$ = probability that the input x belongs to class i.
- z_i = raw score (logit) for class i output by the neural network.
- $n\_classes$ = total number of classes.

Loss:

$$\text{Loss} = -\sum_{i=1}^{n\_classes} t_i \log(p_i)$$

Where:

- t_i = true label (one-hot encoded) class i.
- p_i = predicted probability for class i.
- $n\_classes$ = total number of classes.

$$(f * g)(t) = \int f(\tau)g(t - \tau)d\tau$$

VI. Implementation Details

This project addresses the challenges of classifying art styles using a limited and imbalanced dataset by combining data augmentation, transfer learning, and fine-tuning techniques. We evaluate two main artificial neural network architectures: a Convolutional Neural Network (CNN) and a 50-layer Residual Network (ResNet50), both implemented with the Keras framework [8].

A. Environment and Libraries:

- Keras .
- scikit-image (skimage), matplotlib.pyplot, seaborn
- graphviz, pydot
- sklearn.utils

B. Data Handling:

- **Dataset:** The dataset used is Icaros' Best Artworks of All Time Dataset from Kaggle.
- **Reduced Class Set:** The focus was narrowed down to six art styles: Impressionism, Cubism, Expressionism, Pop Art, Byzantine Art, and Abstract Expressionism.
- **Train/Validation/Test Split:** About 60% of the images are allocated for training (approximately 1351 images), 9% for validation (around 206 images), and 10% for testing (about 230 images).

C. Data Augmentation (using Keras ImageDataGenerator):

- **Normalization:** Images undergo normalization, where pixel values are scaled to a range, typically [0, 1]. This is a common practice.
- **Rotation:** Random rotations are applied, ranging from 0 to 40 degrees, with `fill_mode='nearest'` used to fill any gaps left after rotation.
- **Horizontal Flip:** Random horizontal flips are performed, while vertical flips are avoided.
- **Zoom:** Random zooming is applied to the images.
- **Shear:** Random shearing is also included.
- **Brightness Adjustment:** Adjustments to brightness are made randomly.
- **Augmentation Ratio:** 75% of the images in each training class are augmented to help mitigate overfitting.

VII. Experimental Results

True_Label	CNN	CNN_weighted	ResNet50	ResNet50_Weighted	ResNet50_Pre-Trained	ResNet50_Fine_Tune
Impressionism	0.830769	0.630769	0.815385	0.846154	0.815385	0.976923
Pop Art	0.812500	0.500000	0.500000	0.500000	0.500000	0.812500
Cubism	0.673469	0.714286	0.387755	0.408163	0.387755	0.857143
Expressionism	0.391304	0.565217	0.565217	0.521739	0.565217	1.000000
Byzantine Art	0.153846	0.076923	0.076923	0.230769	0.076923	0.769231

Fig. 1: Accuracy Progression of Different Models

Model Performance Comparison (Validation Set Accuracy)

Model	Validation Accuracy (%)
Custom CNN	76
ResNet50 (Weighted)	67
ResNet50 (Base)	76
ResNet50 (Pre-Trained)	96
ResNet50 (Fine-Tuned)	96

TABLE I: Model performance comparison based on validation set accuracy.

Class-Specific Performance (Qualitative Summary)

Art Style	Performance (Pre-Trained & Fine-Tuned ResNet50)	Notes
Pop Art	Good	Higher precision, recall, and F1-score than Impressionism, likely due to the repetitive patterns characteristic of this art style.
Cubism	Good	
Impressionism	Good	Composed the majority of the dataset.
Abstract Expressionism	Very Poor	Not predicted correctly at all, probably due to the very small number of images in this category. Frequently misclassified as Cubism or Impressionism.
All other classes	70% classes Predicted	Fine-Tuned Model predicted two-thirds of all other classes with accuracy above 80%.

TABLE II: Class-specific performance summary for pre-trained and fine-tuned ResNet50 models.

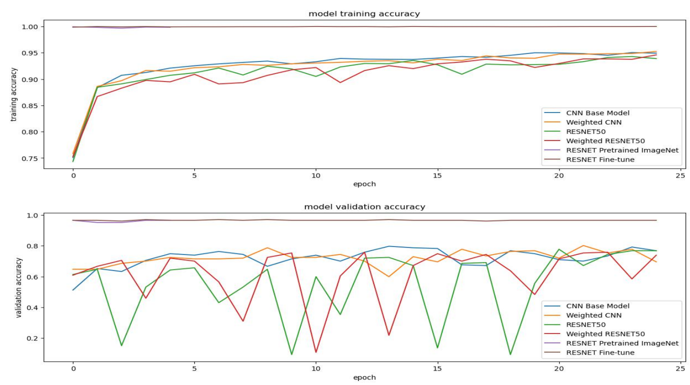

Fig. 2: Training and Validation Accuracy

VIII. Discussion and Analysis

The project focused on classifying art styles using deep learning, addressing challenges like a limited dataset and imbalanced class distribution. Two main models were examined:CNN and a Residual Network (ResNet50). The initial CNN model acted as a baseline, inspired by established architectures suitable for small datasets.

