Exploring Techniques For Age Progression In Facial Images: A Comprehensive Survey

Yugchhaya Galphat
Department of Computer Engineering
Vivekanand Education Society's
Institute of Technology
Chembur, India

Chetaniya Bajaj
Department of Computer Engineering
Vivekanand Education Society's
Institute of Technology
Chembur, India

Gaurav Amarnani
Department of Computer Engineering
Vivekanand Education Society's
Institute of Technology
Chembur, India

Kaplesh Mulchandani
Department of Computer Engineering
Vivekanand Education Society's
Institute of Technology
Chembur, India

Jayesh Repale
Department of Computer Engineering
Vivekanand Education Society's
Institute of Technology
Chembur, India

***Abstract*—Facial age progression has emerged as a critical domain in computer vision, finding applications across various fields including forensics and entertainment. This survey thoroughly investigates the methodologies utilized in age progression, encompassing both traditional and deep learning approaches. Beginning with an examination of traditional approaches, we delineate their strengths and limitations. Subsequently, focusing on deep learning methodologies, particularly generative adversarial networks (GANs), which have significantly enhanced the accuracy and realism of age progression. Additionally, addressing challenges associated with age progression, including dataset biases, privacy concerns, and ethical considerations, emphasizing the imperative to mitigate these challenges for the responsible deployment of age progression technologies. Finally, providing an outlook on future directions, discussing emerging trends such as novel data augmentation techniques, improved interpretability of deep learning models, and considerations for the societal implications of widespread age progression applications.**

Keywords—Age Progression, Image to Image Translation, Generative Adversarial Networks

I. Introduction

Facial age progression, the process of synthetically rendering an individual's face at different ages, has emerged as a crucial area of research in computer vision and pattern recognition, with far-reaching implications across various domains. This technology holds significant potential in fields such as forensics, biometrics, and human-computer interaction, where the ability to accurately predict an individual's facial appearance at different ages can prove invaluable.

Facial age progression technology offers a plethora of real-world applications and relevance across diverse fields due to its capability to synthetically depict how an individual's face may change over time. In law enforcement, this technology serves as a crucial tool for identifying missing persons or suspects, as well as aiding in their recovery. By generating age-progressed images, law enforcement agencies can provide updated representations of individuals' potential appearances at different ages, thereby enhancing the chances of locating and rescuing missing persons or apprehending suspects who may have altered their appearance over time. Furthermore, in forensic investigations, facial age progression plays a pivotal role in reconstructing facial features of suspects or victims based on limited information, such as aged photographs or witness descriptions. This aids investigators in generating leads and narrowing down potential suspects, contributing valuable insights to criminal cases.

Beyond law enforcement and forensics, facial age progression technology finds applications in biometrics and human-computer interaction. In biometrics, accurately predicting an individual's facial appearance at different ages enhances the effectiveness of facial recognition systems for identification and authentication purposes. By accounting for age-related changes in facial features, such as wrinkles, sagging skin, and changes in facial structure, biometric systems can maintain accuracy over time, ensuring reliable identification across various age groups. Additionally, in human-computer interaction, facial age progression facilitates the development of age-aware technologies that adapt to users' changing needs and preferences as they age. This enables the creation of more personalized and intuitive user interfaces in applications ranging from digital assistants to virtual reality environments, enhancing user experience and engagement.

Moreover, facial age progression technology holds significant relevance in healthcare and aging research. By accurately predicting facial aging trajectories, clinicians and researchers can gain insights into the effects of aging on an individual's health and well-being. This includes assessing age-related changes in facial morphology associated with certain medical conditions, monitoring disease progression, and developing personalized interventions for age-related health concerns. Additionally, facial age progression aids in raising awareness about the importance of healthy aging and preventive care by visually demonstrating the potential effects of lifestyle choices, such as sun exposure, smoking, and skincare routines, on facial aging.

Facial age progression is a challenging task due to the complex and highly individualized nature of facial aging. The process involves modeling intricate transformations in facial features, textures, and shapes that occur as individuals age. These changes are influenced by various factors, including genetics, environmental conditions, and lifestyle choices, making it difficult to capture the nuances of aging accurately.

Researchers have explored various methods to tackle facial age progression, each with its unique strengths and limitations. The success of these methods heavily relies on the availability of high-quality and diverse datasets. Several publicly available datasets have been widely used in facial age progression research. However, these datasets often suffer from limitations, including biases in terms of age, ethnicity, and demographic representation, posing significant challenges for researchers in terms of data collection and curation.

The rest of the research paper is structured as follows: Section II provides an in-depth examination of the state-of-the-art methodologies used in facial age progression. It explores both feature-based techniques and deep generative models, delving into their complexities. Section III conducts a critical evaluation of the existing datasets employed for training and validation, highlighting their strengths, limitations, and potential biases. Additionally, ethical considerations surrounding facial age progression datasets are examined, emphasizing the necessity of responsible deployment and the mitigation of risks related to privacy and bias. Section IV centers on the diverse evaluation methods available to assess the accuracy of the algorithms discussed in the paper. Section V concludes the whole paper while Section VI mentions the current challenges faced in face age progression and potential for future work.

