

EE655 - COURSE PROJECT

MULTI STYLE IMAGE
GENERATION USING
STYLEGANS

Presented by : Group 24

Our Team



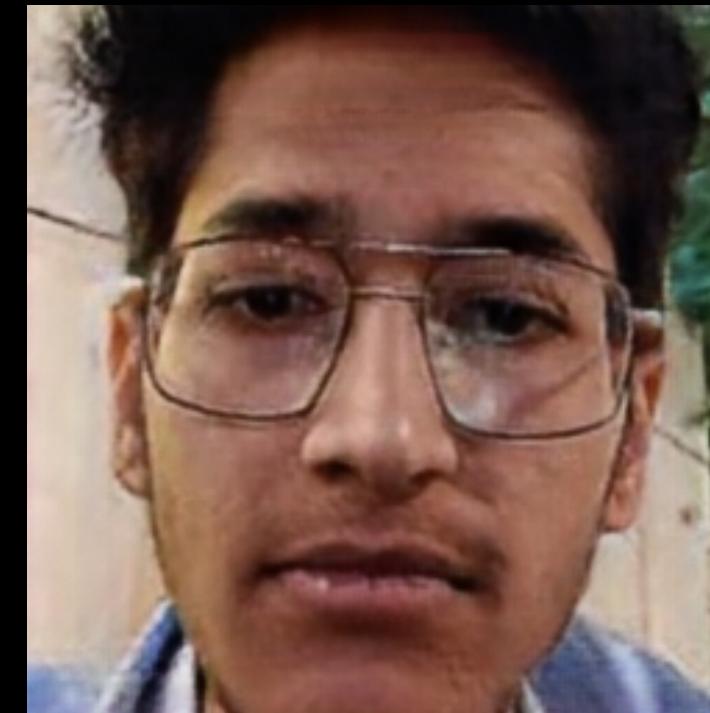
Anirudh Singh
230142



Aarush Singh
Kushwaha
230019



Kshitij Bhardwaj
230580



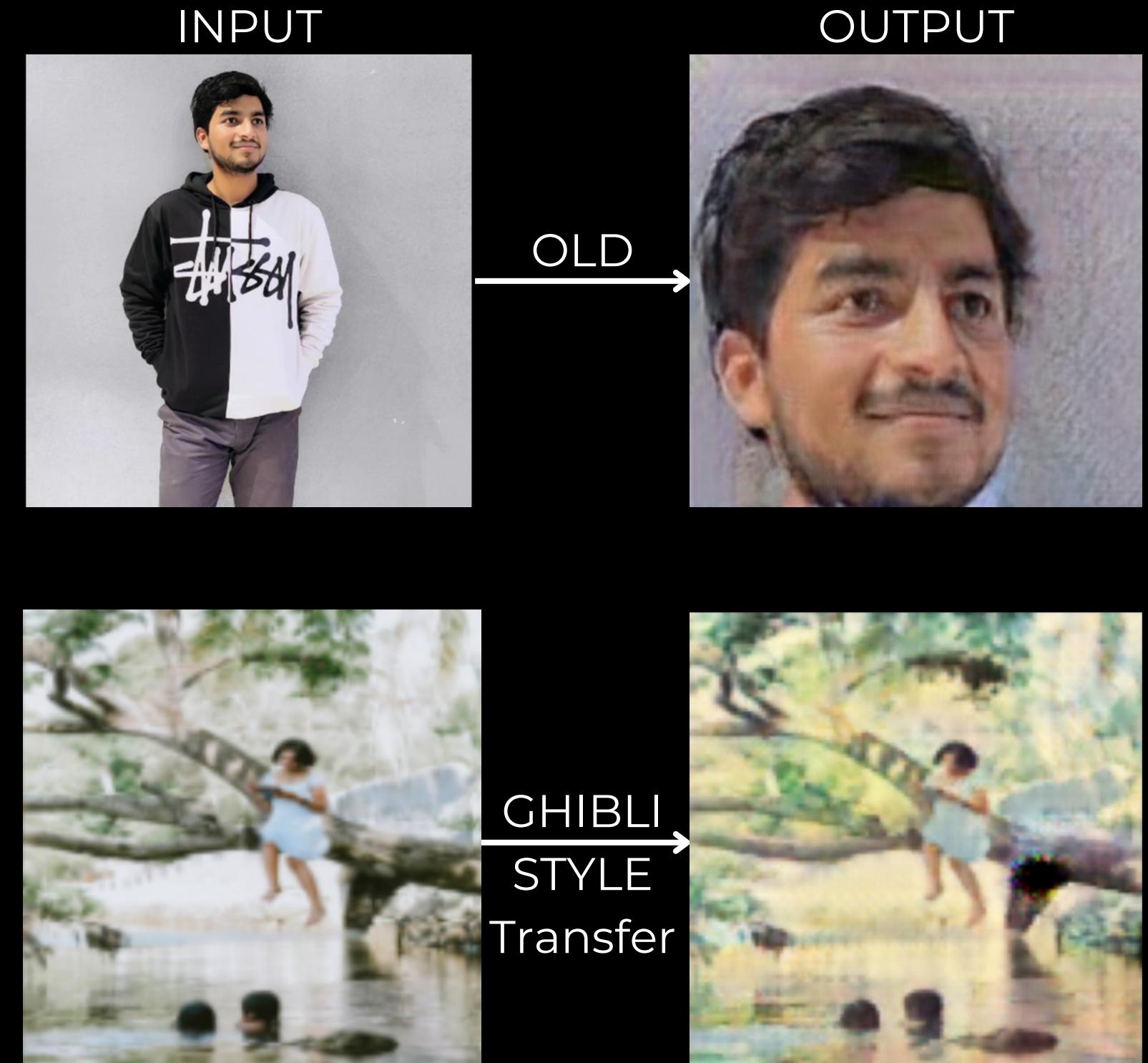
Fariz
Mashhood Ali
230402

Problem Statement

This project addresses two distinct but related tasks in unpaired image-to-image translation:

1. Age Manipulation
2. Ghibli-style Conversion

Both tasks require transforming facial images while preserving core identity features and producing visually realistic results.



Problem Constraints

01

Unpaired Stylized Datasets

Trained on **unpaired** data (i.e. no direct correspondences between input and target) from **UTKFace** (ages 18–120) and **Ghibli-style** images self-generated via **Stable Diffusion** (i.e. limited due to GPU requirements).

02

Resolution and Resource Limitation

All images are resized to 128×128 due to the high computational cost of training GAN-based models.

03

Subjective Evaluation

No standard metrics used; evaluation is based on visual realism and user perception.

Literature Review

What Has Been Done

1. GANs (e.g., DCGAN, CycleGAN, Pix2Pix, StyleGAN)
2. Pioneered realistic image generation, unpaired image translation, and style control
3. StyleGAN introduced **AdaIN** for style manipulation via latent space disentanglement
4. CycleGAN used **cycle-consistency loss** for unpaired domain transfer
5. Attention mechanisms (e.g., in Transformers, basic spatial attention) improved global feature modeling

Existing Gaps

- Limited use of multiple attention blocks in GAN generators
- Struggles with:
 - Fine-grained style control
 - Identity preservation in unpaired settings
 - Spatial precision in transformations

01

Our Novel Approach

Utilisation of **Attention-Augmented GANs** with architectures such as **SE** Blocks, **CBAM** Attention Module, **Coordinate** Attention and **Multi-Head** Attention.

02

Our Novel Approach

Instance Normalization is used instead of Batch Norm, as it normalizes each sample independently—better preserving style cues and improving **stability** in GAN-based style transfer.

