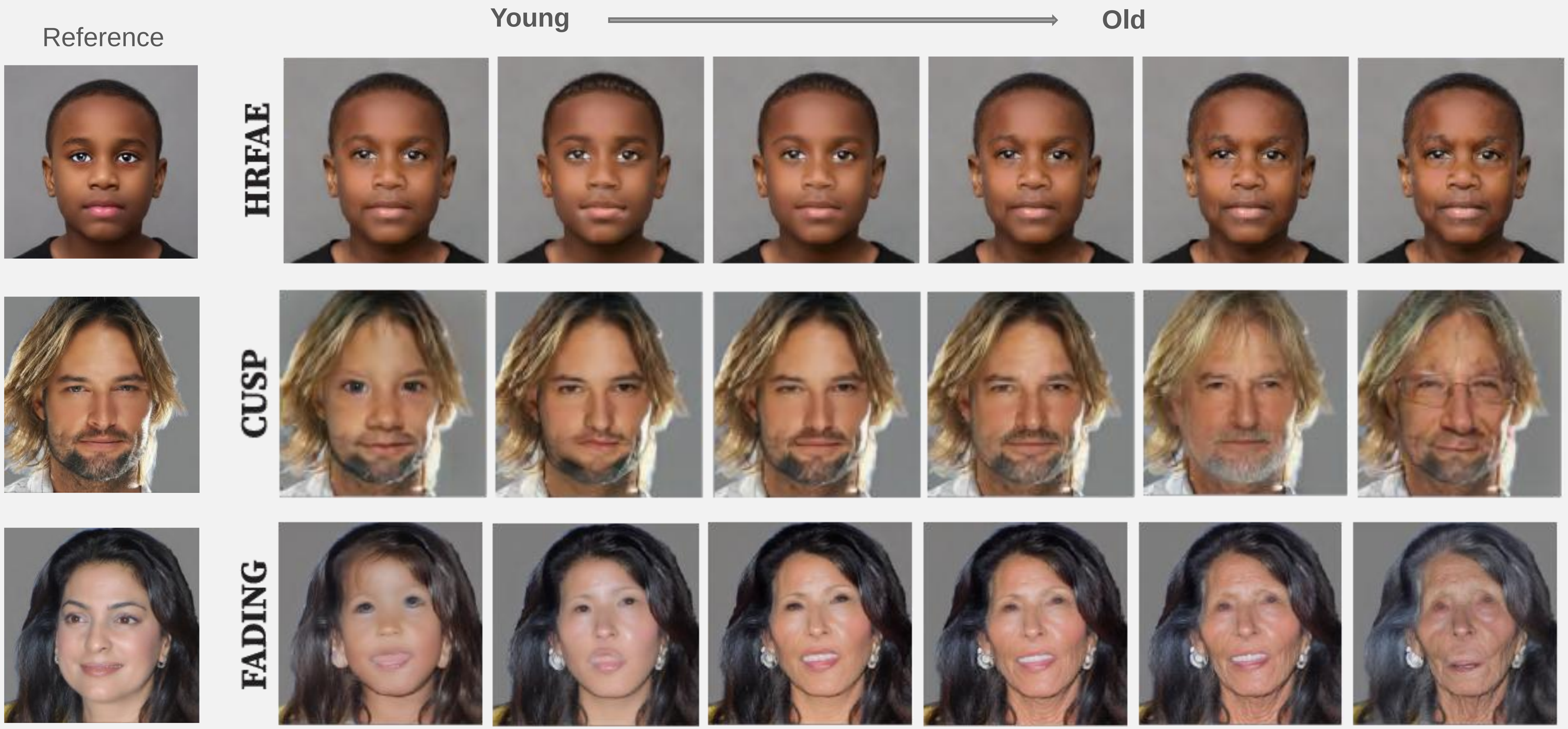
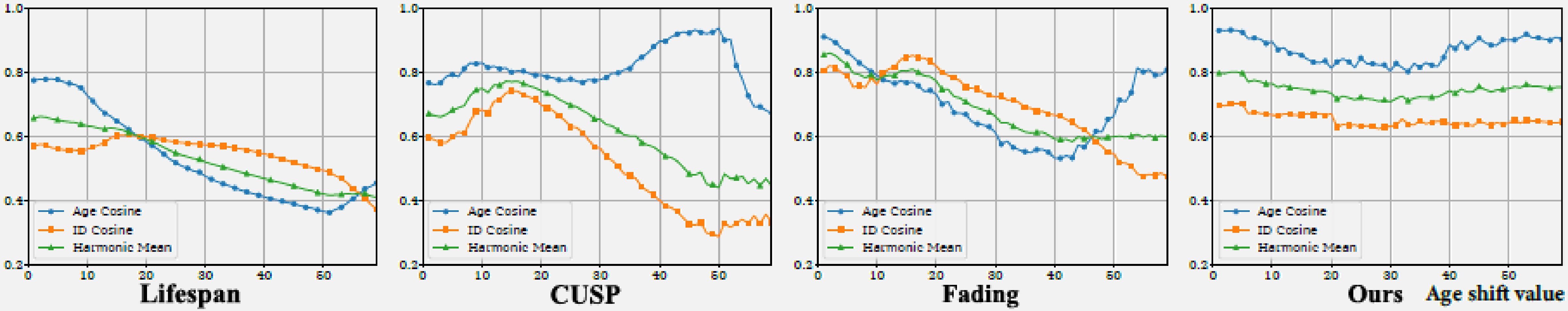
01 Background | Exist Method