II. Methods and Approaches

In the realm of facial age progression, two prominent methodologies emerge: the traditional feature-based approach and the deep learning approach, often powered by Generative Adversarial Networks (GANs). These methodologies embody distinct strategies for predicting how a person's face evolves over time, each offering unique advantages and applications. The feature-based method entails the manual extraction and analysis of specific facial features, such as wrinkles and contours, to estimate aging effects based on statistical models and expert insights. In contrast, the deep learning approach harnesses the power of neural networks, particularly GANs, to autonomously generate realistic aged facial images by learning from large datasets. Fig. 1 illustrates examples of age-progressed images generated using cGAN.

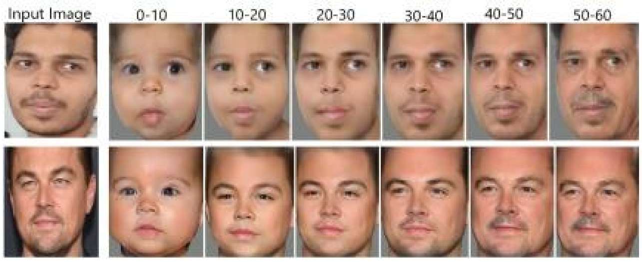

Fig. 1. Face Age Progression

A. Features Based approach

Features based approaches rely on established methods that have been foundational in understanding how a person's face changes with time. These methods often involve the manual or semi-automated analysis of specific facial features, such as wrinkles, skin texture, and facial contours, that are known to undergo alterations with age. Feature-based methods, for instance, extract key landmarks and descriptors from facial images to quantify aging-related changes statistically. They may employ techniques like linear regression or principal component analysis to model the relationship between these features and age labels, providing insights into the aging process. Additionally, template-based methods utilize age-labeled facial images as references for generating aged versions of new faces. By aligning facial features between templates and target images, these methods deform the facial structure to reflect typical aging patterns observed in the templates.

I. Kemelmacher-Shlizerman [1] introduced an innovative, fully automated method for facial age progression, specifically focusing on the challenge of aging young children. This methodology presents a comprehensive approach to automatically progressing facial images across different age ranges. It initiates with the assembly of a substantial dataset covering ages ranging from infancy to adulthood, sourced from diverse online platforms. Leveraging this dataset, the approach involves the creation of aligned and relightable average images for each age group, facilitating the incorporation of realistic shading effects. To generate relightable average images, techniques such as singular value decomposition (SVD) applied to flow-aligned images and the computation of rank-4 approximations are employed. Optical flow methodologies are then utilized to estimate the flow between different age clusters, enabling the computation of age transformations. Additionally, the methodology incorporates algorithms for adjusting aspect ratios and varying skin tones. Because the algorithm's performance heavily depends on the quality and diversity of the training dataset, without sufficient representation across demographics, ethnicities, and facial features the algorithm may struggle to accurately capture age progression for different groups.

X. Shu [2] introduces a novel approach consisting of offline training and online synthesis phases. Short-term aging pairs are collected from various databases to create aging dictionaries covering diverse aging characteristics. A personality-aware coupled dictionary learning model is developed, considering individualized details like birthmarks and scars. Principal Component Analysis (PCA) reduces data dimensionality. In the online synthesis phase, aging faces are rendered iteratively in successive age groups. Sparse coefficients and personalized layers are updated iteratively for optimal aging results, reflecting a short-term coupled learning approach due to limited long-term aging data. The process involves iterative updates until convergence, aiming for consistency with the training phase and producing personalized and natural-looking aging faces through data-driven learning and iterative optimization techniques.

X. Shu [3] presents notable improvements over prior work by addressing the practical challenge of obtaining long-term aging sequences. It achieves this by utilizing dense short-term aging pairs, enhancing the method's applicability for real-world scenarios. The introduction of a bi-level dictionary learning approach, incorporating personalized layers, enhances the capture of individual-specific aging characteristics, resulting in more realistic and personalized aging faces. Furthermore, optimizations in the age progression synthesis process reduce computational complexity, leading to faster convergence and requiring fewer iterations for generating aging faces. Collectively, these enhancements enhance the accuracy, realism, and efficiency of the age progression method, making it suitable for various applications in facial aging analysis and synthesis.

The approach presented by *Shintaro* [4] is for altering facial age in videos while accounting for changes in facial expressions. The methodology primarily focuses on synthesizing aging facial videos while considering the temporal dynamics of wrinkles induced by facial expressions. Initially, the process involves aligning facial expressions between the target video and a database of videos exhibiting similar expressions. This alignment is facilitated using local binary patterns (LBPs) as descriptors for facial expressions, complemented by dynamic time warping (DTW) to ensure temporal coherence. Subsequently, each frame of the database undergoes deformation to match the shape of the corresponding frame in the target video, employing radial basis functions (RBF) interpolation. Additionally, the methodology encompasses the synthesis of an aged video by blending textures from chosen individuals in the database, taking into account attributes like color and wrinkles. The preservation of wrinkles emerges as a critical aspect, entailing the identification of optimal facial regions and the blending of expressions and neutral faces to uphold natural wrinkle depth.

