192.151 Introduction to Deep Learning

2025W

Project 20: Image Style Transfer

1. Background and Motivation

Image style transfer aims to synthesize a new image that preserves the *content* of a given photograph while adopting the *style* (textures, colors, brush strokes) of a reference artwork. It is a classic example of using deep feature representations (e.g., VGG features) for perceptual objectives rather than explicit pixel-wise supervision. Beyond artistic applications, style transfer is a good case study for (i) content–style disentanglement, (ii) perceptual losses, and (iii) quality–efficiency trade-offs between optimization-based methods and fast feed-forward networks.

2. Problem Description

The primary goal is to implement and evaluate an image style transfer system, comparing quality and efficiency across methods.

a) **Data and Preprocessing:**
- Use any set of *content images* (e.g., COCO/your own photos) and a set of *style images* (e.g., WikiArt subset or a small curated collection of paintings).
- Standardize preprocessing (image resolution, normalization) and define a small fixed evaluation set (recommended: 20–50 content images $\times$ 5–10 styles).

b) **Image Style Transfer Model Implementation:**
- Optimization-based NST baseline: the classic neural style transfer objective with content loss + Gram-matrix style loss (iterative optimization).
- Adaptive Instance Normalization (AdaIN) for arbitrary style transfer with a tunable style strength.
- Transformer-based style transfer model (if compute allows).

c) **Evaluation and Analysis:**
- **Content preservation proxy:** perceptual distance between output and content image (e.g., VGG feature distance or LPIPS).
- **Style matching proxy:** Gram-matrix loss (VGG) or feature statistics distance between output and style image.
- **Qualitative gallery:** show representative grids (content $\times$ style) and discuss failure modes (content distortion, style leakage, color shift).
- **Strength ablation:** vary a style-strength knob (e.g., AdaIN interpolation $\alpha \in \{0.2, 0.5, 0.8, 1.0\}$) and analyze the trade-off.

3. Expected Deliverables

Students working in groups (ideally 3 members) are expected to:

- produce a project report,
- deliver a group presentation,
- submit well-commented source code (training + inference + evaluation scripts).

4. Suggested Resources and References

[1] Gatys, L. A., Ecker, A. S., & Bethge, M. (2016). Image Style Transfer Using Convolutional Neural Networks. In *CVPR*. `https://arxiv.org/pdf/1508.06576`

[2] Johnson, J., Alahi, A., & Fei-Fei, L. (2016). Perceptual Losses for Real-Time Style Transfer and Super-Resolution. In *ECCV*. `https://arxiv.org/pdf/1603.08155`

[3] Dumoulin, V., Shlens, J., & Kudlur, M. (2017). A Learned Representation for Artistic Style. In *ICLR*. `https://arxiv.org/pdf/1610.07629`

[4] Huang, X., & Belongie, S. (2017). Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization. In *ICCV*. `https://arxiv.org/pdf/1703.06868`

[5] Multiple Style Transfer (PyTorch reference implementation). `https://github.com/zhanghang1989/PyTorch-Multi-Style-Transfer`