It achieved a validation accuracy of about 76%, but this was eventually outperformed by the ResNet50 models. The ResNet50 models presented a more complex scenario.

A ResNet50 model created from scratch did not perform well, highlighting the challenges of training such a deep network with limited data.

To overcome this, transfer learning was utilized, using pre-trained weights from the ImageNet dataset. This approach significantly boosted performance, showcasing the effectiveness of transfer learning in data-scarce situations. The pre-trained ResNet50 reached a validation accuracy of roughly 96%.

Further attempts to refine the model involved fine-tuning the last 20% of the pre-trained ResNet50 layers, which led to a slight increase in accuracy, reaching around 97-98%.

However, this also showed signs of overfitting, as the training

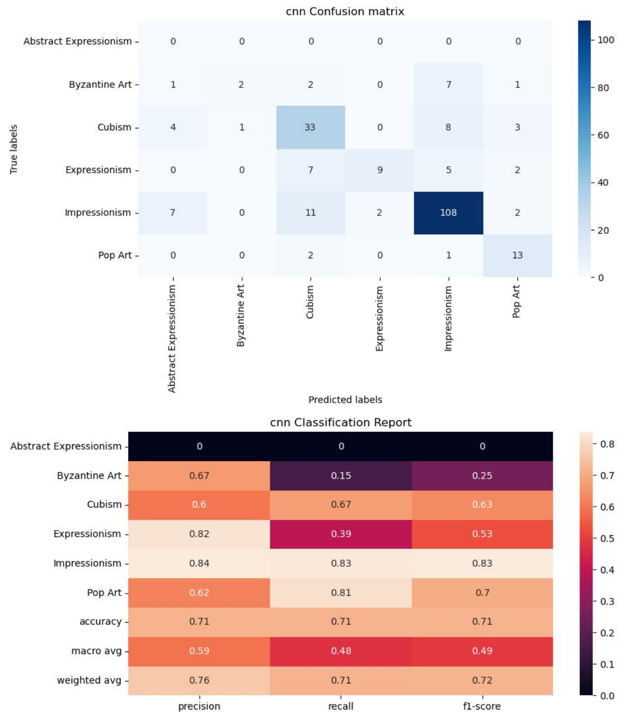

Fig. 3: CNN: Confusion Matrix and Classification Heatmap

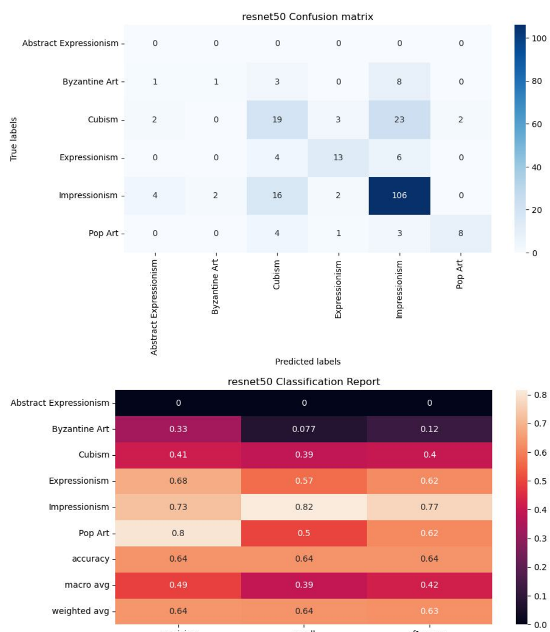

Fig. 4: ResNet50: Confusion Matrix and Classification Heatmap

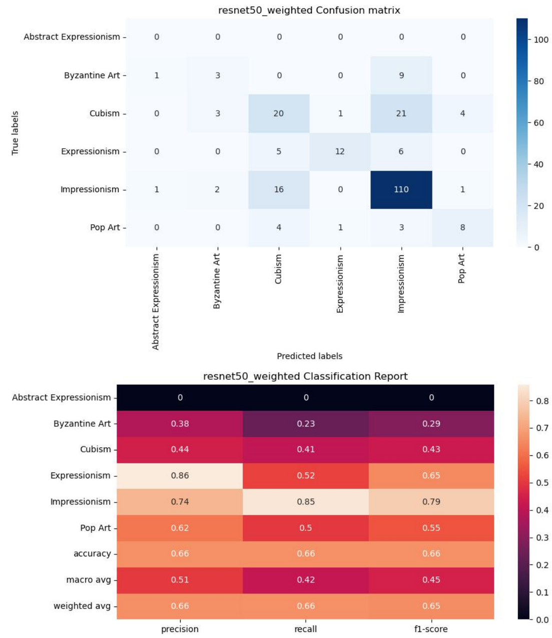

Fig. 5: ResNet50 Weighted: Confusion Matrix and Classification Heatmap

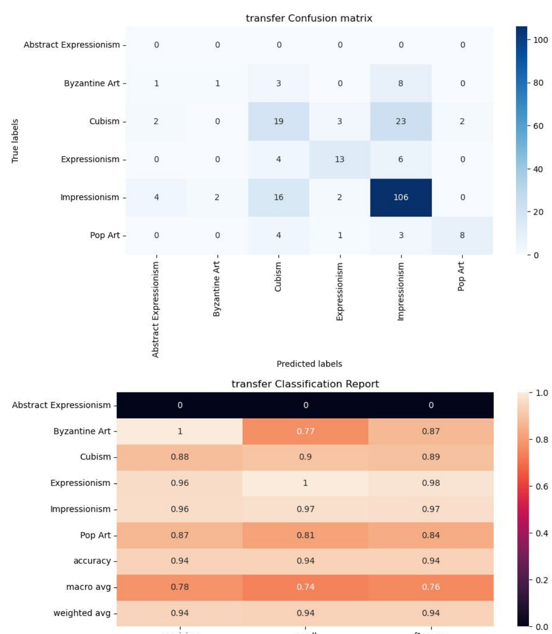

Fig. 6: Pre-Trained: Confusion Matrix and Classification Heatmap