Riaz [5] outlines their methodology for constructing and simulating gender-specific 3D aging models. Initially, they convert 2D images from various datasets into 3D frontal-face models, which serve as the basis for creating aging spaces for both shape and texture. The construction of the 3D aging models involves mapping 2D images to 3D face models using facial landmarks and active shape modeling techniques. Principal Component Analysis (PCA) is separately applied to male and female 3D faces with available age labels to construct shape and texture models. Interpolation methods are utilized to address missing data in the aging patterns. Aging simulation entails fitting input images to the shape aging space and generating corresponding texture patterns. Subsequently, color and flare correction techniques are applied to ensure consistency between the simulated images and the original background. Finally, the age-simulated images are composited onto the original backgrounds using landmark points and an edge-smoothing algorithm to achieve realistic integration.

Elmahmudi [6] proposes an approach that relies on ethnicity-based face templates constructed from age, gender, color, and texture characteristics extracted from faces of principal ethnic groups. The system consists of two main components: firstly, a mathematical method for constructing ethnicity-specific aging templates using average faces; and secondly, the application of these templates to target faces for age generation, incorporating control parameters for color and texture. Additionally, a framework for verifying the accuracy of generated faces through similarity comparison using Convolutional Neural Networks (CNNs) is proposed. Data collection involved multiple phases, including the gathering of images from diverse sources and their normalization to a reference frame. The methodology encompasses techniques for facial landmark detection, template generation, wrinkle mapping, and labeling. Age progression or regression is achieved through image morphing and cross-dissolving, with parameters controlling shape and color fidelity.

While traditional approaches have offered valuable insights into facial aging dynamics, they often rely on simplifications and may struggle to capture the full complexity of age-related changes. As a result, contemporary methods, particularly those based on machine learning and deep learning, have emerged to address these limitations and enhance the accuracy and realism of facial age progression models.

B. Generative Adversarial Networks

In contrast to feature-based approaches, GANs [7] have emerged as a powerful alternative in the realm of face age progression. GANs leverage the concept of adversarial training, where two neural networks, namely the generator and the discriminator, are pitted against each other in a competitive manner. In the context of face age progression, the generator network is tasked with synthesizing realistic images of faces at different ages, while the discriminator network aims to distinguish between real and generated images. In simpler terms, they engage in a min max game. Explained further using equation (1).

$$min\, G\, max\, D\, V(D, G) \;=\; E_{x\sim Pdata(x)}[log\, D(x)] \;+\; E_{z\sim Pz(z)}[log(1 \;-\; D(G(z)))] \tag{1}$$

The Generator G takes random noise samples z from a prior distribution $pz(z)$ and transforms them into data samples $G(z;\, \theta g)$, where θg represents the parameters of the generator. Essentially, $G(z;\, \theta g)$ maps the noise space to the data space, generating synthetic data that ideally mimics the distribution of real data.

The Discriminator D evaluates data samples x and assigns a probability $D(x;\, \theta d)$ representing the likelihood that the sample came from the real data distribution rather than from the generator's distribution. It's represented as a function $D(x;\, \theta d)$, where θd denotes the parameters of the discriminator.

In the training process, the discriminator is trained to maximize the probability of correctly labeling both real data samples and generated samples. This is captured by the term $[log\, D(x)]$ in the equation, where $D(x)$ represents the probability assigned by the discriminator to real data samples. The goal is for $D(x)$ to approach 1 for real data.

Simultaneously, the generator is trained to minimize the discriminator's ability to correctly classify its generated samples as fake. This is achieved by minimizing the log probability of the discriminator outputting 1 when given generated data. This objective is represented by the term $[log(1 \;-\; D(G(z)))]$, where $G(z)$ represents the generated data and $D(G(z))$ represents the probability assigned by the discriminator to generated samples. The generator aims to make $D(G(z))$ approach 0, indicating that the discriminator is unable to distinguish between real and generated data effectively. This adversarial process drives both networks to improve iteratively until an equilibrium is reached, ideally resulting in the generator producing high-quality synthetic data samples.

The Conditional Adversarial Autoencoder (CAAE) network, introduced in Paper [8], Unlike traditional GANs which often suffer from instability and produce noisy outputs, integrates an encoder, generator, and discriminators to effectively learn a manifold of facial features. This manifold facilitates smooth transitions in age while preserving unique individual characteristics. By converting input facial images into feature vectors using the encoder, and then combining them with age labels for guidance during generation by the generator, the CAAE ensures accurate and personalized image synthesis. Furthermore, the objective function of the CAAE balances several factors, including reconstruction accuracy, distribution regularity, and image realism, to optimize the quality of the generated images. What distinguishes the CAAE from other generative models, such as Variational Autoencoders (VAEs) [9] and Adversarial Autoencoders (AAEs) [10], is its innovative use of discriminators on both the encoder and generator. This unique approach sets the CAAE apart, making it a promising upgrade over traditional GANs.

G. Antipov [11] introduces two significant contributions: Age-cGAN, an iteration of GANs tailored to craft impeccable synthetic images categorized by age, and an innovative method for optimizing latent vectors to uphold the distinct identity of the original person during facial reconstruction. The proposed approach to facial aging unfolds in a methodical two-step process: initially approximating the latent vector, followed by optimization. The intricacies of Age-cGAN are detail\\ed, showcasing its adept use of conditional data to generate images falling within predefined age brackets. Additionally, the technique for approximating facial reconstruction places a strong emphasis on preserving identity, effectively mitigating issues like blurriness and irrelevant details.