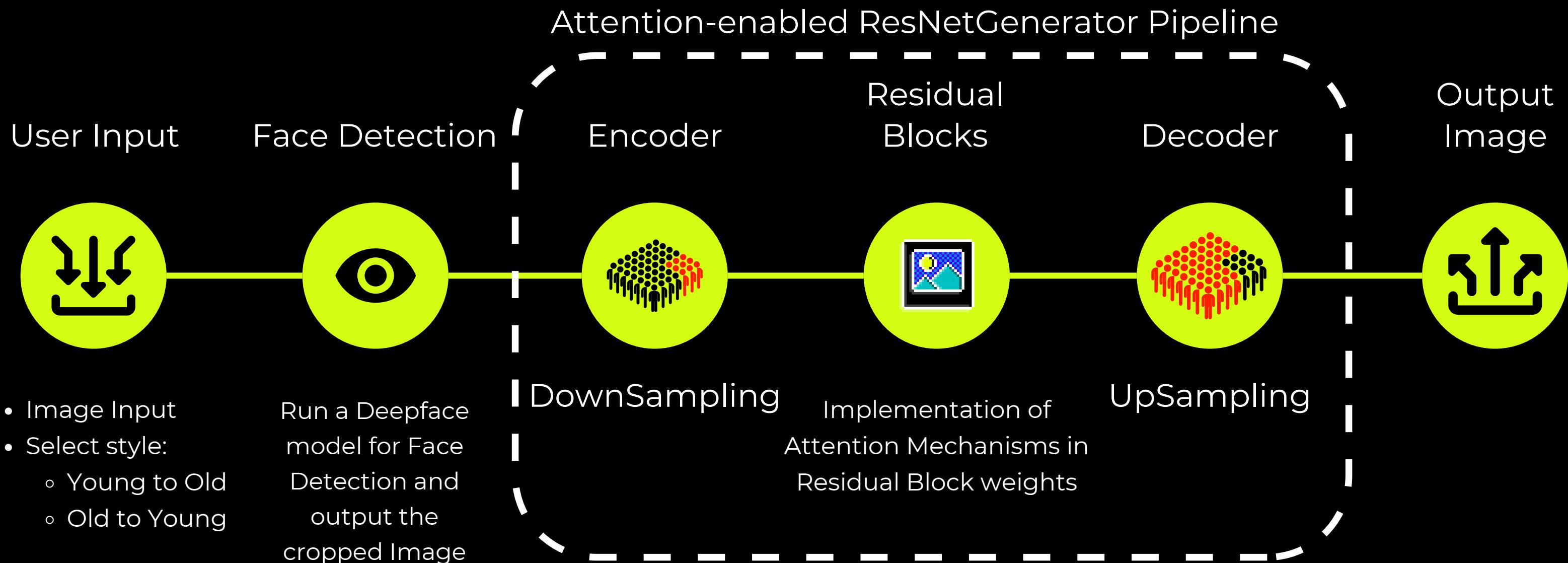
03

Our Novel Approach

PatchGAN Discriminator classifies overlapping **70×70** patches as real or fake, encouraging the **Generator** to produce fine-grained realism and stylization in generated faces.

