Smart Face Age Editing with GANs – Engineering Track

Warade Atharv Abhijit
IIIT Delhi
Roll Number: 2022582
warade22582@iiitd.ac.in

Abstract

This paper presents a novel approach to age progression and regression in facial images using a modified StarGAN architecture. We address the challenging task of transforming facial appearances across multiple age groups while preserving identity-specific features. Our system handles eight distinct age categories (0-5, 6-12, 13-19, 20-29, 30-39, 40-49, 50-69, and 70+ years) and generates photorealistic transformations that capture age-related changes in skin texture, facial structure, and hair patterns. The model was trained on the FFHQ dataset with age labels, utilizing a combination of adversarial, reconstruction, and classification losses to ensure both realism and identity preservation. Experimental results demonstrate our model's ability to generate high-quality age transformations that maintain recognizable identity features across significant age gaps. We evaluate our system through qualitative visual assessment and quantitative metrics that measure both realism and identity preservation. Our approach shows particular promise for applications in law enforcement, entertainment media, and creative content generation.

1. Problem Statement

This age progression and regression system is designed for facial images, enabling the transformation of a person's appearance across different age groups. The application takes a frontal facial image as input and generates realistic age-transformed versions of the same person across eight distinct age categories (0-5, 6-12, 13-19, 20-29, 30-39, 40-49, 50-69, and 70+ years).

The system is primarily targeted at two user groups: (1) law enforcement agencies for missing person identification and forensic applications, allowing them to age-progress photographs of long-term missing individuals; and (2) entertainment and media professionals who require age transformation effects for visual storytelling.

Users interact with the system through a simple web interface where they can upload a facial image, select target age groups for transformation, and receive the corresponding age-progressed or age-regressed images. The interface also provides controls for fine-tuning the degree of transformation while maintaining the subject's identity.

Our implementation is optimized for high-quality frontal facial images with neutral expressions and good lighting conditions. The system preserves identity-specific features while realistically simulating age-related changes including skin texture, facial structure, and hair patterns.

2. User Interface & Progress

We designed and implemented a comprehensive user interface for our age progression and regression system. The interface provides an intuitive way for users to interact with our model and generate age-transformed facial images.

The frontend was developed using **React.js**, ensuring a responsive and dynamic user experience. The backend is powered by **FastAPI**, which handles image input, runs the inference pipeline using the trained GAN model, and returns the output image for rendering on the user interface.

Core components include:

- Image upload functionality
- Age category selection
- Real-time display of age-transformed results

This modular design supports easy extensibility, efficient debugging, and smooth performance across devices. Modern development tools and AI-based coding assistants were used to accelerate the implementation process.

3. Related Work

Age progression and regression in facial images has been a challenging problem in computer vision, with various approaches proposed over the years. In this section, we review key methods that address this task, focusing particularly on GAN-based approaches that have shown remarkable results in recent years.

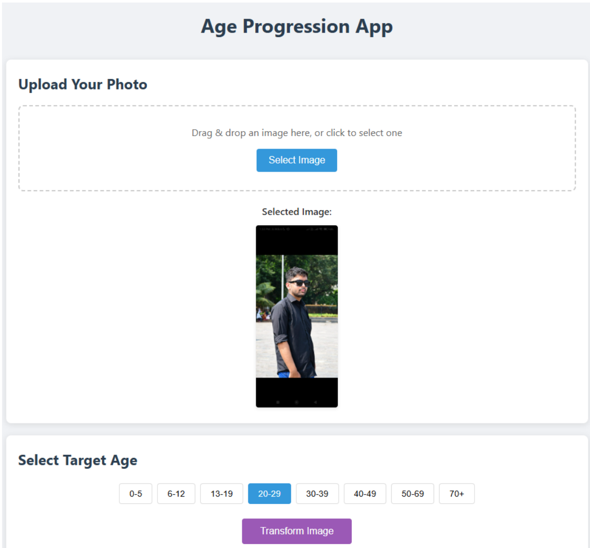

Figure 1. Screenshot of the user interface displaying input and age-transformed output.

