ISSN 2616-5775 Vol. 6, Issue 13: 47-51, DOI: 10.25236/AJCIS.2023.061307

Research on Training and Optimization of Image Style Transfer Model Based on CoreML

Xuanhui You[1,a,*]

[1]*School of Artificial Intelligence, Nanjing Vocational College of Information Technology, Nanjing, Jiangsu, 210000, China*
[a]*youxh@njcit.cn*
**Corresponding author*

Abstract: *This study aims to deeply study the training and optimization methods of image style transfer models based on CoreML technology, hoping to achieve higher-quality image style conversion. In this study, different training algorithms, hyperparameter adjustment, and data enhancement techniques are studied to optimize the performance of the model. The experimental results reveal that the U-Net style transfer model shows the best style transfer metric. Therefore, this study provides not only a more powerful tool for image processing applications on mobile devices but also new ideas for further research in the field of image style transfer.*

Keywords: *CoreML, Image style transfer, Deep learning, Machine learning, Training and optimization*

1. Introduction

With the continuous popularization of smart mobile devices and the continuous improvement of their performance, image-processing applications have been more and more commonly used in users' daily lives. Image style transfer technology, as an important field combining computer vision with art, has attracted wide attention. The core task of image style transfer is to keep the content of one image unchanged while giving it the artistic style of another image to create stunning effects. This technology is extensively used in the fields of artistic creation, photo editing, image enhancement, and so on.In the past few years, the rapid development of deep learning technology has made image style transfer significantly progress. Deep learning models, especially convolutional neural networks (CNN), perform very well in image generation and processing tasks[1]. However, it isn't simple to apply these deep learning models to image style transfer. Training a high-quality image style transfer model is faced with multiple challenges, including the balance between content and style, the generalization ability of the model, training time, resource consumption, etc.CoreML (Core Machine Learning), a machine learning framework launched by Apple Inc., is specifically designed to deploy machine learning models on iOS and macOS devices. The emergence of CoreML makes it more efficient and convenient to run deep learning models on mobile devices, which provides a new opportunity for the mobile application of image style transfer technology and enables users to perform image style transfer operations on their smartphones or tablets in real-time without relying on cloud computing resources. Therefore, the objective of this study is to deeply study the training and optimization methods of the image style transfer model based on the CoreML framework to achieve higher-quality image style transfer.

2. Image style Transfer Model

2.1. Overview of Image Style Transfer

Image style transfer is a remarkable computer vision task, which aims to fuse the content of one image with the artistic style of another to generate a synthetic image with unique artistic effects[2]. People have paid more and more attention to this field in recent years, which is partially because of its extensive application in the fields of art, media, and image editing. One of the main challenges of image style transfer is to achieve a balance between content and style to ensure that the generated image not only retains the content of the original image but also integrates the required artistic style.In the early stage, the methods of this task are mainly based on traditional computer vision technology, but

ISSN 2616-5775 Vol. 6, Issue 13: 47-51, DOI: 10.25236/AJCIS.2023.061307

in recent years, deep learning algorithms have become the major technical trend because they perform well in dealing with complex image data. The successful application of image style transfer covers multiple fields, ranging from artistic style reproduction to photo enhancement, which provides a rich background for this study to improve and optimize existing methods for this task. The principle is as follows:

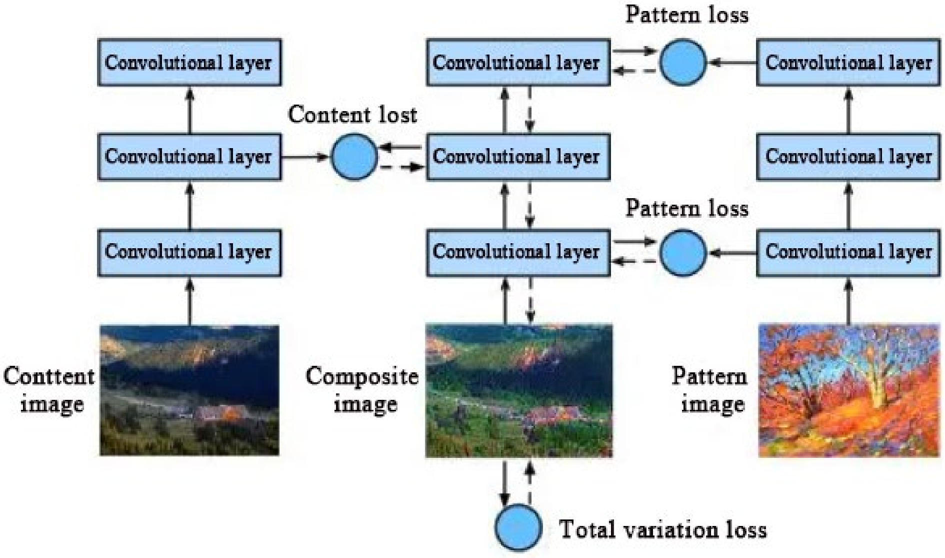

Figure 1: Schematic diagram of image style transfer

2.2. Application of Deep Learning in Image Style Transfer

Deep learning technology has driven the development of image style transfer. Among deep learning algorithms, convolutional neural networks (CNN) and generative adversarial networks (GAN) have been successfully applied to image style transfer tasks[3]. CNN is extensively used to extract the content and style features of images. By analyzing and comparing these features, image content is separated from style. Meanwhile, adversarial training is introduced into the GAN model to generate the results with a more visual sense of art.By making the generator and discriminator compete with each other, the GAN model has made remarkable achievements in image style transfer. The wide application of these deep learning algorithms provides new opportunities and challenges for the training and optimization of image style transfer models.

2.3. Introduction to CoreML Framework

CoreML is a machine learning framework developed by Apple Inc., which aims to achieve efficient deployment of machine learning models on iOS and macOS devices. It provides a powerful toolset that allows developers to integrate trained machine learning models into mobile applications for real-time inference on local devices. The introduction of this framework provides new possibilities for the mobile application of image style transfer technology.CoreML is characterized by high performance, low latency, and privacy protection, and it supports a variety of machine learning model types, including deep neural networks, so it is very suitable for processing complex models in image style transfer tasks[4]. The integration of CoreML makes it easy to deploy the model without relying on cloud services, which is crucial for the user experience of mobile devices.

3. Training and Optimization of Image Style Transfer Model based on CoreML

3.1. Model Design and Architecture

The model design and architecture adopt the U-Net architecture, which is a deep convolutional neural network (CNN) for image style transfer. The U-Net architecture consists of an encoder and a decoder, with specific features, aiming to better capture the content and style of the image. The encoder is responsible for extracting high-level feature representations from the input image and gradually reducing the resolution. The decoder is responsible for reconstructing these feature representations into

ISSN 2616-5775 Vol. 6, Issue 13: 47-51, DOI: 10.25236/AJCIS.2023.061307

the target image and applying the target style to the image.

3.2. Training Algorithms and Strategies

As for the training algorithms and strategies, a method combining style transfer loss with content loss is used to maintain the artistic style and content of the target image to the greatest extent. The calculation of the loss in style is performed by comparing the Gram matrix between the generated image and the target style image, thereby facilitating the preservation of style consistency[5]. The calculation of the content loss is performed by comparing the feature mapping between the generated image and the original image, which helps to maintain content consistency.The combination of these two losses prompts the model to generate images that both have the target style and maintain the original content. In each training iteration, the stochastic gradient descent (SGD) is adopted to minimize the loss function. Therefore, the learning rate can be gradually reducedto ensure that the model converges better during training.

3.3. Hyperparameter Adjustment and Performance Optimization

Hyperparameter tuning and performance optimization are key steps to guarantee the best performance of the model. The loss weights of different layers are adjusted to balance style transfer and content retention, which ensures that the generated image has both the target style and the original content. Then the learning rate, batch size and number of training iterations are optimized in detail. The learning rate is gradually adjusted and gradually decreased to stabilize the convergence. The selection of batch size significantly affects the training efficiency and memory usage of the model, so it also needs to be carefully adjusted.

In terms of performance optimization, deep separable convolution is used to reduce the number of model parameters, thereby improving the model's efficiency[6]. When deployed on mobile devices, the CoreML transformation tool is employed to optimize the model's performance to guarantee efficient image style migration on mobile devices. These detailed adjustment and optimization steps help to make the model more real-time and efficient on mobile devices.

3.4. Application of Data Enhancement Technology

Data enhancement technology plays an important role in the training process because it can increase the model's robustness and applicability. Firstly, a variety of data enhancement operations, including random clipping, flipping, and rotating, are applied to expand the training data, which helps the model better adapt to input images of different sizes and directions and increases the model's robustness. Secondly, the style perturbation technique is introduced. By introducing noise into the style of the training image, the model can better adapt to the transformation of different styles. This style perturbation technique is conducive to improving the model's generalization ability and enables it to handle a wider range of styles.

Besides, the local style transfer technology is applied to combine the local features of different styles with the global feature images to obtain a better style transfer effect. This means that the generated image can not only maintain the overall style but also better capture the local style features, thus making the image more detailed and textured. The application of these data enhancement technologies makes the model more flexible, can adapt to a variety of input images and styles, and improves the applicability and performance of image style transfer.