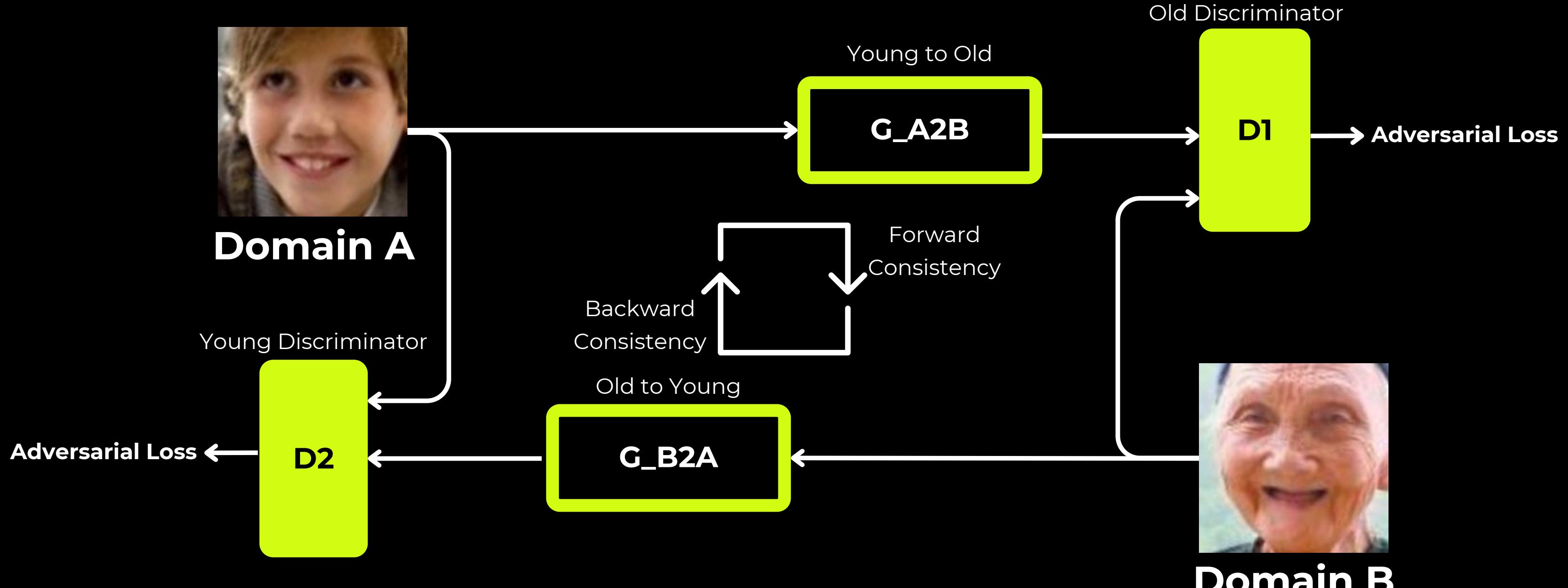
Proposed Method

A Walkthrough of the Entire **Age Manipulation** pipeline



Proposed Method

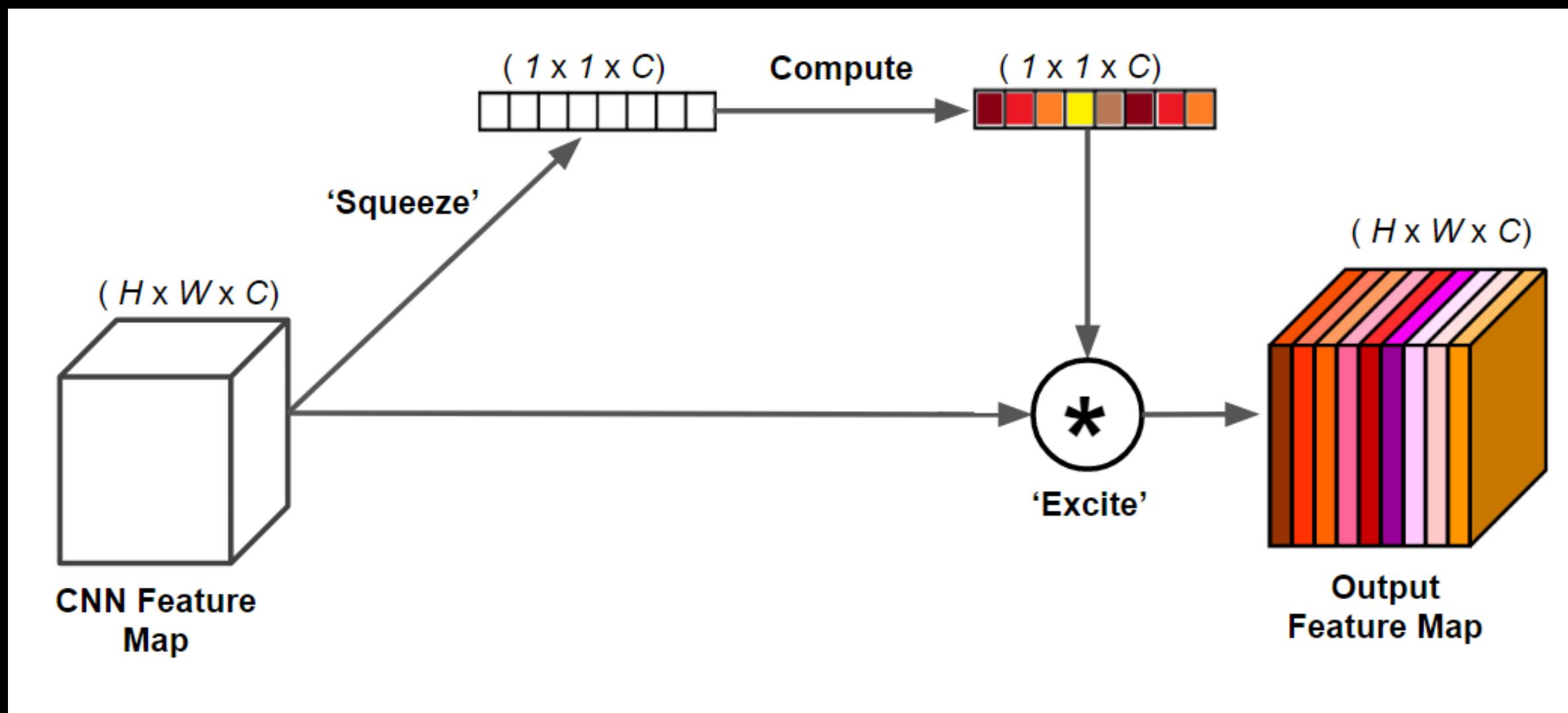
System Architecture - CycleGAN-Based Dual Translation Pipeline



- **Backbone:** ResNet-based Generator
- **Cycle Consistency** ensures reversibility
- **PatchGAN Discriminator:** Evaluates local 70×70 patches \rightarrow sharp details

