Enhancing Methods for Restorable Arbitrary Style Transfer in Image Stylization

Final talk

Karatyshchev Dmitry

Department of Mathematical Forecasting Methods, Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University

December 21, 2024

Goals

- Revisit and implement the Restorable Arbitrary Style Transfer (RAST) framework.
- Conduct an ablation study to evaluate the significance of various loss components.
- Introduce an idempotency loss to enforce consistent style transfer behavior.
- Adapt the multirestoration loss for low-resolution images to maintain visual fidelity.
- Explore architectural simplifications to enhance model efficiency without compromising performance.
- Provide insights and recommendations for future improvements in arbitrary style transfer techniques.

Arbitrary Image Style Transfer (AIST)

- **Definition:** Arbitrary Image Style Transfer (AIST) is the task of transferring the artistic style of one image (style image) onto another image (content image) while preserving the content of the original image.
- **Objective Function:**

$$\mathcal{L} = \lambda_{\text{content}}\mathcal{L}_{\text{content}} + \lambda_{\text{style}}\mathcal{L}_{\text{style}}$$

  - $\mathcal{L}_{\text{content}}$: Content loss ensuring preservation of content features.
  - $\mathcal{L}_{\text{style}}$: Style loss ensuring accurate replication of style features.
  - λ_{content} and λ_{style}: Weighting factors balancing content and style contributions.
- **Applications:**
  - Creation of digital art and enhanced photography.
  - Improving visual content for virtual and augmented reality environments.
  - Automating design and image editing processes.

Challenges in Arbitrary Image Style Transfer

- **Content Preservation:** Ensuring that the essential structures and objects in the content image remain intact after style transfer.
- **Style Accuracy:** Precisely replicating the artistic style, including colors, textures, and brushstrokes, from the style image.
- **Scalability:** Enabling the model to handle an unlimited variety of styles without the need for separate models for each style.
- **Real-Time Performance:** Achieving fast processing speeds suitable for applications requiring instant style transfer, such as mobile apps and interactive systems.
- **Artifact Minimization:** Reducing visual artifacts and inconsistencies that may arise during the style transfer process.

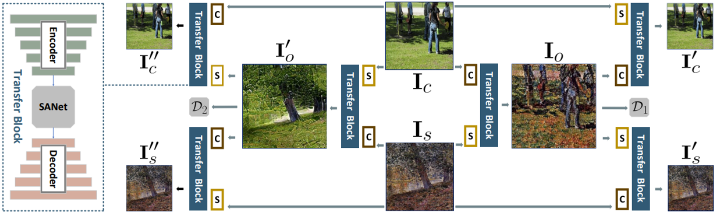

Figure: Overview of the RAST Architecture

RAST Objective Function

- **Overall Objective:**

$$\mathcal{L} = \lambda_{contra} L_{contra} + \lambda_{identity} L_{identity} + \lambda_{diff} L_{diff}^{-1} + \lambda_{multi} L_{multi} + \lambda_{adv} L_{adv}$$

- **Explanation of Loss Terms:**
  - L_{contra}: Enforces contrastive learning to enhance content preservation.
  - $L_{identity}$: Ensures the identity consistency between input and output images.
  - L_{diff}^{-1}: Minimizes the inverse difference to reduce content distortion.
  - L_{multi}: Balances multiple restoration pathways for arbitrary style transfer.
  - L_{adv}: Adversarial loss improves perceptual quality by leveraging a discriminator.

Proposed Method Overview

- Enhancing the Restorable Arbitrary Style Transfer (RAST) framework through targeted modifications.
- Key Contributions:
  1. Ablation Study to assess the significance of various loss components.
  2. Introduction of Idempotency Loss to enforce consistent style transfer behavior.
  3. Adaptation of Multirestoration Loss tailored for low-resolution images.
  4. Architectural Simplifications to streamline the model without compromising performance.

Ablation Study

- **Objective:** Determine the contribution of each loss component in the RAST framework.
- **Methodology:**
  - Systematically remove each loss term from the total loss function.
  - Evaluate the impact on style transfer quality and restoration capability.
- **Total Loss Function:**

$$\mathcal{L} = \lambda_{\text{contra}}\mathcal{L}_{\text{contra}} + \lambda_{\text{identity}}\mathcal{L}_{\text{identity}} + \lambda_{\text{diff}}\mathcal{L}_{\text{diff}}^{-1} + \lambda_{\text{multi}}\mathcal{L}_{\text{multi}} + \lambda_{\text{adv}}\mathcal{L}_{\text{adv}}$$