X. Tang [12] introduces Identity-Preserved Conditional Generative Adversarial Networks (IPCGANs) for facial aging. It proposes a method to divide facial images into age groups and generate lifelike faces within specific age ranges while maintaining the original identity. IPCGANs consist of three essential components: CGANs for realistic face synthesis, an Identity-Preserved module to retain identity features, and an age classifier to ensure the generated faces match the desired age group. The CGANs module employs Conditional Least Squares Generative Adversarial Networks (LSGANs) to produce high-quality images, while the Identity-Preserved module utilizes perceptual loss to safeguard identity characteristics. Additionally, an age classification module is incorporated to enforce age coherence in the generated faces. The combined objective function integrates adversarial loss, identity preservation loss, and age classification loss. The network architecture comprises a generator, discriminator, and age classification network, each customized for its specific role. The generator employs residual blocks and integrates age conditions before the initial convolution layer, while the discriminator's architecture draws inspiration from invertible conditional GANs, incorporating condition injection after the first convolution layer. Lastly, the age classification network, based on AlexNet [13], is enhanced with fully connected layers and dropout to prevent overfitting.

H. Yang [14] combines GANs, age-specific feature extraction, and identity preservation techniques. The framework involves a generator network, encoder-decoder architecture, and discriminator network. Identity preservation is addressed by measuring the input-output distance in a feature space sensitive to identity changes. The loss function comprises adversarial loss, pixel-wise loss for color aberration, and identity loss, with weighting factors regulating their contributions.

Y. Liu [15] introduces a framework based on GANs to tackle the challenge of learning age-specific transformations in unpaired face image datasets. This framework consists of a generator network and a discriminator based on wavelet analysis. The generator is designed to incorporate both low-level image details and high-level semantic facial attributes, aiming to stabilize the translation process between young and aged faces. To capture age-related textures effectively, the model utilizes wavelet packet transform, enabling multi-scale texture analysis while keeping computational demands low. The training process involves minimizing adversarial loss, pixel loss to maintain image-level consistency, and identity loss to preserve personalized facial features.

The S2GAN framework introduced by *Z. He* [16] is a novel approach to face aging by incorporating personalized aging factors and age-specific transformations. It comprises three main components: establishing personalized aging bases using a deep encoder, transforming these bases into age representations for different age groups, and decoding these representations to generate aged faces. Unlike traditional GANs, S2GAN optimizes the model with three objectives: age group classification loss, L1 reconstruction loss for identity preservation, and adversarial loss for fidelity. The age-specific transforms, shared across individuals but distinct for different ages, enable continuous aging interpolations, providing more natural and practical results compared to methods with discrete age groups. Furthermore, S2GAN offers lower computational cost and storage requirements by utilizing a single model for all target ages and sharing the personalized basis across ages, making it more efficient and scalable for face aging tasks.

The approach by *M. Sheng* [17] aims to improve the precision of face aging through the utilization of a conditional GAN under the supervision of Ranking-CNN. It categorizes facial images into five distinct age groups and employs a combination of a generator, discriminator, pretrained Alexnet network [13], and Ranking-CNN within its framework. Diverging from conventional GANs, this technique integrates a perceptual loss mechanism to uphold identity preservation and integrates Ranking-CNN to enforce more rigorous age constraints on the generator. Additionally, it adopts conditional Least Squares GAN (LSGAN) [18] for adversarial loss implementation to ensure consistent and high-quality image synthesis.

Shi [18] introduces The CAN-GAN framework. At its core lies the Conditional Attention Normalization (CAN) which is integrated into both the generator and discriminator units. Unlike conventional approaches, CAN utilizes age disparities instead of age labels for normalization, thereby adeptly capturing age-related facial characteristics while diminishing irrelevant ones. Furthermore, the model integrates the Conditional Age Attribute Classifier (CAAC) to assess the significance of individual facial attributes in age determination, thereby enhancing accuracy. The model's objective function merges adversarial loss, reconstruction loss, and age classification loss to effectively train the CAN-GAN model.

Sharma [19] introduces an integrated approach Leveraging techniques such as CycleGAN for age progression and Enhanced Super-Resolution GAN (ESRGAN) for image enhancement. The method aims to transform input face images into aged versions while preserving original features and improving image quality. By incorporating advancements like age-conditional GANs, progressive GANs, and cGANs, the model ensures better resolution and guidance in data generation. Although requiring substantial computational resources for training, This integration of face age progression and super-resolution techniques offers a promising solution to generating high-quality aged face images.

The progressive face aging framework proposed by Zhizhong [20] diverges from traditional GANs by concentrating specifically on capturing the aging dynamics of facial images. While conventional GANs typically generate images from random noise, this framework redefines the aging process through a progressive neural network structure composed of multiple sub-networks. These sub-networks specialize in learning the aging effects between neighboring age groups, facilitating controlled and lifelike age transitions. Through the integration of residual skip connections and binary gates, the framework effectively preserves facial identity and mitigates overfitting issues during the aging process. Moreover, by training the model end-to-end, it addresses cumulative error concerns,