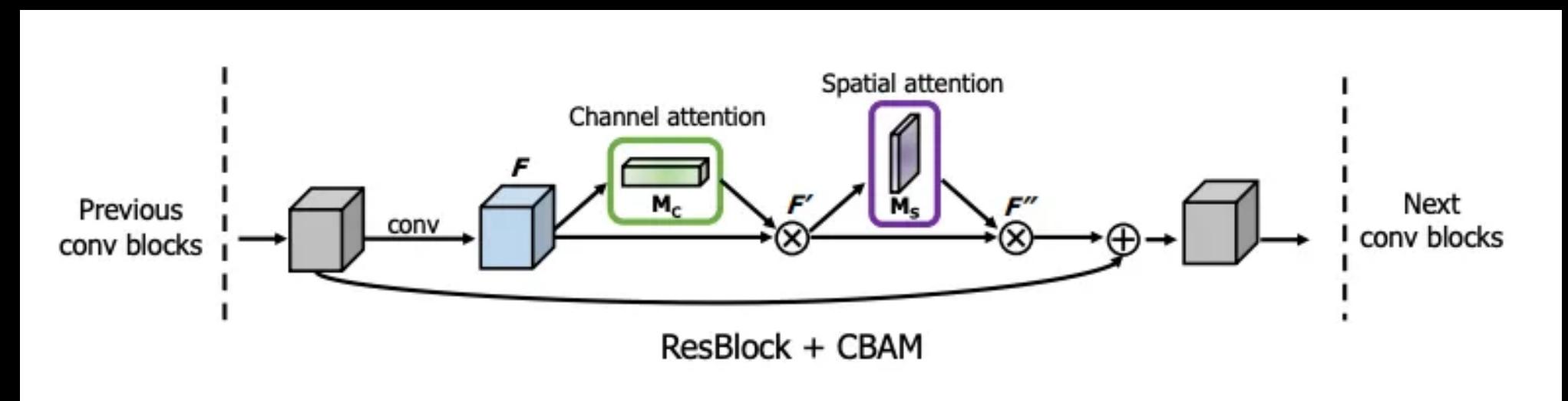
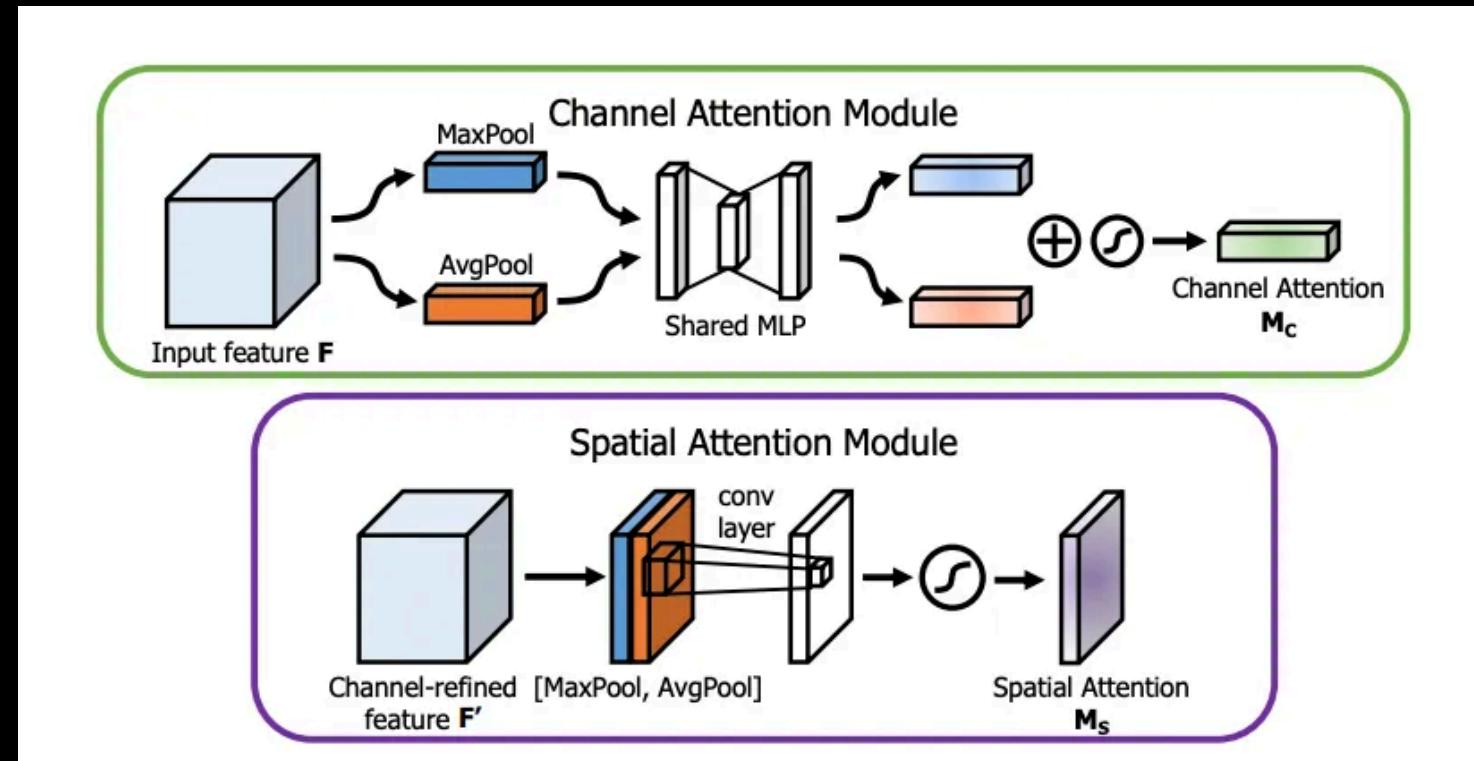
Proposed Method

