

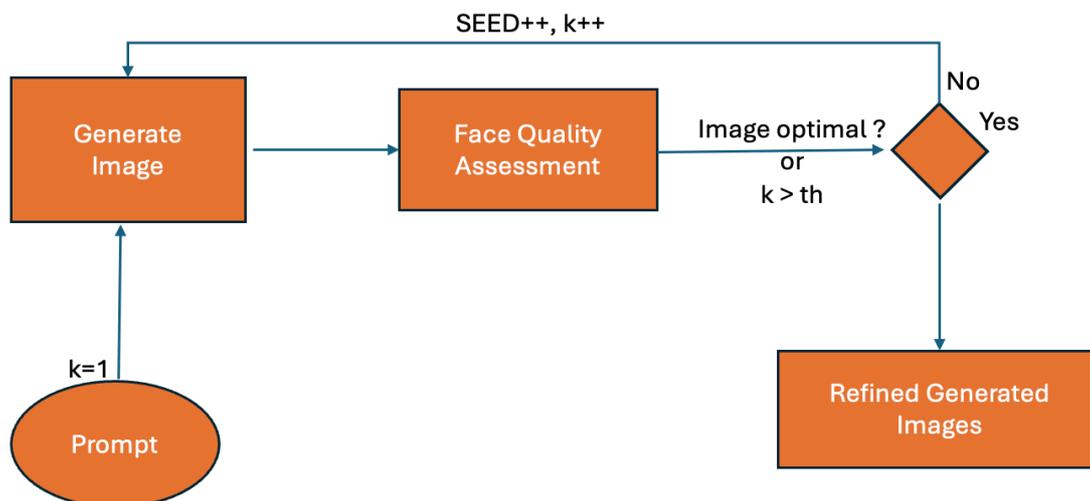
FACE QUALITY TRANSFORMER: A FACE QUALITY ASSESSMENT AND ENHANCEMENT FRAMEWORK

Appendix



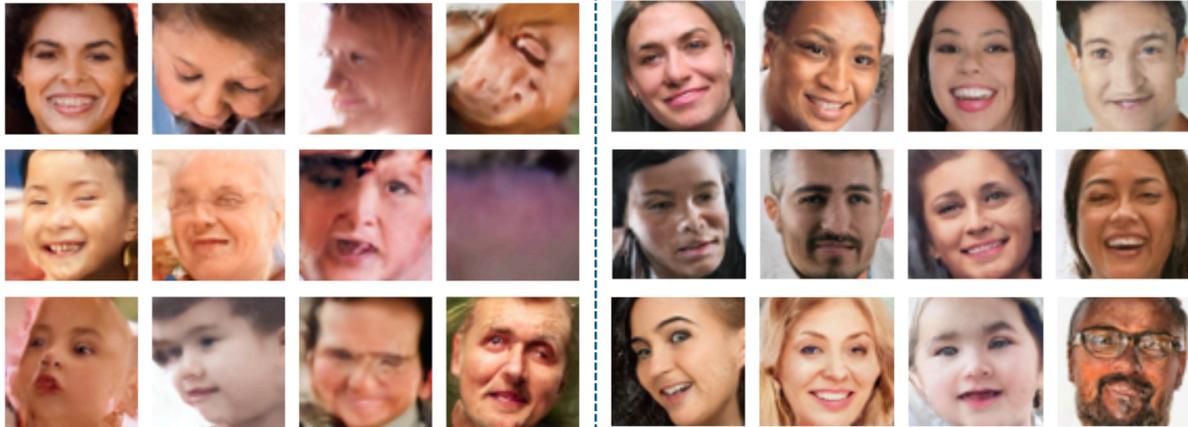
Sample images shown here are human faces generated by various state-of-the-art diffusion models including Firefly, DALL-E, Stable Diffusion and Flux. Green boxes show high-quality faces and red boxes show poor-quality faces labeled by Our FaceQuality Transformer.

Correction Pipeline and User acceptance



Architecture diagram of user acceptance experiment

More Examples of Generated Data by diffusion model trained using Face quality estimation loss



Images on the right are the output of diffusion model trained using Face Quality Estimation loss, on the left are trained using simple MSE loss

Model Training Details : We conducted all the experiments on NVIDIA A10 GPUs (24GB VRAM) with BinaryFocalCrossentropy loss and AdamW optimizer (starting with learning rate of $8e-5$ for Imagenet pre-trained weights till learning rate reaches $1e-7$ using ReduceLRonPlateau Learning Rate annealing). We observe that using the high learning rate makes the network forget information learned by pre-trained model architecture on imagenet dataset. Apart from this we also did data augmentation including adding random gaussian noise, converting images to grayscale, random resizing, ColorJitter, Random Sharpening, flip and Random shadows introduction to cover different lighting conditions of faces during training.