- Existing methods either prioritize age transformation **at the expense of identity consistency or vice versa**.
- The generated results is not able to able to imitate the **natural change**, such as hair color, beard.



01 Background | Exist Method

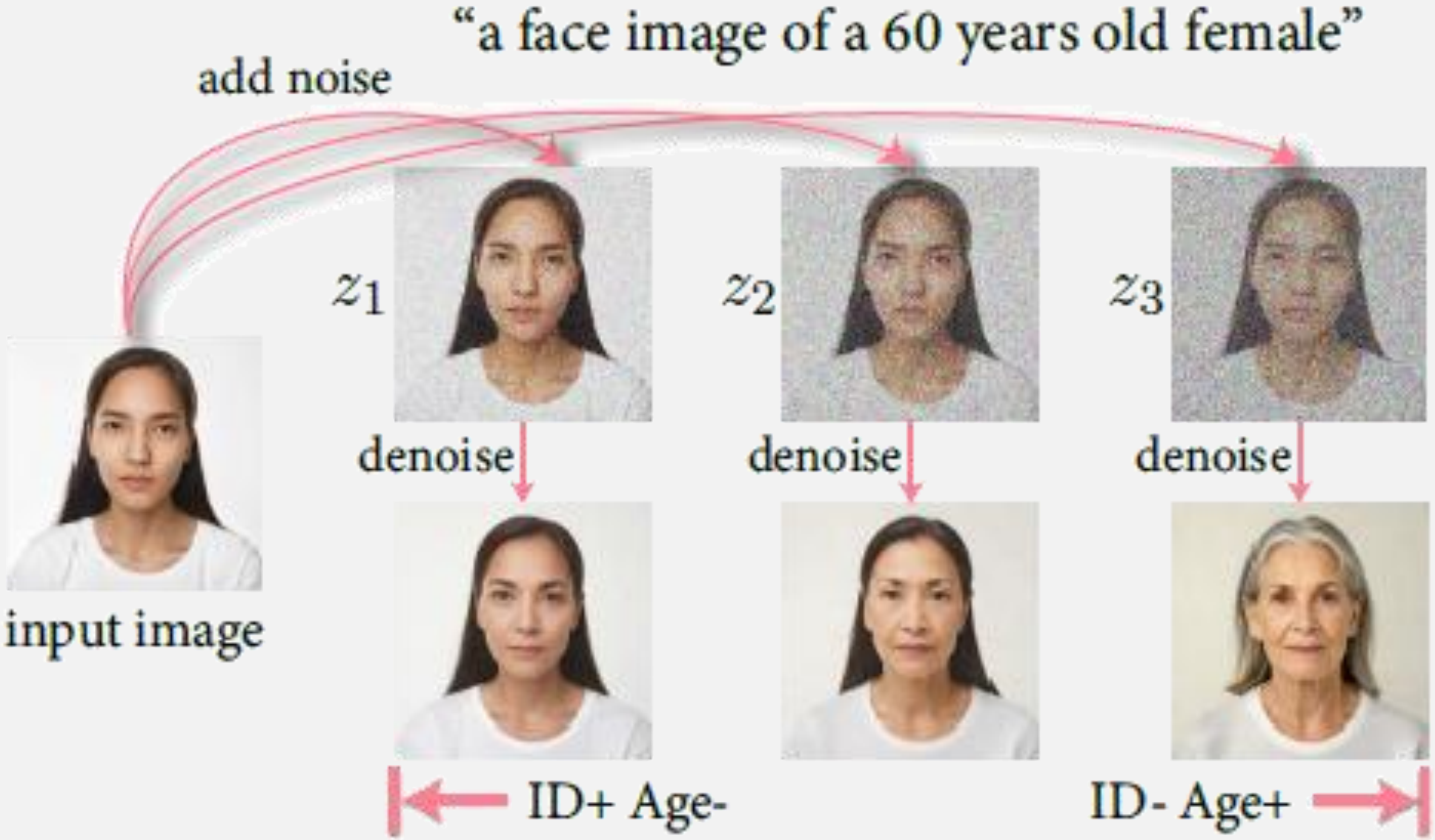


We compute the **Age/ID cosine similarities** over 100 human faces across 1-60 age shift values and the corresponding harmonic means. **Existing approaches tend to favor either age accuracy or identity consistency**, resulting in imbalanced performance across the entire lifespan ages.

In contrast, our method Cradle2Cane achieves a better balance between the two objectives.