TABLE I
Summary of Age Progression Algorithms and Methods

Paper Title	**ML Model**	**Dataset**	**Parameters**	**Results**	**Comments**
Personalized Age Progression with Aging Dictionary - 2015 [2]	PCA	CACD, MORPH, FGNET	Comparison to ground truth	Out of 12,300 visual comparisons done by 50 people 45.35% preferred this model, 36.45 preferred the prior works, while 18.20% found them comparable.	One of the earliest methodologies provided for age progression
Personalized Age Progression with Bi-Level Aging Dictionary Learning - 2018 [3]	PCA	CACD, MORPH	Comparison to ground truth	Out of 13,050 visual comparisons done by 50 people 36.5% preferred BDL-PAP, 34.8% preferred CDL-PAP, 26.7% preferred prior works and 2.0% were not satisfied.	BDL-PAP is more time efficient than CDL-PAP.
A framework for facial age progression and regression using exemplar face templates - 2021 [6]	CNN, GAN	FEI, MORPH II	CNN Face recognition	Similarity value of higher than 70% on the FEI dataset.	Ethnicity-based age progression method validated by CNN
Age Progression/Regression by Conditional Adversarial Autoencoder - 2017 [8]	CAAE	MORPH, CACD	Comparison to ground truth	Out of 1508 votes from 47 people 52.77% preferred CAAE, 28.99% preferred prior works and 18.24% thought they were equal.	Introduced Conditional Adversarial Autoencoder (CAAE) and integrated encoder in GAN
Face aging with conditional generative adversarial networks - 2017 [11]	CGAN	AGE-eGAN	Identity-Preserving Face Reconstruction and Aging	After Face reconstruction, The software 'OpenFace' gave these scores After Initial Reconstruction 53.2%, Pixelwise Optimization 59.8%, Identity-Preserving Optimization 82.9%.	GAN iteration for generating high-quality synthetic images
Face Aging with Identity-Preserved Conditional Generative Adversarial Networks - 2018 [12]	IPCGAN	CACD	Face Verification	After conducting face verification testing, CAAE scored 91.53%, acGAN scored 85.83%, IPCGAN scored 96.90%.	Integrated CGANs with an Identity-Preserved module and an age classifier
Attribute-Aware Face Aging With Wavelet-Based Generative Adversarial Networks - 2019 [15]	GAN	MORPH, CACD	Face Verification	On the MORPH dataset, the face verification score for CAAE was 11.77%, GLCA-GAN was 95.39%, PAG-GAN was 97.33%, Proposed model 99.42%.	Utilized wavelet analysis in the discriminator to stabilize translation
S2GAN: Share Aging Factors Across Ages and Share Aging Trends Among Individuals - 2019 [16]	S2GAN	MORPH, CACD	Aging Accuracy	On the MORPH dataset, Aging accuracy score for CAAE was 47.38%, IPCGAN was 64.42%, Proposed S2GAN was 93.0%.	Optimized with classification, reconstruction, and adversarial losses
Face Aging with Conditional Generative Adversarial Network Guided by Ranking-CNN - 2020 [17]	Ranking-CNN, GAN	CACD	Aging Accuracy	On the CACD dataset, Aging accuracy score for CAAE was 27.01%, IPCGAN was 47.98%, Proposed model was 52.08%.	Implemented Conditional GAN supervised by Ranking-CNN
CAN-GAN: Conditioned-attention normalized GAN for face age synthesis - 2020 [18]	CAN-GAN, CAAC	CACD, MORPH, FGNET	Face Verification	On the MORPH dataset, the face verification score for GLCA-GAN 95.39%, Yang et al. 97.00%, WaveletGLCA-GAN 99.80%, CAN-GAN 99.99%	Used CAN for age-based normalization and CAAC to improve age determination accuracy
An Improved Technique for Face Age Progression and Enhanced Super-Resolution with Generative Adversarial Networks - 2020 [19]	Cycle-GAN, ESRGAN	IMDB-WIKI, CACD, UTKFace, FGNET, CELEB-A	Age Estimation	On the IMDB-WIKI dataset, FACE++ software estimates the average age of the synthesized images as 30.2 for the age group 19 - 35. average age of 36.5 for age group 35 - 60 and average age of 61.7 for age group 60 and above	Combined CycleGAN for age progression and ESRGAN for image enhancement
PFA-GAN: Progressive Face Aging With Generative Adversarial Network - 2021 [20]	PFA-GAN	MORPH, CACD	Face Verification	On the MORPH dataset, FACE++ is used for Face Verification PFA-GAN 99.70%, CAAE 44.02%, IPCGAN 99.21%, WGLC-CAN 99.27%.	Introduced multiple sub-networks to capture aging dynamics, preserve facial identity, and mitigate overfitting
Face aging using global and pyramid generative adversarial networks - 2021 [23]	GAN	UTKFace, CACD	Age Classification	The age classification accuracy for GFA-GAN is 27.45%, for PFA-GAN it is 39.72%, for CAAE it is 17.64%, and for IPCGANs it is 17.88%.	Introduced weight sharing in GFA-GAN and pyramid weight sharing in PFA-GAN

ensuring robust age progression across various age groups and conditions, a distinct feature from the separate training paradigm of traditional GANs.

H. Tang [21] proposes Attention-Guided Generation Scheme I and Scheme II. In Scheme I, attention-guided generators G and F are employed to learn mappings between image domains X and Y. Here, attention masks are generated to selectively adjust foreground content while preserving background elements. However, Scheme I encounters challenges with complex tasks due to its reliance on a single attention and content mask generation process. To overcome these limitations, Scheme II introduces distinct sub-networks for generating attention and content masks, enabling more adaptable learning and translation of both foreground and background content. Moreover, attention-guided discriminators are proposed to emphasize discriminative content and enhance translation quality. Both schemes incorporate cycle-consistency loss [22] and supplementary regularization methods to effectively optimize the translation process.