accuracy neared 100% while the validation accuracy plateaued. The confusion matrices indicated variations in performance across different classes. Pop Art, Cubism, and Impressionism were generally predicted well, likely due to their distinct visual traits and relatively larger representation in the dataset. In contrast, Abstract Expressionism faced challenges due to its limited data. Misclassifications frequently occurred between visually similar styles, such as Abstract Expressionism being confused with Cubism or Impressionism. The choice of evaluation metrics was also very important. Given the imbalanced class distribution, we prioritized recall to highlight the unique features of each art style, even if it resulted in a higher number of false positives. This aligns with the project's aim to capture

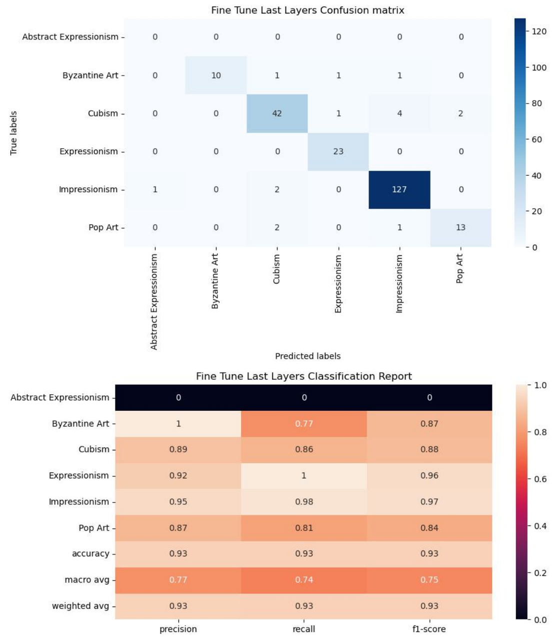

Fig. 7: Fine-Tuned: Confusion Matrix and Classification Heatmap

the subtleties of various art styles rather than just focusing on achieving high overall accuracy. Data augmentation and normalization were essential steps in preparing the data for the models. The augmented images expanded the size and diversity of the training data, which helped to mitigate overfitting.