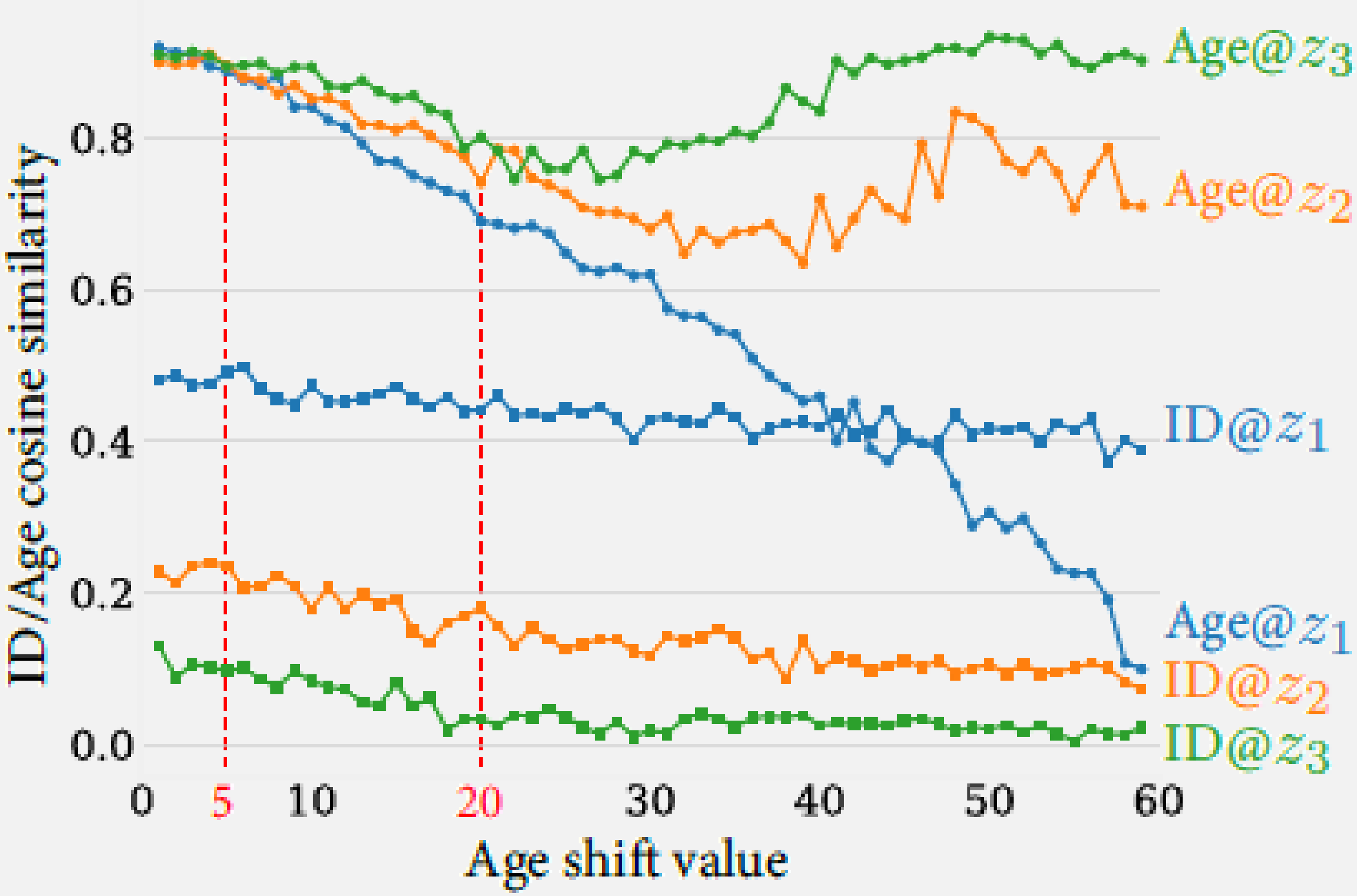
01 Background | SDXL-turbo

- We illustrate the effects of injecting three different levels of noise into the input image
- As visually evident, higher noise levels lead to more pronounced age transformations **at the cost of reduced identity preservation.**



01 Background | SDXL-turbo

- We present a statistical analysis on 100 human faces, that