The methodology proposed by Pantraki [23] treats age progression as an unsupervised task of translating images across different age groups, employing the UNIT network [24] for this purpose. Each age group is represented through three key components: an encoder, a decoder/generator, and a discriminator. In GFA-GAN, weight sharing is introduced using a combination of encoders and generators, capturing both localized and global features. PFA-GAN builds upon this approach by introducing a pyramid weight sharing mechanism to emulate the gradual aging process. Throughout the training process, the framework aims to minimize various losses such as cycle consistency and total variation, ensuring faithful translations between diverse age groups.

Table I presents a comprehensive summary of the algorithms employed in facial age progression research, along with the datasets utilized to train and evaluate these algorithms. Each algorithm is meticulously analyzed in terms of its year of publication, the specific dataset employed, and any noteworthy features or contributions.

III. Dataset

Collecting a comprehensive dataset relevant to facial aging is a challenging task, necessitating careful consideration of various factors. One of the critical criteria in the data collection process is ensuring diversity in age representation among subjects. This ensures that the dataset covers a broad spectrum of aging patterns, facilitating the development of robust facial aging models. However, while sequential images of the same individual at different ages are often preferred, certain methodologies in facial aging research can effectively discern aging patterns without this requirement.

The success of Generative Adversarial Networks (GANs) in producing realistic facial aging progression hinges on the diversity and quality of the training dataset, crucial for capturing the nuanced complexities of aging with accuracy. The various datasets commonly used in research related to facial age progression and related fields vary in terms of size, subject type, labeling criteria, and distribution of images across age groups. Several datasets focus on celebrity faces, such as CACD [26], IMDB-WIKI [27], and CelebA [28], with large numbers of images spanning different age categories. Other datasets, like FFHQ [30], UTKFace [8], VGG Face 2 [34], Morph [29], AgeDB [35] are collected from general populations, encompassing a range of ages and genders. The labeling criteria predominantly include age and gender, although some datasets also include additional attributes such as race. While most datasets exhibit non-monotone distributions across age groups, indicating a varied representation of ages, some datasets, like FGNET [31] and Olivetti [33], feature monotone distributions primarily focused on specific age ranges. Additionally, there are datasets like WebFace260M [32], WebFace42M [32], DigiFace1M [36], MegaFace1M [37] that provide extensive collections of facial images without specific labels, catering to various research needs.

Ethical concerns surrounding facial aging datasets, especially those containing images of children, are paramount. While these datasets offer valuable insights, they raise significant privacy and consent issues. Mitigating these concerns is essential to ensure responsible research practices, often prompting the exploration of alternative methods like generating synthetic images. One notable paper addressing this challenge is ChildGAN [25]. It introduces a novel method for generating synthetic images of children that closely resemble real faces while avoiding the need to use actual images of minors.

The methodology proposed in the aforementioned paper comprises three main phases: firstly, gathering synthetic data for initializing training datasets; secondly, separately training ChildGAN models tailored for boys and girls using StyleGAN2; and finally, employing techniques for editing the latent space to modify facial attributes. Synthetic data from diverse origins undergoes a rigorous filtering process and is categorized into two distinct classes. ChildGAN, constructed upon StyleGAN2, undergoes training via transfer learning to generate high-fidelity facial images of children. The manipulation of facial attributes, such as expressions and lighting conditions, is facilitated through latent space editing. This approach seamlessly combines sophisticated deep learning techniques with meticulous data curation to create lifelike synthetic facial images of children.

IV. Conclusion

In conclusion, facial age progression research represents a dynamic and evolving field at the intersection of computer vision, machine learning, and ethics. The methodologies discussed, ranging from traditional feature-based approaches to advanced deep learning techniques utilizing Generative Adversarial Networks (GANs), underscore the complexity and diversity of strategies employed to predict facial aging accurately. While each approach offers distinct advantages and contributions, the success of these methods critically hinges on the quality and diversity of datasets utilized for training and evaluation. Ethical considerations surrounding data collection, particularly concerning the inclusion of images of children, necessitate careful navigation to ensure privacy protection and mitigate potential harms. Moreover, the comprehensive overview of publicly available datasets provided herein underscores the foundational importance of diverse and well-annotated data repositories in advancing facial age progression research.

Moving forward, continued interdisciplinary collaboration and ethical awareness will be paramount in harnessing the full potential of facial age progression technologies for societal benefit while addressing associated ethical concerns.

V. Challenges and future work

Challenges and future directions in the domain of facial age progression (FAP) encompass several critical aspects that researchers are actively addressing. One significant challenge lies in achieving a balance between visual fidelity, aging accuracy, and identity preservation in synthesized images. While advancements have been made in improving these aspects individually, achieving optimal results across all dimensions simultaneously remains a complex task. Future research efforts can be focused on developing novel algorithms and techniques that strike a harmonious balance between these competing objectives.

