Using CNNs to Classify Paintings: A First Step Towards Perfecting Art Genres

Muhammad Ammar Haider Baloch
National University of Computer and Emerging Sciences
Islamabad, Pakistan
ammarbaloch2015@gmail.com

***Abstract*—Art genre classification is a challenging task in the field of computer vision. This paper explores the use of Convolutional Neural Networks (CNNs) for classifying paintings into six distinct art genres: 'Abstract,' 'Baroque,' 'Cubism,' 'Digital Art,' 'Impressionism,' and 'Abstract Expressionism.' We present two different CNN architectures and evaluate their performance on a dataset of paintings. Our findings demonstrate the potential of CNNs in automating art genre classification tasks, revealing insights into the inherent complexities and nuances in the fusion of art and technology.**

I. Introduction

Art genre classification presents a unique challenge in computer vision, blending the intricacies of artistic expression with advanced computational techniques. Paintings, as a medium of visual art, exhibit a range of styles and characteristics that define their genres. This research delves into the realm of automated classification of these genres, employing Convolutional Neural Networks (CNNs) — a class of deep, feed-forward artificial neural networks known for their prowess in image recognition and classification tasks. By focusing on 'Abstract,' 'Baroque,' 'Cubism,' 'Digital Art,' 'Impressionism,' and 'Abstract Expressionism,' this study aims to explore and quantify the efficacy of CNNs in identifying and categorizing these diverse art styles, thereby contributing to the broader field of computational art analysis.

II. Related Work

The application of CNNs in image classification tasks has been a topic of extensive research, given their ability to perform feature extraction and pattern recognition in complex datasets. In the context of art genre classification, these deep learning models have shown promise in discerning artistic styles by analyzing visual elements. Prior studies have focused on developing robust CNN architectures capable of handling the high dimensionality and variability inherent in art images. Notable works include Zhao et al. (2021), who investigated the performance of various CNN models in art classification across multiple genres and styles, emphasizing the significance of optimizing model architecture for improved accuracy. Similarly, Lee and Cha (2016) explored the use of self-organizing maps, another neural network technique, for categorizing and visualizing art genres, thus demonstrating the interdisciplinary nature of this research domain.

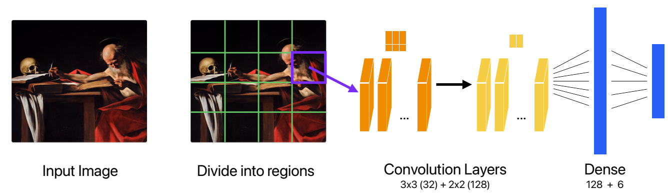

Fig. 1. Architecture 1

III. Data Collection

The dataset, curated specifically for this study, encompasses an array of 360 paintings, meticulously gathered from authoritative and verifiable sources. Each of the six genres — 'Abstract,' 'Baroque,' 'Cubism,' 'Digital Art,' 'Impressionism,' and 'Abstract Expressionism' — is represented by 60 paintings, ensuring a balanced distribution for unbiased training and evaluation. The dataset's construction was guided by the principles of diversity and representativeness, aiming to encompass a wide range of artistic expressions within each genre.

IV. Methodology

Our methodology introduces two novel CNN architectures designed to capture the multifaceted nature of art genre classification. These architectures are engineered to analyze and process the unique features of paintings, such as texture, color distribution, and structural elements.

A. Architecture 1

Architecture 1 innovatively segments each painting into a 4×4 grid, treating each cell as a distinct input for localized feature extraction. Employing a compact CNN within each cell, consisting of 6 convolutional layers and subsequent max-pooling, this architecture captures fine-grained details. The individual outputs are then concatenated to form a comprehensive feature vector, encapsulating a holistic view of the painting. This vector undergoes further processing through a series of dense layers, culminating in the final classification. The architecture is visualized in Figure 1.