quantitatively demonstrates the Age-ID trade-off inherent in face aging tasks. Specifically, we evaluate three representative noise injection levels and measure their corresponding impacts on age accuracy and identity consistency.

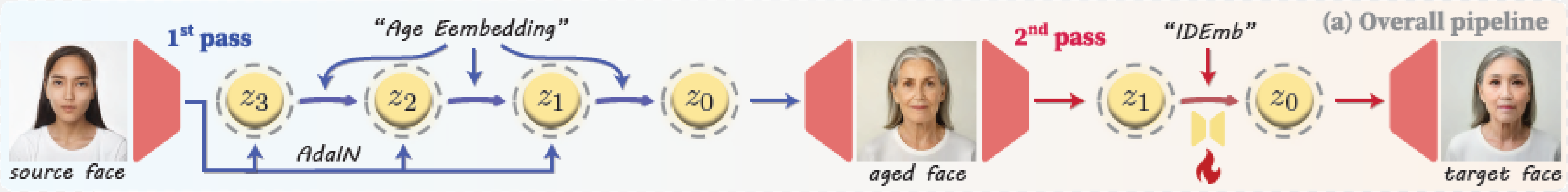


02 Solution | Pipeline

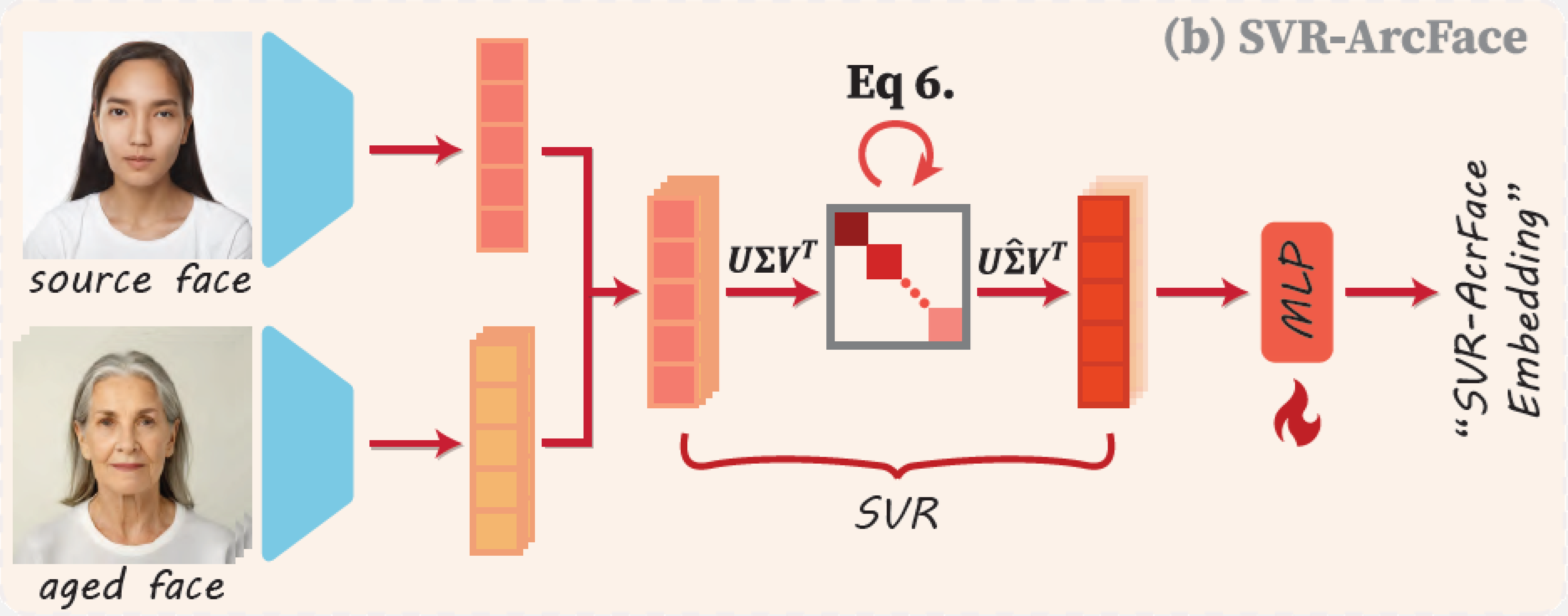
Our method Cradle2Cane consists of two passes:

- The first pass employs adaptive noise injection (AdaNI) to enhance age accuracy. For AdaNI injection, we divide the age transformation magnitude into three categories, using ages 5 and 20 as boundaries.

- The second pass incorporates identity-aware embeddings (IDEmb), including SVR-ArcFace and Rotate-CLIP embeddings, to improve identity consistency.

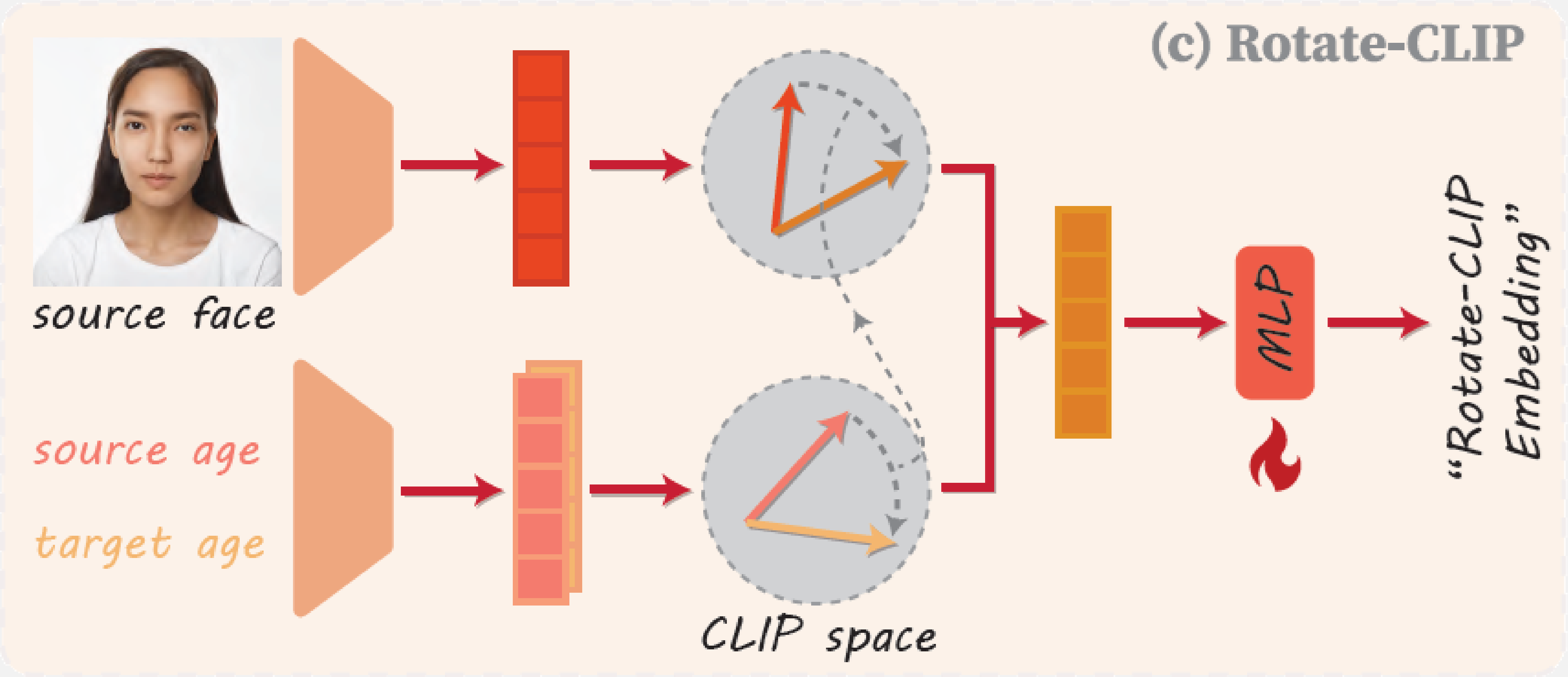


02 Solution | SVR-Arafcae



This method enhances identity features by extracting **ArcFace embeddings** from a source face and several synthetically aged versions of it.

02 Solution | Rotate-Clip