Another challenge in facial age progression is the availability and quality of training data. Current datasets often suffer from biases, such as underrepresentation of certain age groups or ethnicities, which can impact the performance and generalization capabilities of facial age progression models. Addressing these biases and curating more diverse and comprehensive datasets will be crucial for improving the robustness and reliability of facial age progression systems.

Ethical considerations also pose significant challenges in the development and deployment of facial age progression technology. Concerns related to privacy, consent, and potential misuse of synthesized images underscore the need for robust ethical frameworks and guidelines. Future research will need to explore ethical implications in greater depth and develop mechanisms to ensure responsible use of facial age progression technology.

Furthermore, scalability and computational efficiency are ongoing challenges in facial age progression, particularly as the demand for real-time or large-scale age progression applications grows. Optimizing algorithms for faster inference and reducing computational resource requirements will be essential for practical deployment in various domains, including law enforcement, entertainment, and healthcare.

Looking ahead, future work in facial age progression will likely focus on exploring new avenues such as generative models with improved interpretability, incorporating domain knowledge from related fields like psychology and gerontology, and leveraging emerging technologies such as augmented reality (AR) and virtual reality (VR) for more immersive and interactive age progression experiences. By addressing these challenges and pursuing innovative research directions, the field of facial age progression holds tremendous potential for transformative advancements with far-reaching societal impacts.

References

[1] I. Kemelmacher-Shlizerman, S. Suwajanakorn and S. M. Seitz, "Illumination-Aware Age Progression," 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 2014, pp. 3334-3341, doi: 10.1109/CVPR.2014.426.

[2] X. Shu, J. Tang, H. Lai, L. Liu and S. Yan, "Personalized Age Progression with Aging Dictionary," 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 2015, pp. 3970-3978, doi: 10.1109/ICCV.2015.452.

[3] X. Shu, J. Tang, Z. Li, H. Lai, L. Zhang and S. Yan, "Personalized Age Progression with Bi-Level Aging Dictionary Learning," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 4, pp. 905-917, 1 April 2018, doi: 10.1109/TPAMI.2017.2705122.

[4] Shintaro Yamamoto, Pavel A. Savkin, Takuya Kato, Shoichi Furukawa, and Shigeo Morishima. 2017. Facial video age progression considering expression change. In Proceedings of the Computer Graphics International Conference (CGI '17). Association for Computing Machinery, New York, NY, USA, Article 5, 1–5. https://doi.org/10.1145/3095140.3095145

[5] Riaz, Sidra & Park, Unsang & Choi, Jongmoo & Natarajan, Prem. (2019). Age progression by gender-specific 3D aging model. Machine Vision and Applications. 30. 10.1007/s00138-018-0975-2.

[6] Elmahmudi, A., Ugail, H. A framework for facial age progression and regression using exemplar face templates. Vis Comput 37, 2023–2038 (2021). https://doi.org/10.1007/s00371-020-01960-z

[7] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2020. Generative adversarial networks. Commun. ACM 63, 11 (November 2020), 139–144. https://doi.org/10.1145/3422622

[8] Z. Zhang, Y. Song and H. Qi, "Age Progression/Regression by Conditional Adversarial Autoencoder," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 4352-4360, doi: 10.1109/CVPR.2017.463.

[9] Kingma, Diederik P. and Max Welling. “Auto-Encoding Variational Bayes.” CoRR abs/1312.6114 (2013): n. pag.

[10] Makhzani, A., Shlens, J., Jaitly, N., Goodfellow, I. and Frey, B., 2015. Adversarial autoencoders. arXiv preprint arXiv:1511.05644.

[11] G. Antipov, M. Baccouche and J. -L. Dugelay, "Face aging with conditional generative adversarial networks," 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 2017, pp. 2089-2093, doi: 10.1109/ICIP.2017.8296650.

[12] X. Tang, Z. Wang, W. Luo and S. Gao, "Face Aging with Identity-Preserved Conditional Generative Adversarial Networks," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018, pp. 7939-7947, doi: 10.1109/CVPR.2018.00828.

[13] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2017. ImageNet classification with deep convolutional neural networks. Commun. ACM 60, 6 (June 2017), 84–90. https://doi.org/10.1145/3065386

[14] H. Yang, D. Huang, Y. Wang and A. K. Jain, "Learning Face Age Progression: A Pyramid Architecture of GANs," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, (2018), pp. 31-39, doi: 10.1109/CVPR.2018.00011.

[15] Y. Liu, Q. Li and Z. Sun, "Attribute-Aware Face Aging With Wavelet-Based Generative Adversarial Networks," in 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, (2019) pp. 11869-11878. doi: 10.1109/CVPR.2019.01215

[16] Z. He, M. Kan, S. Shan and X. Chen, "S2GAN: Share Aging Factors Across Ages and Share Aging Trends Among Individuals," 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South), 2019, pp. 9439-9448, doi: 10.1109/ICCV.2019.00953.

[17] M. Sheng, Z. Ma, H. Jia, Q. Mao and M. Dong, "Face Aging with Conditional Generative Adversarial Network Guided by Ranking-CNN," 2020 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), Shenzhen, China, 2020, pp. 314-319, doi: 10.1109/MIPR49039.2020.00071.

[18] Shi, Chenglong & Zhang, Jiachao & Yao, Yazhou & Sun, Yunlian & Rao, Huaming & Shu, Xiangbo. (2020). CAN-GAN: Conditioned-attention normalized GAN for face age synthesis. Pattern Recognition Letters. 138. 10.1016/j.patrec.2020.08.021.