Novel Approach - SE Block in ResNet Generator



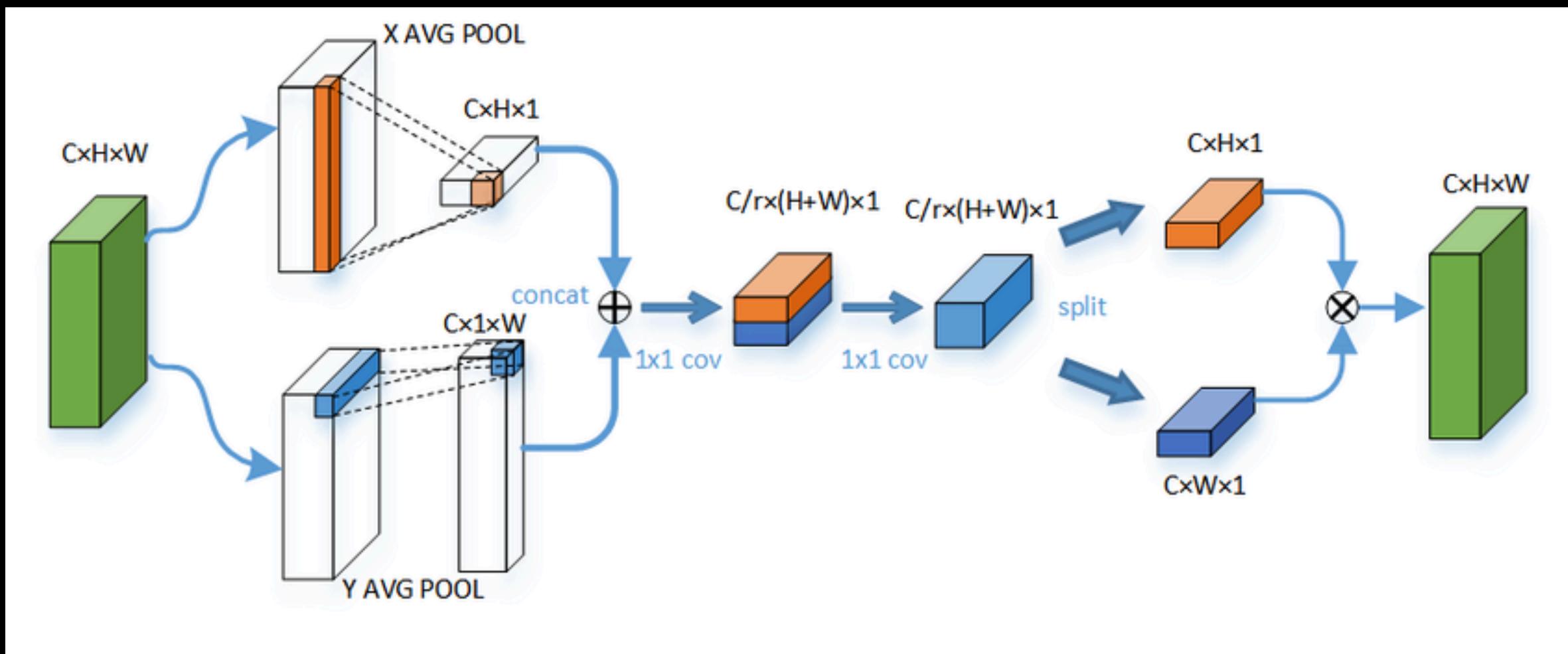
Proposed Method

Novel Approach - CBAM Attention in ResNet Generator



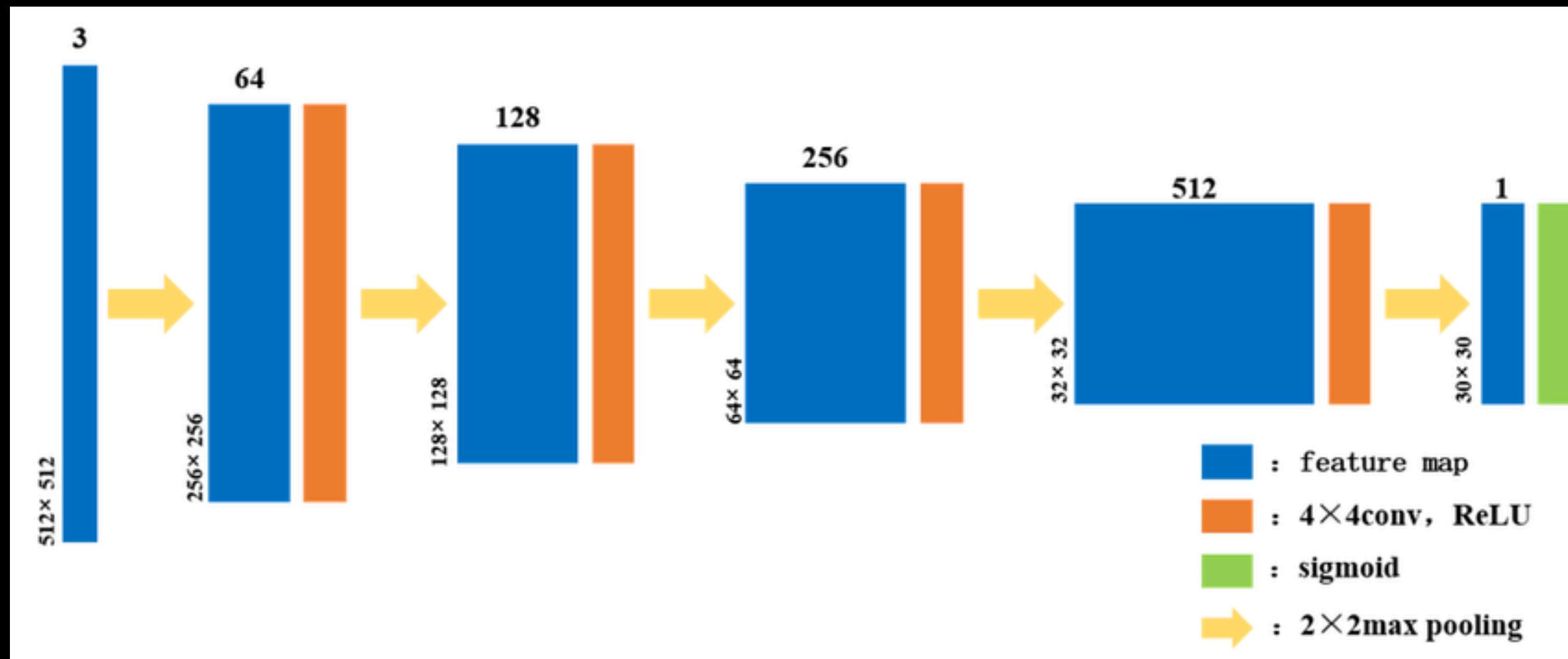
Proposed Method

Novel Approach - Coordinate Attention in ResNet Generator



Proposed Method

Novel Approach - PatchGAN Discriminator



Dataset Used

UTK Dataset

Large-Scale Face Dataset

Parameter	Value
NO. OF IMAGES	4000
RESOLUTION	128x128



Dataset Used

Custom Ghibli Images Dataset

Generated
StableDiffusion
Source Models

Parameter	Value
NO. OF IMAGES	4023
RESOLUTION	128x128



Evaluation and Results

Age Manipulation

Web-app deployed on Hugging Face for Demo

Evaluation based on **Visual Inspection**

Aging Deaging App

Input

Upload an image

Drag and drop file here
Limit 200MB per file • JPG, JPEG, PNG