Rotate-CLIP adjusts the age characteristics of a **face image embedding** by calculating a rotational projection between the text embeddings of the source and target ages.

02 Solution | Loss

We jointly optimize identity, age, and quality using:

- **Identity Loss**: MS-SSIM + ArcFace cosine similarity

$$\mathcal{L}_{\text{id}} = \lambda_1 \cdot (1 - \text{MS-SSIM}(\mathbf{x}_a, \mathbf{x}_b)) + \lambda_2 \cdot (1 - \cos(f_{\text{Arc}}(\mathbf{x}_a), f_{\text{Arc}}(\mathbf{x}_b)))$$

- **Age Loss**: MiVOLO embedding + age prediction error

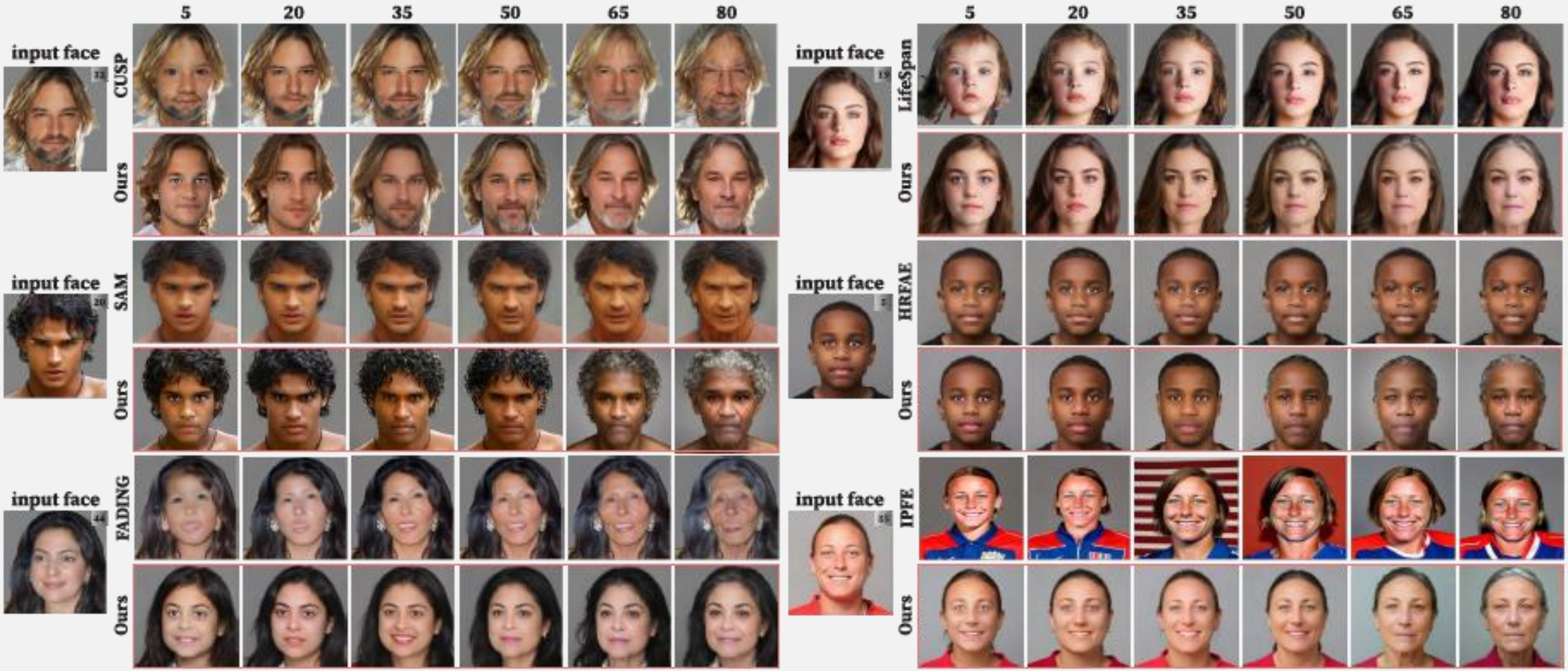
$$\mathcal{L}_{\text{age}} = \lambda_3 \cdot (1 - \cos(f_{\text{Mi}}(\hat{\mathbf{x}}_b), f_{\text{Mi}}(\mathbf{x}_b))) + \lambda_4 \cdot \|g_{\text{Mi}}(\mathbf{x}_b) - b\|_2^2$$