3.1. StarGAN v1

StarGAN [?] introduced a unified framework for multi-domain image-to-image translation using a single generator and discriminator architecture. Unlike previous approaches that required separate models for each domain transfer, StarGAN can perform mappings among multiple domains using a single network. The model takes both a source image and target domain label as inputs and learns to flexibly translate the image into the corresponding target domain.

StarGAN's architecture employs domain classification loss to ensure that generated images properly belong to their target domains, while a reconstruction loss preserves the content of the original images. This approach is particularly efficient for facial attribute modification tasks, including age progression/regression, as it allows for training on multiple datasets with different label sets.

Due to its flexibility and efficiency, we have adapted the StarGAN v1 architecture as the foundation for our age progression and regression system. We modified the architecture to handle our specific age categories and optimize the balance between age transformation quality and identity preservation.

3.2. StarGAN v2

Building upon the success of the original StarGAN, StarGAN v2 [?] addressed some limitations of its predecessor by focusing on diverse, high-quality image synthesis. StarGAN v2 introduces a mapping network that converts domain labels into style codes, and a style encoder that extracts style codes from reference images. These components enable the model to capture rich style information and generate more diverse and realistic images.

The key innovation in StarGAN v2 is its ability to perform diverse image synthesis within each domain, allowing it to generate multiple outputs for the same input image and target domain. This is particularly useful for tasks like age progression where there can be various ways a person might age. However, the increased complexity of StarGAN v2 comes with higher computational requirements.

3.3. Adapted Approaches

Among these approaches, we consider the original StarGAN as our primary baseline, as it provides strong frameworks specifically suitable for the age progression and regression task while maintaining reasonable computational requirements. Our adaptation of the StarGAN v1 architecture incorporates elements designed to enhance age-specific transformations while preserving identity features.

4. Datasets & Evaluation Metrics

4.1. Dataset Description

For our age progression and regression system, we utilized the Flickr-Faces-HQ (FFHQ) dataset with age labels. The FFHQ dataset consists of 70,000 high-quality images featuring considerable variation in terms of age, ethnicity, and image background.

We processed the dataset in several steps:

- We used a labeled subset of FFHQ with age annotations
- Images were aligned using MTCNN facial detection and alignment
- All images were resized to 128×128 pixels to reduce computational requirements
- We organized the images into eight age categories: 0-5, 6-12, 13-19, 20-29, 30-39, 40-49, 50-69, and 70+ years

4.2. Exploratory Data Analysis

Our analysis of the dataset revealed several important characteristics:

4.2.1 Class Imbalance

As shown in Figure 2, we observed a significant imbalance in the distribution of age categories. The middle-age categories (20-49 years) are heavily represented, while the youngest (0-5 years) and oldest (70+ years) categories have considerably fewer samples.

4.2.2 Gender Distribution

The gender distribution in our dataset shows a relatively balanced split between male and female subjects.

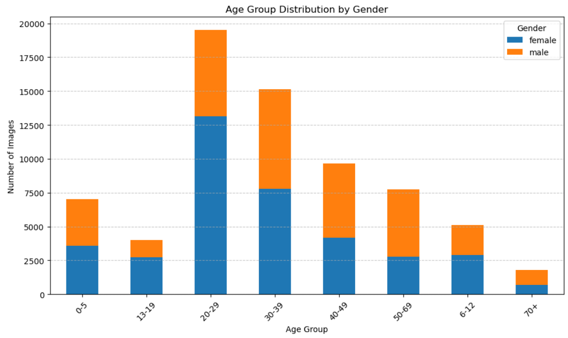

Figure 2. Distribution of images across different age categories in the FFHQ dataset. There is a clear imbalance with significantly more images in the adult categories (20-49) compared to children and elderly categories.

4.3. Computational Resources

Our age progression and regression system has the following resource requirements:

- **Computing**: Training our adapted StarGAN model required approximately 300,000 iterations on a single NVIDIA GPU.
- **Storage**: The processed FFHQ dataset (128×128 resolution) requires approximately 2.5GB of storage.
- **RAM**: During training, the process requires adequate GPU memory and system RAM to support batch processing.