  - $\mathcal{L}_{\text{contra}}$: Contrastive Loss
  - $\mathcal{L}_{\text{identity}}$: Identity Loss
  - $\mathcal{L}_{\text{diff}}^{-1}$: Style Difference Loss
  - $\mathcal{L}_{\text{multi}}$: Multi-Restoration Loss
  - $\mathcal{L}_{\text{adv}}$: Adversarial Loss

Ablation Study Results

- **Findings:**
  - Each loss component significantly contributes to the overall performance.
  - Removal of the Multi-Restoration Loss leads to a notable decline in restoration quality.
- **Visual Impact:**

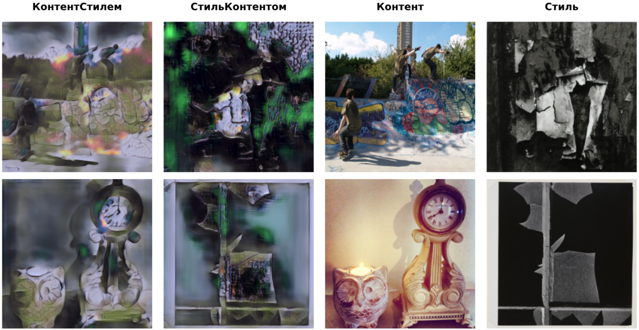

Figure: Effect of Removing Multi-Restoration Loss

Idempotency Loss

- **Purpose:** Ensure consistent stylization results when the style transfer operation is applied multiple times.
- **Mathematical Formulation:**

$$T(T(I_c, I_s), I_s) = T(I_c, I_s)$$

$$\mathcal{L}_{\text{idemp}} = \| T(T(I_c, I_s), I_s) - T(I_c, I_s)\|^2$$

  - T: Style Transfer Operation
  - I_c: Content Image
  - I_s: Style Image
- **Implementation:**
  - Adds a constraint that the output remains unchanged upon repeated applications of the same style.

Idempotency Loss Impact

- **Evaluation:**
  - Compare stylization results with and without Idempotency Loss.
  - Apply style transfer operation twice on the same content and style images.
- **Results:**
  - **With $\mathcal{L}_{\textbf{idemp}}$:** Consistent stylized images upon repeated applications.
  - **Without $\mathcal{L}_{\textbf{idemp}}$:** Subtle changes in style intensity and minor artifacts upon multiple applications.

Idempotency Loss Impact

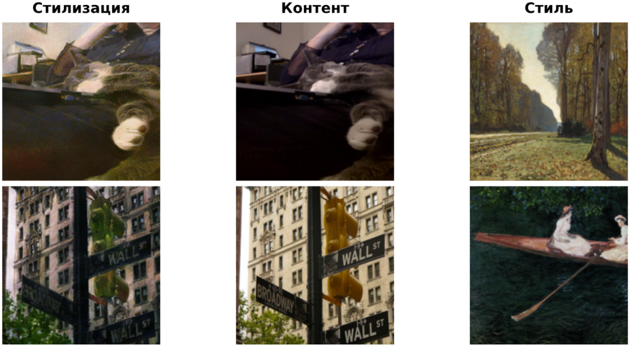

Figure: Stylization with Idempotency Loss ($\mathcal{L}_{\text{idemp}}$)

Idempotency Loss Impact

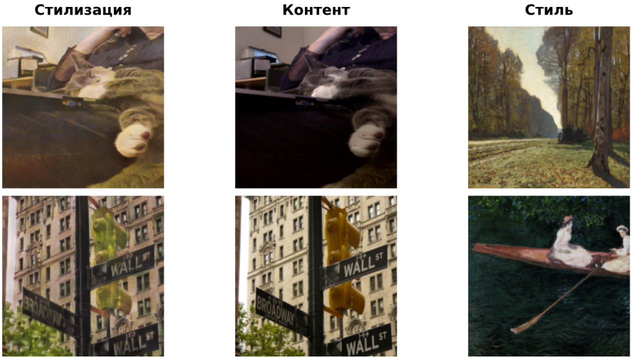

Figure: Stylization without Idempotency Loss

Multirestoration Loss for Low-Resolution Images

- **Challenge:** Standard Multi-Restoration Loss leads to excessive colorization in low-resolution images, degrading style transfer quality.
- **Proposed Solution:**
  - Adapt the Multi-Restoration Loss to better handle low-resolution scenarios.
  - Incorporate a weighting mechanism to emphasize restoration accuracy at lower resolutions.
- **Mathematical Formulation:**

$$\mathcal{L}_{\text{multi\_low\_res}} = \mathcal{L}_{\text{multi}}(\text{LowRes}(I_c), \text{LowRes}(I_c))$$

  - LowRes: Function to obtain low-resolution representations.
- **Outcome:** Aims to maintain color distribution and visual fidelity without over-altering colors.