- **Quality Loss**: LPIPS + GAN loss

$$\mathcal{L}_{\text{per}} = \lambda_5 \cdot \text{LPIPS}(\mathbf{x}_a, \mathbf{x}_b) + \lambda_6 \cdot \mathcal{L}_{\text{GAN}}(\mathbf{x}_b)$$

Total loss: weighted sum of all components

03 Results | Qualitative comparison

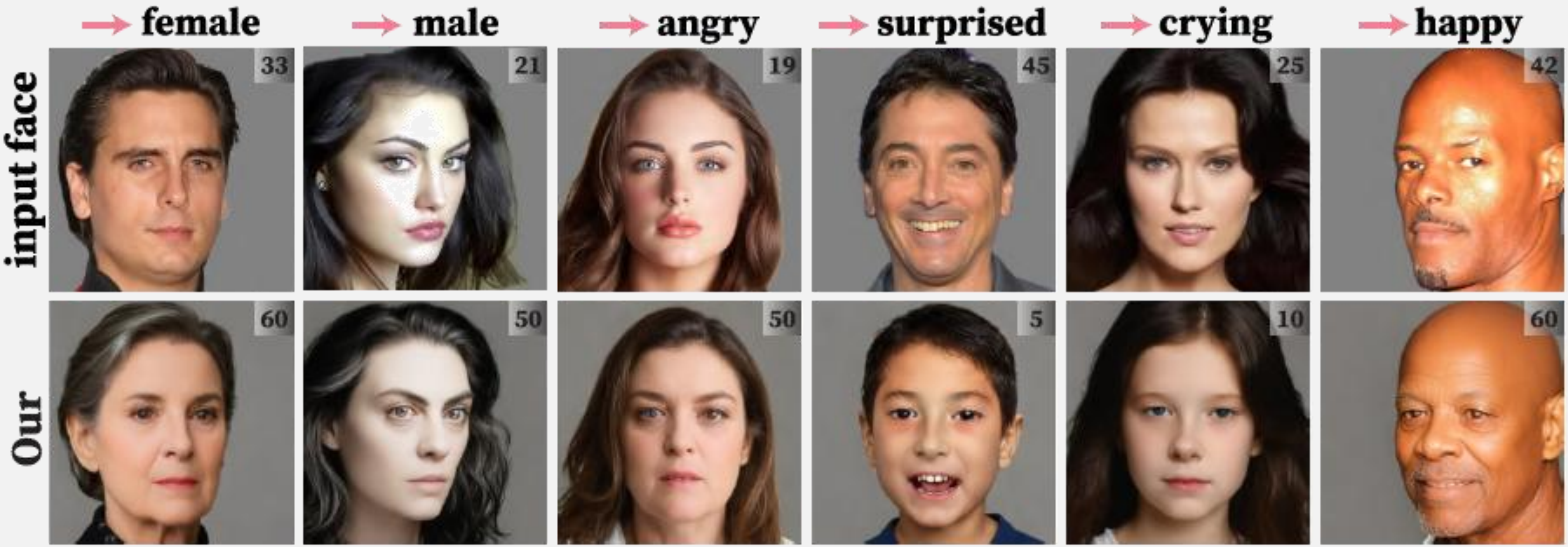


Qualitative comparison with existing face aging methods across lifespan ages.

03 Results | In-the-wild image

While applying to in-the-wild real human faces, Cradle2Cane demonstrates better performance while the existing methods often fail.