Browse files

 AARUSH SINGH KUSHWAHA - IIT ...
133.3KB

Output

Select age conversion option

Old to Young
 Young to Old

Selected conversion: Young to Old

Generate

 Old you

Uploaded Image

Aging Deaging App

Input

Upload an image

Drag and drop file here
Limit 200MB per file • JPG, JPEG, PNG

Browse files

 Screenshot 2025-04-19 234011.png
420.0KB

Output

Select age conversion option

Old to Young
 Young to Old

Selected conversion: Young to Old

Generate

 Old you

Uploaded Image

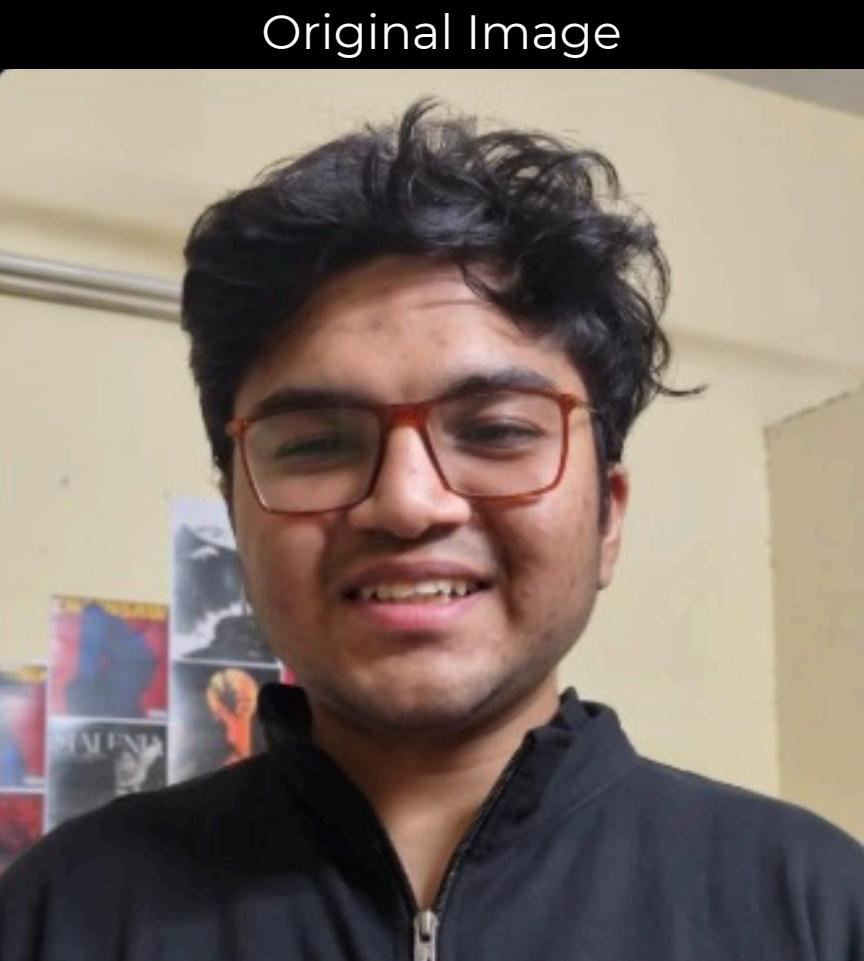
Evaluation and Results

Ghibli Style Transfer

Using Coordinate Attention Module with Residual Blocks in ResNet generator.

Web-app deployed on Hugging Face for Demo

Attention Mechanism	Avg. Generator Loss	Epochs
None	1.75	130
Coordinate	1.05	66



Original Image



Ghibli Image

Conclusion

We developed a **multi-style image generation framework** using **advanced GAN architectures** enhanced with attention mechanisms like ***Squeeze-and-Excitation, CBAM, and Coordinate Attention***. The system enables **unpaired translation** between *real and Ghibli-style faces*, and identity-preserving *age manipulation*.

Through **curated datasets**, targeted augmentation, and **architectural innovations**, we addressed challenges related to ***domain mismatch*** and ***style fidelity***. Despite minor artefacts in occluded or low-light conditions, our model consistently produces visually convincing and identity-aware outputs. This work underscores the power of **attention-augmented GANs** for creative style and attribute transfer in digital art and human-centered design.

THANK YOU

