Style Transfer and Contrastive Learning for Image Classification

Aybike Battal
Department of Artificial Intelligence and Data Engineering
Istanbul Technical University
Istanbul, Turkey
Student No: 150200343
Email: battal20@itu.edu.tr

Nazrin Abdinli
Department of Artificial Intelligence and Data Engineering
Istanbul Technical University
Istanbul, Turkey
Student No: 150220925
Email: abdinlin22@itu.edu.tr

Ilayda Kara
Department of Artificial Intelligence and Data Engineering
Istanbul Technical University
Istanbul, Turkey
Student No: 150220747
Email: karai22@itu.edu.tr

***Abstract*—This project explores the application of style transfer and contrastive learning to improve image classification. The pre-trained deep learning models are leveraged, including VGG19 for style transfer and InceptionV3 for feature extraction. The method involves transferring the style from a source image to a content image and using contrastive learning to enhance the classification accuracy of stylized images. The data for training and testing includes images from publicly available datasets. The high classification accuracy is achieved, demonstrating the effectiveness of the approach.**

I. Introduction

In recent years, there have been big improvements in using deep learning for image classification. But, older models have a hard time dealing with different styles and looks in images. This project tries to solve this problem by using two techniques: neural style transfer (NST) and contrastive learning. These methods aim to make image classification more accurate and reliable.

II. Problem Statement and Hypothesis

A. Problem Statement

The objective of this project is to enhance image classification performance through style transfer and contrastive learning. Traditional image classification models often struggle with variations in style and appearance. By applying style transfer, it is aimed to create more diverse training data, and through contrastive learning, and seek to improve the model's ability to distinguish between different image styles.

1) Hypothesis: It is hypothesized that style transfer can generate varied training data that, when combined with contrastive learning, will improve the robustness and accuracy of image classification models.

2) Literature Survey: Recent advancements in self-supervised learning, such as SimCLR, MoCo, and BYOL, have demonstrated the efficacy of contrastive learning in obtaining meaningful representations from unlabeled data.
SimCLR emphasizes the importance of data augmentation and contrastive learning to acquire robust representations by maximizing agreement between different views of the same image while minimizing agreement between views of different images.
MoCo introduces a momentum encoder to refine representations, achieving state-of-the-art performance. Neural Style Transfer (NST) techniques, pioneered by Gatys et al., separate content and style representations, enabling the transfer of artistic styles onto different content images.

III. Methods, Data, Results, and Discussion

A. Methods

1) Style Transfer: VGG19 model is employed to extract style and content features from images. The style transfer process involves optimizing a generated image to minimize the difference between its content and a given content image while matching the style of a style image.

2) Preprocessing: Images are resized to 512x512 pixels and preprocessed for the VGG19 model.

3) Feature Extraction: Style and content features are extracted from specific layers of the VGG19 model.

4) Loss Calculation: The total loss is computed as a weighted sum of content and style losses. Content loss measures the difference between the generated and content images, while style loss compares the Gram matrices of the generated and style images.

5) Optimization: The generated image is optimized using gradient descent to minimize the total loss.

6) Contrastive Learning: We use the InceptionV3 model for feature extraction and apply contrastive learning to improve classification accuracy. Contrastive loss encourages the model to distinguish between positive (similar) and negative (dissimilar) image pairs.

7) Feature Extraction: Features are extracted from content and stylized images using the InceptionV3 model.

8) Positive and Negative Pairs: Positive pairs consist of features from the content and stylized images, while negative pairs use features from the same image.

9) Loss Calculation: Contrastive loss is computed for both positive and negative pairs, promoting similarity for positive pairs and dissimilarity for negative pairs.

10) Training a Classifier: A simple linear classifier is trained on the extracted features to classify images as original or stylized.

B. Data

Publicly available images are used from TensorFlow's example images and Kaggle datasets for training and testing. The content images are natural photographs, while the style image is an artistic painting.

C. Results

1) Style Transfer: The style transfer successfully generates images that combine the content of natural photographs with the artistic style of paintings. The generated images exhibit the desired stylistic patterns while retaining the original content structure.

Fig. 1. Content Image

Fig. 2. Style Image

Fig. 3. Output Image

2) Contrastive Learning and Classification: The linear classifier achieves high accuracy in distinguishing between original and stylized images. The training accuracy reaches 100%, with perfect precision, recall, and F1 scores.

Metric	Training Set	Test Set
Accuracy	100%	100
Precision	1.0	1.0
Recall	1.0	1.0
F-1 Score	1.0	1.0

D. Discussion

The findings show that combining style transfer and contrastive learning is effective in making image classification better. Style transfer makes different training data, which helps the classifier understand different styles. Contrastive learning makes the model better at telling original and stylized images apart by focusing on their similarities and differences.

1) Limitations: The style transfer process takes a lot of computer power and time to process each image. Moreover, the method used in the project depends on chosen style and content images, which might make the training data less diverse.

2) Future Work: The project could be tried using different styles to create a broader range of training data. Additionally, contrastive learning could be made faster and capable of handling more data. The method could be tested with larger and more diverse datasets to assess its real-life effectiveness.

IV. Conclusion

The approach of this project demonstrates that combining style transfer and contrastive learning can significantly enhance image classification performance. The style transfer method generates diverse training data, while contrastive learning improves the model's ability to generalize across different styles. The high accuracy and perfect precision, recall, and F1 scores indicate the robustness of our method.

In future work, it is planned to explore the impact of using different style images and further optimize the contrastive learning process and aimed to evaluate the approach on larger and more diverse datasets to validate its effectiveness in real-world scenarios.

By leveraging advanced techniques like style transfer and contrastive learning, the project contributes to the field of image classification and provides a foundation for further research and development.

V. References

1. Chen et al. (2020). "A Simple Framework for Contrastive Learning of Visual Representations" - SimCLR

2. He et al. (2020). "Momentum Contrast for Unsupervised Visual Representation Learning" - MoCo

3. Gatys et al. (2016). - Neural Style Transfer (NST)

4. https://storage.googleapis.com/download.tensorflow.org/example_images/YellowLabradorLooking_new.jpg

5.https://storage.googleapis.com/kaggle-datasets-images/133/275/a6c2e389f50dfb5bbc76705e893a6c10/dataset-card.png

VI. Link of the Source Code of the Project

https://github.com/itu-itis22-abdinlin22/BLG_454E_project/blob/main/blg454e-project-last.ipynb