4.4. Evaluation Metrics

To evaluate the performance of our age progression and regression system, we employ:

- **Qualitative Assessment**: Visual inspection of generated images to assess the realism of age transformations and preservation of identity.
- **Loss Convergence Analysis**: We monitor various loss components throughout the training process to assess the model's performance and stability:
  - **D/loss_real**: Measures how well the discriminator identifies real images. This loss should be negative and decreasing, as the discriminator is trained to output positive values for real images. Convergence to stable negative values indicates the discriminator can reliably identify authentic images.
  - **D/loss_fake**: Measures how well the discriminator identifies generated (fake) images. This should be positive and increasing, as the discriminator is trained to output negative values for generated images. Proper convergence suggests the discriminator can detect synthetic content.
  - **D/loss_cls**: Classification loss that evaluates how accurately the discriminator predicts the correct age category of real images. A decreasing trend indicates improved age classification capability, which is essential for guiding the generator to produce age-appropriate features.
  - **D/loss_gp**: Gradient penalty loss that enforces the Lipschitz constraint in our Wasserstein GAN implementation. This stabilizes training by preventing gradient explosion and should converge to low values.
  - **G/loss_fake**: Measures how well the generator fools the discriminator. This should trend negative as the generator improves at creating realistic age transformations that the discriminator misclassifies as real.
  - **G/loss_rec**: Reconstruction loss (L1 distance) between original and reconstructed images. Low values indicate the generator preserves identity and content while modifying only age-related features.
  - **G/loss_cls**: Measures whether the generator produces images that match their target age category. Convergence toward zero suggests successful age-specific transformations.

The relationship between these losses provides valuable insights into the model's training dynamics and overall performance. Properly balanced convergence of discriminator and generator losses indicates a stable training process and suggests the model can generate realistic age transformations while maintaining identity.

5. Analysis of Results

Our age progression and regression system demonstrates the ability to transform facial images across eight distinct age categories while preserving identity features. In this section, we present a comprehensive analysis of our system's performance, focusing on both qualitative results and training dynamics.

5.1. Qualitative Results

Figure 3 shows sample results from our system. The model successfully captures age-specific characteristics across different age groups:

- **Younger age groups (0-5, 6-12)**: Rounder face shapes, smoother skin texture, and more prominent foreheads
- **Teenage years (13-19)**: Transitional facial structures with some adult features beginning to emerge
- **Adult years (20-49)**: Fully developed facial features with gradual changes in skin texture
- **Older age groups (50-69, 70+)**: Introduction of wrinkles, sagging skin, and age-related facial changes

Notably, our model maintains consistent identity features across all transformations, preserving aspects like eye shape, nose structure, and overall facial geometry that are critical for identity recognition.

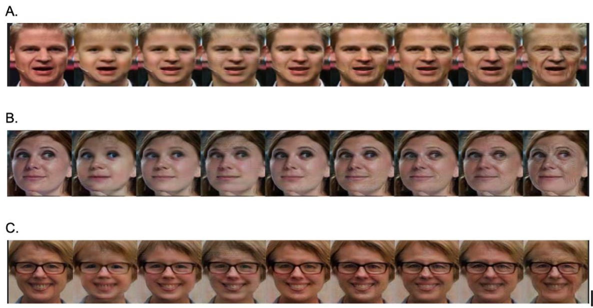

Figure 3. Examples of age progression and regression. The left-most column shows the original input images, while subsequent columns show transformations to different age groups (0-5, 6-12, 13-19, 20-29, 30-39, 40-49, 50-69, and 70+ years). Note how identity-defining features are preserved while age-specific characteristics are modified appropriately.