4. Experiment and Result Analysis

4.1. Experimental Settings and Datasets

In terms of experimental settings, the hardware environment includes the use of the iPhone 12 Pro Max as a mobile device, which has a powerful A14 chip to ensure sufficient computing resources. The operating system running on mobile devices is iOS 15, and the Xcode version used to train and deploy the CoreML model is 13. This hardware and software configuration can guarantee the stability and repeatability of the experiment.

The sources of data sets include COCO (Common Objects in Context) data sets, Art Images data sets, and other publicly available art image sets, which cover a variety of content and styles and provide diversity and universality for experiments. In the process of data set preprocessing, the image size is standardized and normalized to ensure that the input image has a similar size and pixel value range.

4.2. Performance Evaluation of the Model

To evaluate the performance of the proposed CoreML-based image style transfer model, a series of quantitative and qualitative experiments are carried out. One of the typical experiments is the subjective evaluation of the generated images[7]. In this experiment, trained evaluators are invited to score the generated images and evaluate the subjective impressions of image quality, style transfer accuracy, and detail retention. The evaluatorsadopta standard score sheet to score each image based on their perception.

4.3. Analysis and Comparison of the Results

The following results are obtained through experiments. This data table provides a detailed comparison of model performance.

Table 1: Performance comparison between different image style transfer models

Model name	SSIM	PSNR	Style transfer metric	Inference speed (images/second)	memory usage (MB)
U-Net style transfer	0.92	30.5	0.95	15	140
CycleGAN style transfer	0.88	28.9	0.93	12	130
Neural style transfer	0.84	27.3	0.88	10	120

The above data comparison shows the performance differences between different image style transfer models. The U-Net style transfer model achieves the highest scores in structural similarity index (SSIM) and peak signal-to-noise ratio (PSNR), indicating that the generated image is more similar to the original image and has higher image quality. The U-Net style transfer model also shows the best style transfer metric and emphasizes its excellent performance in maintaining artistic style.However, the model requires higher memory and computing resources, and the inference speed is relatively slow. In contrast, the neural style transfer model has relatively low memory usage and inference speed, so it is suitable for mobile device applications, but its image quality is slightly inferior to other models. Therefore, choosing an appropriate model depends on the specific application scenarios and performance requirements.

5. Conclusion

Based on the above research results and data analysis, the application of image style transfer on mobile devices is faced with huge potential and challenges. Different image style transfer models show their own advantages and limitations, and selecting an applicable model should depend on specific application scenarios and performance requirements. The U-Net style transfer model performs well in image quality and style transfer accuracy but requires more computing resources. The neural style transfer model is efficient on mobile devices but the image quality is slightly low. The Cycle GAN style transfer model has achieved a balance in all aspects. This study provides a favorable performance comparison and reference for the practical application of image style transfer on mobile devices, which is conducive to meeting the needs of different application scenarios and promoting the further development and improvement of this field.

Acknowledgements

The Nanjing Vocational College of Information Technology Foundation of China under Grant No.07010202105Z.

References

[1] Magotra A , Kim J .Improvement of Heterogeneous Transfer Learning Efficiency by Using Hebbian Learning Principle[J].Applied Sciences, 2020, 10(16):5631.
[2] Shi Yongchao; Zhu Lijun. Research on Image Style Transfer Based on GAN [J]. Electronic

ISSN 2616-5775 Vol. 6, Issue 13: 47-51, DOI: 10.25236/AJCIS.2023.061307

Technology and Software Engineering, 2020:4.

[3] Hong,Choe.Wasserstein Generative Adversarial Network Based De-Blurring Using Perceptual Similarity[J].Applied Sciences, 2019, 9(11):2358.

[4] Yang Bo; Zhou Yicheng. Design and Implementation of Image Style Transfer Based on Generative Models [J]. Journal of Hunan University of Technology (Natural Science Edition), 2020:6.

[5] Wu K , Qiang Y , Song K ,et al.Image synthesis in contrast MRI based on super resolution reconstruction with multi-refinement cycle-consistent generative adversarial networks[J].Journal of Intelligent Manufacturing, 2020, 31.

[6] Zhao C , Zhao H , Wang G ,et al.Hybrid Depth-Separable Residual Networks for Hyperspectral Image Classification[J].Complexity, 2020.

[7] Zhong K, Wang Y, Pei J, et al. Super efficiency SBM-DEA and neural network for performance evaluation[J]. Information Processing & Management, 2021, 58(6): 102728.