B. Architecture 2

Architecture 2 takes a preprocessing-centric approach, transforming each painting into a grayscale image for edge detection, while simultaneously extracting prominent colors via

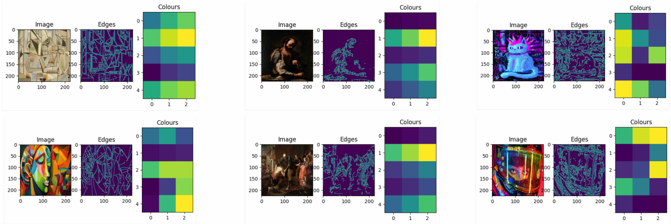

Fig. 2. Architecture 2 - Preprocessing

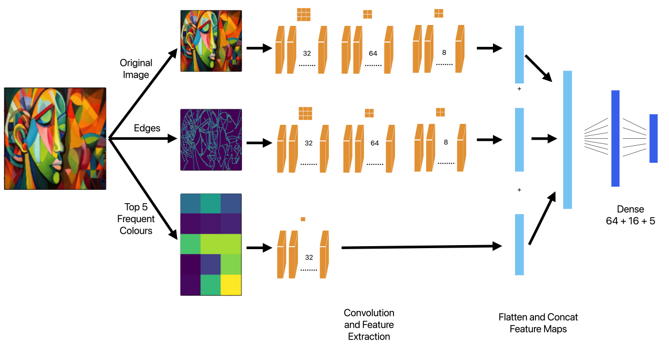

Fig. 3. Architecture 2

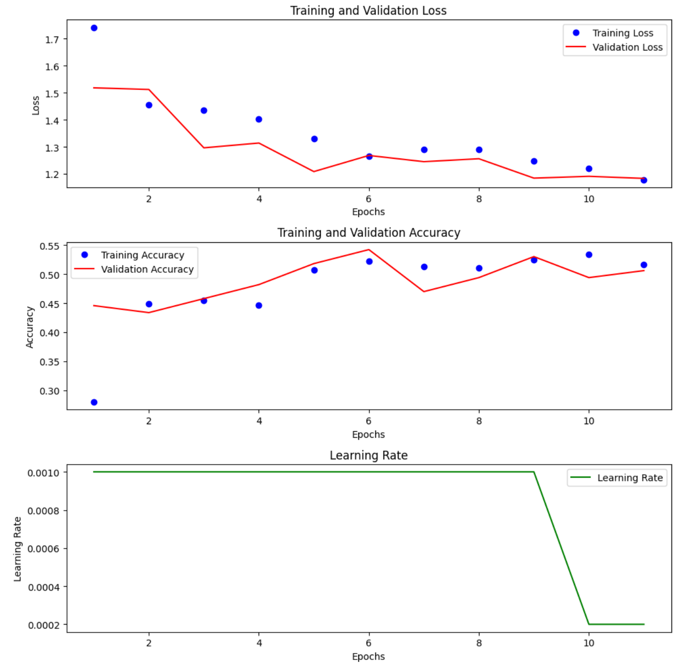

Fig. 4. Outcomes of Training Architecture 1

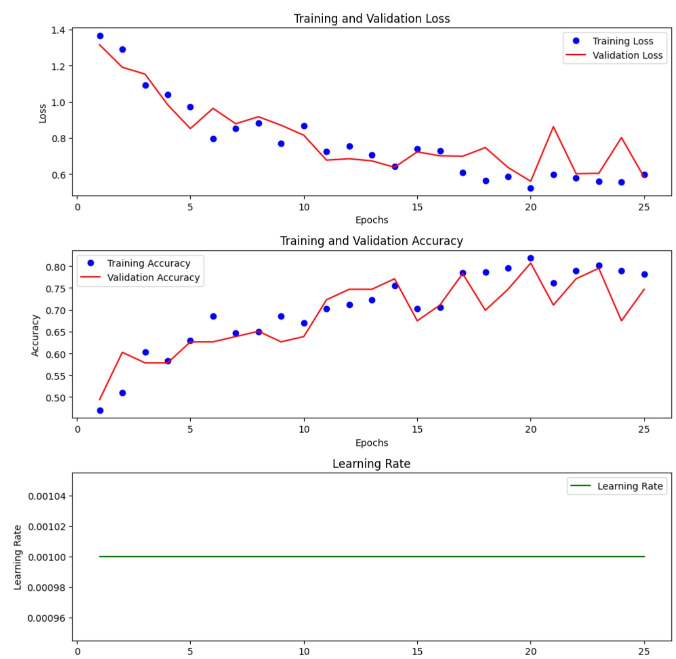

Fig. 5. Outcomes of Training Architecture 2

K-Means clustering. These preprocessed elements serve as inputs to three dedicated CNNs, each specializing in a particular aspect of the painting. The individual CNN outputs are merged, creating a rich feature amalgamation that is then funneled through dense layers for final classification. This architecture highlights the significance of preprocessing in enhancing the model's sensitivity to crucial artistic attributes. The architecture is visualized in Figure 3.

V. Experiments and Results

Experiments were conducted to evaluate the performance of the proposed architectures. The dataset was partitioned into training, validation, and test sets, with a focus on implementing comprehensive data augmentation strategies to simulate a wide range of artistic variations. The training process utilized the Adam optimizer, known for its efficiency in handling sparse gradients on noisy problems, and categorical cross-entropy as the loss function, a standard choice for multi-class classification tasks. The models' training was further refined through callbacks including ModelCheckpoint, EarlyStopping, and ReduceLROnPlateau, ensuring optimal performance and convergence. The experimental setup leveraged Google Colab's T4 GPU, providing the necessary computational power for efficient model training and evaluation. Architecture 1, Figure 4 shows the training outcomes. For Architecture 2, Figure 5 shows the training outcomes. For Architecture 1, we can see that the model very quickly achieves maximum accuracy and does not improve on it. For Architecture 2, we can see that the model slowly achieves a maximum and then wanders around that maximum.

A. Results

B. Results