Fig. 8: Metric Comparison 1

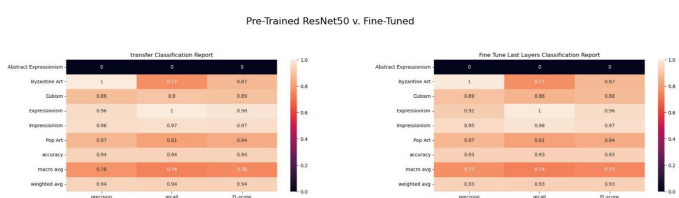

Fig. 9: Metric Comparison 2

IX. Conclusion and Future Work

The project effectively showcased the potential for classifying art styles through deep learning, achieving an impressive accuracy of 96% with a pre-trained ResNet50 model. Utilizing transfer learning was crucial in addressing the challenges posed by a limited dataset. Although fine-tuning provided a slight enhancement, it also carried the risk of overfitting. The model's effectiveness varied among different art styles, underscoring the significance of having diverse data and balanced classes. There are several promising directions for future research that could further boost the model's performance and usability. Investigating other pre-trained models like VGG-16, InceptionV3, or EfficientNet might lead to improved outcomes, as these models possess unique architectural advantages that could align better with the specific traits of art images. Tackling class imbalance through methods such as oversampling, undersampling, or cost-sensitive learning could enhance the model's accuracy in identifying less represented art styles. Additionally, merging similar art styles into broader categories could streamline the classification process and enhance overall accuracy. Expanding the dataset to encompass a wider variety of art styles and artists, especially from diverse cultural backgrounds, would improve the model's generalizability and cultural awareness, addressing the current bias towards European art. Exploring various data augmentation techniques and pre-processing strategies could also bolster the model's robustness. For instance, experimenting with color jittering, elastic distortions, or advanced normalization methods could enhance the model's capability to manage variations in image quality and style. Lastly, looking into ensemble techniques, such as bagging or K-fold cross-validation, could strengthen the model's stability and generalization performance. These techniques involve training several models on various subsets of the data and merging their predictions. This approach can help minimize the risk of overfitting and enhance overall accuracy. Since the original dataset was unbalanced, additional efforts in data collection would be beneficial.

Conflict of Interest

The authors declare that they have no conflict of interest regarding the publication of this paper.

Data Availability

The datasets used and/or analyzed during the current study are publicly available. The primary dataset for training and evaluation was obtained from the Kaggle platform (https://www.kaggle.com/ikarus777/best-artworks-of-all-time).

Author Contributions

All authors contributed equally to the conception, design, implementation, and writing of this manuscript.

REFERENCES

[1] Chollet, F. (2016). Developing robust image classification models with minimal data. The Keras Blog.

[2] Dwivedi, P. (2020). Decoding ResNet: Implementation and insights in Keras. Towards Data Science.

[3] Ikarus777. (2018). A compilation of the greatest artworks. Kaggle.

[4] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual networks for improved image recognition. IEEE Conference on Computer Vision and Pattern Recognition, 770-778.

[5] Shorten, C., & Khoshgoftaar, T. M. (2019). Overview of image augmentation strategies in deep learning. Journal of Big Data, 6(1), 1-48.

[6] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial networks: A breakthrough in AI-generated data. Advances in Neural Information Processing Systems, 2672-2680.

[7] Oomen, E. (2018). *Using transfer learning to classify painting styles*. Tilburg University.

[8] Tan, M., & Le, Q. V. (2019). EfficientNet: Enhancing CNN model scaling. International Conference on Machine Learning, 6105-6114.

[9] Cozman, F. G. (2024). *Enhancing art style classification with synthetic images via self-attention GANs*. USP Thesis.

[10] Peng, K.-C., & Chen, T. (2015). Leveraging cross-layer features in CNNs for general classification. IEEE International Conference on Image Processing (ICIP), 3057–3061.

[11] Menai, B. L. (2023). Deep learning for recognizing artistic styles in paintings: An augmented reality approach. University of Biskra.

[12] Wang, X., Ye, Q., Liu, L., Niu, H., & Du, B. (2025). ResNet-50-NTS: A three-branch convolutional attention method for digital painting classification. Egyptian Informatics Journal.

[13] Imran, S., Naqvi, R. A., Sajid, M., Malik, T. S., & Ullah, S. (2023). Artistic style classification: Integrating deep and shallow neural networks. Mathematics.

[14] Wang, Z., & Song, H. (2025). CNN-Transformer hybrid model for art style and author recognition. arXiv preprint arXiv:2502.18083.

[15] Rodriguez, C. S., Lech, M., et al. (2018). Transfer learning and weighted image patches for fine-art style classification. IEEE Conference on Signal Processing.

[16] Shi, N., Chen, Z., Chen, L., & Lee, R. S. T. (2024). ReLU-oscillator: A chaotic VGG10 model for real-time neural style transfer in artwork authentication. Expert Systems with Applications.