03 Results | Face edit



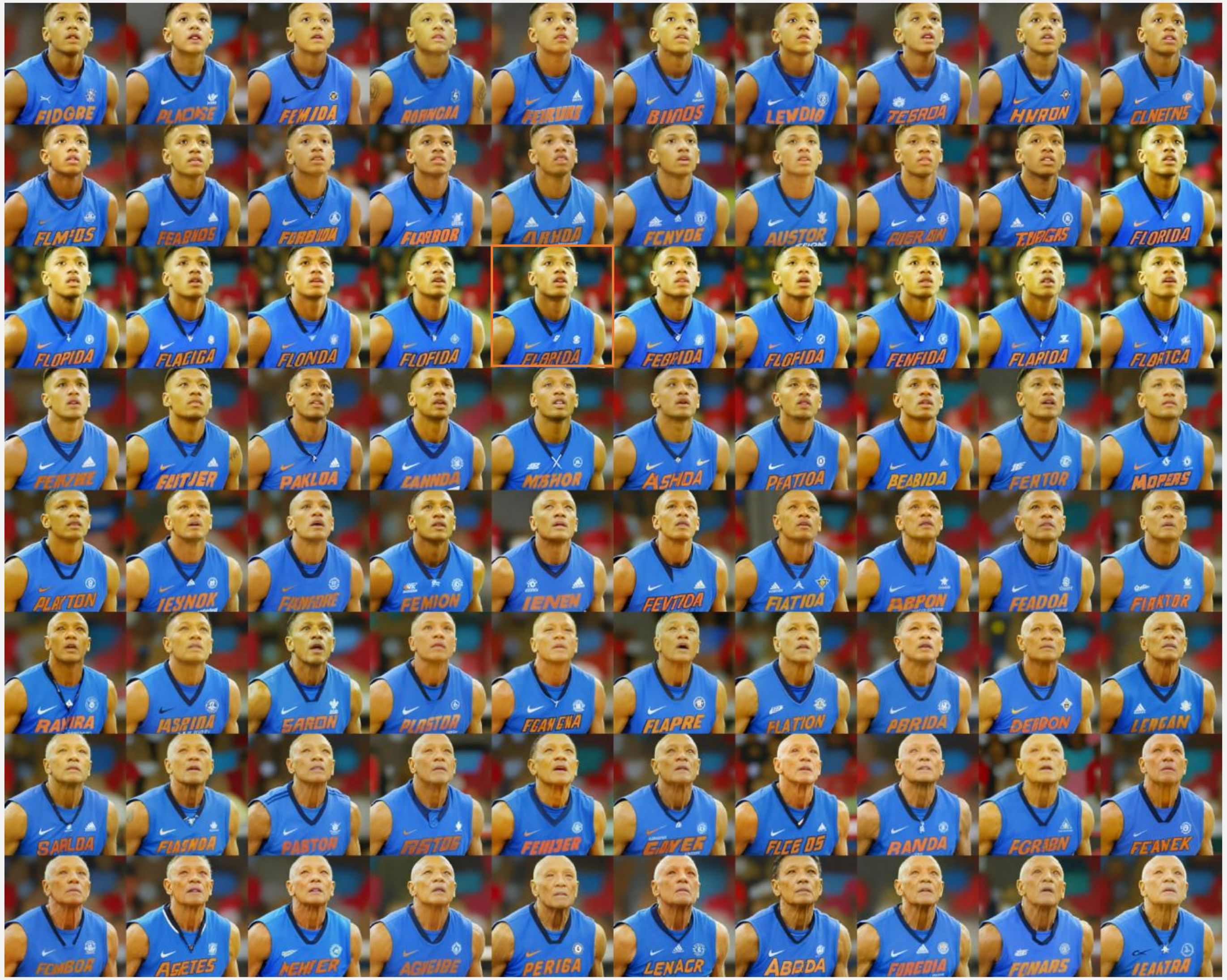
Our Cradle2Cane can also be applied to modify gender and emotion attributes while performing age transformation on human faces.

03 Results | All age

Samsung Research

Face aging results from 1 to 80 years old. Reference images are marked in orange.

03 Results | All age



Face aging results from 1 to 80 years old of in-the-wild images. Reference images are marked in orange.

03 Results | Diverse reference image

Samsung Research

Aging results generated from diverse reference ages. For each reference image, we synthesize faces at six target ages: 5, 20, 35, 50, 65, and 80.

03 Results | Diverse reference image