[19] Sharma, N., Sharma, R. & Jindal, N. An Improved Technique for Face Age Progression and Enhanced Super-Resolution with Generative Adversarial Networks. Wireless Pers Commun 114, 2215–2233 (2020). https://doi.org/10.1007/s11277-020-07473-1

[20] Zhizhong Huang, Shouzhen Chen, Junping Zhang, and Hongming Shan. 2021. PFA-GAN: Progressive Face Aging With Generative

Adversarial Network. Trans. Info. For. Sec. 16 (2021), 2031–2045. https://doi.org/10.1109/TIFS.2020.3047753

[21] H. Tang, H. Liu, D. Xu, P. H. S. Torr and N. Sebe, "AttentionGAN: Unpaired Image-to-Image Translation Using Attention-Guided Generative Adversarial Networks," in IEEE Transactions on Neural Networks and Learning Systems, vol. 34, no. 4, pp. 1972-1987, April 2023, doi: 10.1109/TNNLS.2021.3105725.

[22] J. -Y. Zhu, T. Park, P. Isola and A. A. Efros, "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks," 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017, pp. 2242-2251, doi: 10.1109/ICCV.2017.244.

[23] Pantraki, E., Kotropoulos, C. Face aging using global and pyramid generative adversarial networks. Machine Vision and Applications 32, 82 (2021). https://doi.org/10.1007/s00138-021-01207-4

[24] Liu, M.Y., Breuel, T. and Kautz, J., 2017. Unsupervised image-to-image translation networks. Advances in neural information processing systems, 30.

[25] Farooq, Muhammad Ali, Wang Yao, Gabriel Nicolae Costache and Peter Corcoran. “ChildGAN: Large Scale Synthetic Child Facial Data Using Domain Adaptation in StyleGAN.” IEEE Access 11 (2023): 108775-108791.

[26] Chen, BC., Chen, CS., Hsu, W.H. (2014). Cross-Age Reference Coding for Age-Invariant Face Recognition and Retrieval. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds) Computer Vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science, vol 8694. Springer, Cham. https://doi.org/10.1007/978-3-319-10599-4_49

[27] Rothe, R., Timofte, R. & Van Gool, L. Deep Expectation of Real and Apparent Age from a Single Image Without Facial Landmarks. Int J Comput Vis 126, 144–157 (2018). https://doi.org/10.1007/s11263-016-0940-3

[28] Z. Liu, P. Luo, X. Wang and X. Tang, "Deep Learning Face Attributes in the Wild," 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 2015, pp. 3730-3738, doi: 10.1109/ICCV.2015.425.

[29] K. Ricanek and T. Tesafaye, "MORPH: a longitudinal image database of normal adult age-progression," 7th International Conference on Automatic Face and Gesture Recognition (FGR06), Southampton, UK, 2006, pp. 341-345, doi: 10.1109/FGR.2006.78.

[30] Or-El, R., Sengupta, S., Fried, O., Shechtman, E., Kemelmacher-Shlizerman, I. (2020). Lifespan Age Transformation Synthesis. In: Vedaldi, A., Bischof, H., Brox, T., Frahm, JM. (eds) Computer Vision – ECCV 2020. ECCV 2020. Lecture Notes in Computer Science(), vol 12351. Springer, Cham. https://doi.org/10.1007/978-3-030-58539-6_44

[31] A. Lanitis, C. J. Taylor and T. F. Cootes, "Toward automatic simulation of aging effects on face images," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 4, pp. 442-455, April 2002, doi: 10.1109/34.993553.

[32] Zhu, Zheng, Guan Huang, Jiankang Deng, Yun Ye, Junjie Huang, Xinze Chen, Jiagang Zhu, Tian Yang, Jiwen Lu, Dalong Du and Jie Zhou. “WebFace260M: A Benchmark Unveiling the Power of Million-Scale Deep Face Recognition.” 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2021): 10487-10497.

[33] Mirzal, Andri. (2014). NMF versus ICA for blind source separation. Advances in Data Analysis and Classification. 10.1007/s11634-014-0192-4.

[34] Q. Cao, L. Shen, W. Xie, O. Parkhi and A. Zisserman, "VGGFace2: A Dataset for Recognising Faces across Pose and Age," in 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), Xi'an, China, 2018 pp. 67-74. doi: 10.1109/FG.2018.00020

[35] S. Moschoglou, A. Papaioannou, C. Sagonas, J. Deng, I. Kotsia and S. Zafeiriou, "AgeDB: The First Manually Collected, In-the-Wild Age Database," 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 2017, pp. 1997-2005, doi: 10.1109/CVPRW.2017.250.

[36] Bae, Gwangbin, M. D. L. Gorce, Tadas Baltrušaitis, Charlie Hewitt, Dong Chen, Julien P. C. Valentin, Roberto Cipolla and JingJing Shen. “DigiFace-1M: 1 Million Digital Face Images for Face Recognition.” 2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) (2022): 3515-3524.

[37] Kemelmacher-Shlizerman, Ira, Steven M. Seitz, Daniel Miller and Evan Brossard. “The MegaFace Benchmark: 1 Million Faces for Recognition at Scale.” 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2015): 4873-4882.