# Emotion-based Style Transfer

Dishant Padalia
University of Massachusetts Amherst
Amherst, MA
dpadalia@umass.edu

Rahul Saxena
University of Massachusetts Amherst
Amherst, MA
rahulsaxena@umass.edu

## Abstract

*Emotion-based style transfer using deep CNNs introduces an innovative approach integrating emotions into style transfer, creating emotionally enriched images.*

## 1. Group Members and Contributions

- Dishant will be responsible for building, training, and fine-tuning the emotion recognition model.
- Rahul will work on developing the style transfer network, and Dishant will support him in improving the model performance.
- We plan to share the workload for writing the project report and collaborate while integrating emotion recognition and style transfer models.

## 2. Motivation

Emotion-based style transfer has many applications, like art creation, content personalization, digital storytelling and designing in-game styles. It allows content creators to amplify the emotional impact of their visuals, offering new possibilities for emotionally engaging multimedia content.

## 3. Literature Review

B. Nojavanasghari et al. [6] proposed "EmoReact," an approach for recognizing emotional responses in children, highlighting behavioural cues, and comparing emotion recognition models. Gatys et al. [2] introduced neural style transfer using CNNs that blend content and style from images. We are also investigating the latest advancements in neural network architectures and pre-trained models like VGG [7], ResNet [3], and more.

## 4. Data

We plan to use established emotion detection datasets, namely AffectNet [5] and Fer2013 [4], to train our emotion detection and style transfer models. To further improve the model generalization and performance, we may explore techniques like data augmentation and class weighting to offset the imbalance in the datasets, and we plan to use Google Colab GPU resources for model training.

## 5. Approach

### 5.1. Stage 1: Emotion Detection

We will train a CNN-based model for emotion detection from images, which will classify input images into emotion labels (e.g., happy, sad, anger), serving as a basis for the style to be used in the subsequent stage.

### 5.2. Stage 2: Style Transfer

Here, we will design a set of style transfer models. The process in this stage involves:

1. **Model Training** - For each emotion class, we will curate emotion-specific datasets of images, along with a corresponding artistic style (e.g., melancholic for sad, vibrant for joyful).
2. **Adaptive Style Transfer** - Upon detecting the emotion class from Stage 1, we will dynamically apply the relevant style model [1] associated with that class to the image, producing an output image aligned with the detected emotional state.

## 6. Evaluation Metrics

- **Style Transfer Quality:** Mean Squared Error (MSE), Structural Similarity Index (SSIM), or Peak Signal-to-Noise Ratio (PSNR) can be used to measure the quality of style transfer.
- **Emotion Recognition:** Accuracy and F1-score will be used for model evaluation.
- We will plot training curves that depict the change of style loss and content loss over iterations and across various trained models.

## References


[1] Daniel Bossett, David Heimowitz, Nidhi Jadhav, Leilani Johnson, Arti Singh, Helen Zheng, and Sabar Dasgupta. Emotion-based style transfer on visual art using gram matrices. In *2021 IEEE MIT Undergraduate Research Technology Conference (URTC)*, pages 1–5. IEEE, 2021. 1

[2] Leon A Gatys, Alexander S Ecker, and Matthias Bethge. A neural algorithm of artistic style. *arXiv preprint arXiv:1508.06576*, 2015. 1

[3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016. 1

[4] Yousif Khaireddin and Zhuofa Chen. Facial emotion recognition: State of the art performance on fer2013. *arXiv preprint arXiv:2105.03588*, 2021. 1

[5] Ali Mollahosseini, Behzad Hasani, and Mohammad H Mahoor. Affectnet: A database for facial expression, valence, and arousal computing in the wild. *IEEE Transactions on Affective Computing*, 10(1):18–31, 2017. 1

[6] Behnaz Nojavanasghari, Tadas Baltrušaitis, Charles E Hughes, and Louis-Philippe Morency. Emoreact: a multimodal approach and dataset for recognizing emotional responses in children. In *Proceedings of the 18th acm international conference on multimodal interaction*, pages 137–144, 2016. 1

[7] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. 1