Multirestoration Loss Impact on Low-Resolution Images

- **Evaluation:**
  - Conduct experiments on images resized to 32×32 pixels.
- **Results:**
  - **Performance:** Both baseline and enhanced models perform poorly on low-resolution images.
  - **Issues:** Noticeable artifacts and significant color deviations; stylized outputs lack detail and coherence.

Multirestoration Loss Impact on Low-Resolution Images

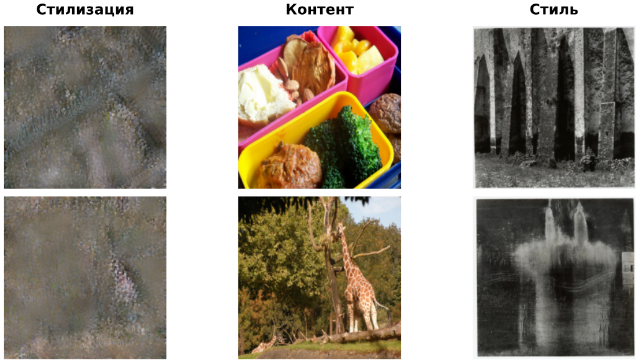

Figure: Stylization and Restoration on Low-Resolution Images (32×32 pixels)

Architectural Simplification

- **Objective:** Assess the impact of simplifying the RAST architecture by removing non-essential components.
- **Approach:**
  - Retain only core modules responsible for style transfer and restoration.
  - Replace encoder module with pretrained EfficientNet **tan2020efficientnetrethinkingmodelscaling**.
- **Findings:**
  - **Performance Degradation:** Simplified model exhibits poorer stylization results.
  - **Efficiency vs. Quality:** While the model is lighter, the quality of style transfer is adversely affected.
- **Conclusion:**
  - Certain architectural components are critical for maintaining high-quality style transfer.
  - Simplifications may introduce inefficiencies and degrade performance.

Overall Objective Function

- **Combined Objective:**

$$\mathcal{L} = \lambda_{\text{contra}}\mathcal{L}_{\text{contra}} + \lambda_{\text{identity}}\mathcal{L}_{\text{identity}} + \lambda_{\text{diff}}\mathcal{L}_{\text{diff}}^{-1} +$$

$$+\lambda_{\text{multi\_low\_res}}\mathcal{L}_{\text{multi\_low\_res}} + \lambda_{\text{adv}}\mathcal{L}_{\text{adv}} + \lambda_{\text{idemp}}\mathcal{L}_{\text{idemp}}$$

  - **Purpose:** Balances contributions of each loss component during training.
  - **Impact:** Ensures that all aspects of style transfer and restoration are adequately addressed.

Conclusion

Key Contributions:

- Enhanced the RAST framework by introducing the Idempotency Loss.
- Demonstrated that Idempotency Loss ensures consistent and reliable style transfer across multiple applications.

Insights:

- Other proposed modifications, such as adapting multirestoration loss for low-resolution images and architectural simplifications, did not yield the desired improvements.
- Idempotency Loss stands out as a significant advancement in ensuring stability and consistency in style transfer.

Future Work:

- Further refine the Idempotency Loss mechanism for enhanced performance.
- Explore alternative approaches to improve low-resolution stylization.
- Investigate additional architectural optimizations to build upon the strengths of Idempotency Loss.

Final Thoughts:

- This study provides a clear pathway for developing more robust and consistent arbitrary style transfer models, with Idempotency Loss being a pivotal component.

References I

[1] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2414–2423.

[2] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *European Conference on Computer Vision*, Springer, 2016, pp. 694–711.

[3] X. Huang and S. Belongie, "Arbitrary style transfer in real-time with adaptive instance normalization," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 1501–1510.

[4] Y. Li, C. Wang, and X. Tang, "Universal style transfer via feature transforms," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 4505–4513.

References II

[5] Z. Ma, Y. Zhang, and X. Li, "Rast: Restorable arbitrary style transfer via multi-restoration," in *Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV)*, Available at https://openaccess.thecvf.com/content/WACV2023/papers/Ma_RAST_Restorable_Arbitrary_Style_Transfer_via_Multi-Restoration_WACV_2023_paper.pdf, 2023, pp. xxxx–xxxx.

[6] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 2223–2232.

[7] I. Goodfellow, J. Pouget-Abadie, M. Mirza, *et al.*, "Generative adversarial nets," *Advances in Neural Information Processing Systems*, vol. 27, pp. 2672–2680, 2014.