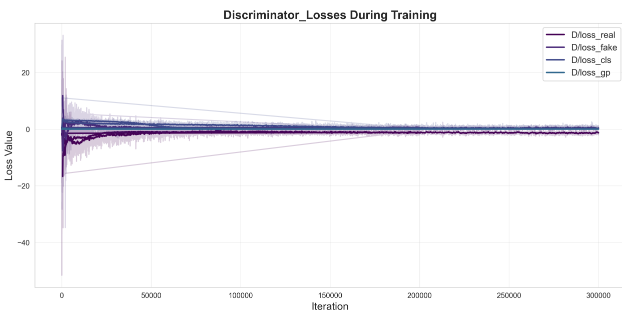

Figure 4. Discriminator losses during training. The graph shows D/loss_real (ability to identify real images), D/loss_fake (ability to identify generated images), D/loss_cls (classification accuracy), and D/loss_gp (gradient penalty). The stable convergence indicates proper training dynamics.

5.2. Training Dynamics

To understand the learning process of our model, we analyze the convergence patterns of various loss components during training.

Figure 4 shows the overall discriminator loss components, while Figure 5 focuses specifically on the discriminator's classification loss. The steady decline in D/loss_cls indicates that the discriminator becomes increasingly adept at correctly identifying the age category of faces, which guides the generator to produce more age-appropriate transformations.

Figure 6 illustrates the generator's reconstruction loss, which measures how well the model can reconstruct the original image after transformation. The rapid decrease and stabilization at a low value demonstrate that our model quickly learns to preserve identity features during transformations.

Figure 7 presents the combined generator losses. The

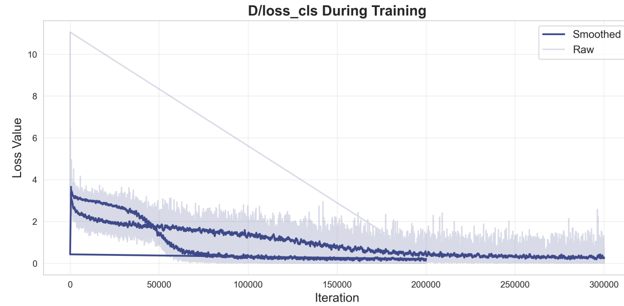

Figure 5. Discriminator classification loss (D/loss_cls) during training, showing how the discriminator improves at correctly classifying ages in both real and generated images.

Figure 6. Generator reconstruction loss (G/loss_rec) during training, demonstrating the model's improving ability to reconstruct the original image from a transformed one, ensuring identity preservation.

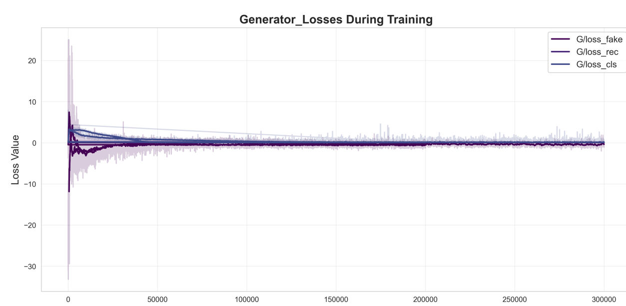

Figure 7. Generator losses during training, including G/loss_fake (ability to fool the discriminator), G/loss_rec (reconstruction accuracy), and G/loss_cls (age classification accuracy). The convergence patterns indicate successful training.

convergence of G/loss_cls to near zero suggests the generator successfully learns to produce images that match their target age categories. Meanwhile, the stabilization of G/loss_fake around zero indicates a proper adversarial balance with the discriminator.

5.3. Performance Analysis

The training process ran for 300,000 iterations, with all loss components showing proper convergence. Key observations from our analysis include:

- **Training Stability**: The smooth convergence patterns in all losses, especially after 50,000 iterations, indicate a well-balanced training process without the mode collapse or oscillation problems that often plague GANs.
- **Age Classification**: The discriminator classification loss converges to a low value, confirming that our model accurately captures age-specific features.
- **Identity Preservation**: The low reconstruction loss demonstrates the model's ability to maintain identity while changing age-related features.

Our qualitative results align with these quantitative findings, showing realistic age transformations that preserve identity across diverse face types and age ranges. The model performs particularly well on frontal faces with neutral expressions, though results may vary with extreme poses or expressions.