The experimental results revealed that both architectures achieved moderate success, with Architecture 1 attaining a validation accuracy of 55% and Architecture 2 reaching 78%. These outcomes underscore the potential of CNNs in art genre classification, yet also highlight the challenges in capturing the full spectrum of artistic diversity. Factors such as dataset limitations, model complexity, and the inherent subjectivity of art interpretation were identified as potential areas for improvement.

VI. Discussion

The study's findings open up avenues for further exploration in the intersection of computer vision and art analysis. The moderate success of the CNN architectures in classifying art genres points to the complexity of this task and the need for more sophisticated models and larger, more diverse datasets. The exploration of hyperparameter tuning, advanced architectural designs, and the integration of additional art-related features could enhance the models' performance. Furthermore, the interpretability of these models and their ability to provide meaningful insights into the artistic styles and elements warrants additional research, potentially leading to new methodologies in art analysis and curation.

VII. Model Implementation Prospects

Despite the limitations in genre classification, the developed model exhibits a notable proficiency in distinguishing between Digital and Non-Digital Art. This suggests potential applications in digital art authentication, curation, and archival, where rapid and accurate classification can aid in managing large digital art repositories. Furthermore, the model's ability to discern digital art styles opens up possibilities for its use in educational and research settings, providing a tool for studying and understanding the evolving landscape of digital art.

VIII. Conclusion

This paper presents a foray into the use of Convolutional Neural Networks for classifying paintings into distinct art genres. The proposed CNN architectures, while demonstrating the potential of deep learning in this domain, also highlight the challenges and complexities inherent in the fusion of art and technology. The results and insights gained from this study contribute to the burgeoning field of computational art analysis, paving the way for further advancements and explorations at the nexus of computer vision and art.

References

[1] Zhao, W., Zhou, D., Qiu, X., Jiang, W. (2021). Compare the performance of the models in art classification. PLoS ONE 16(3): e0248414.

[2] Lee, S.G., Cha, E.Y. (2016). Style classification and visualization of art painting's genre using self-organizing maps. Hum. Cent. Comput. Inf. Sci. 6, 7.