Check for updates

Comparative analysis of CycleGAN and AttentionGAN on face aging application

NEHA SHARMA[1,*], REECHA SHARMA[1] and NEERU JINDAL[2]

[1]Department of Electronics and Communication Engineering, Punjabi University, Patiala, Punjab, India
[2]Department of Electronics and Communication Engineering, Thapar Institute of Engineering and Technology, Patiala, Punjab, India
e-mail: nehaleo_sharma@yahoo.com; reecha@pbi.ac.in; neeru.jindal@thapar.edu

Abstract. Recently, there is incredible progress in the arena of machine learning with generative adversarial network (GAN) methods. These methods tend to synthesize new data from input images that are highly realistic at the output. One of its applications in the image-to-image transformation way is the face aging task. In the face aging process, new face images are synthesized with the help of the input images and desired target images. Face aging can be beneficial in several domains such as in biometric systems for face recognition with age progression, in forensics for helping to find the missing children, in entertainment, and many more. Nowadays, several GANs are available for face aging applications and this paper focuses on the insight comparison among the frequently used image-to-image translation GANs which are CycleGAN (Cycle-Consistent Adversarial Network) and AttentionGAN (Attention-Guided Generative Adversarial Network). The first model (CycleGAN) comprises two generators, two discriminators, and converting an image from one domain to another without the need for paired images dataset. The second is AttentionGAN, which consists of attention masks and content masks multiplied with the generated output in one domain to generate a highly realistic image in another domain. For comparison, these two are trained on two dataset which is CelebA-HQ (CelebFaces Attributes high-quality dataset) and FFHQ (Flickr Faces HQ). Efficacy is evaluated quantitatively with identity preservation, five image quality assessment metrics, and qualitatively with a perceptual study on synthesized images, face aging signs, and robustness. It has been concluded that overall CycleGAN has better performance than AttentionGAN. In the future, a more critical comparison can be performed on the number of GANs for face aging applications.

Keywords. Generative adversarial network (GAN); CycleGAN; attentionGAN; image to image transformation; face age progression.

1. Introduction

Since widespread utilization has been shown for deep learning (DL), excessive development has been achieved in face recognition and face detection tasks. Further, various studies related to face manipulation such as face synthesis, face swapping, facial attributes, and their analysis has been extensively studied [1–3]. Face recognition is already in place but still needs more attention. Moreover, the COVID-19 crisis has forced many more daily tasks to be digitalized in the world, even after the post-COVID-19 some tasks may remain digitalized. Thus, it needs precise face recognition and verification for security at many places. For example, at airports for baggage drops, security screening, gate check-ins, this may consume large processing time. Thus, face ID can provide comfort at these overcrowded points. Similarly,

in banks face ID feature is offered to customers for their respective log-in, also in ATM's facial features are used for the withdrawal of the money. Face recognition with age progression can be helpful in the above-mentioned applications. Thus, the application of the automatic face aging method with machine learning to handle the large database can come to play the role of auto-readjustment of e-records. And in many more cases like avoiding contact at attendance system for offices and degree colleges, passports renewal, electronic customer-retailer business, abduction of children where after many years person's biological face appearance changes. So, it becomes necessary to go deeper into the existing face aging methods and their current scenario, so that in the future various face aging method challenges can be solved. That is why face aging is the topic of utmost interest. Based on these considerations, two face aging methods CycleGAN and AttentionGAN are compared for face aging application and evaluated. As they have

*For correspondence

achieved tremendous attention in translation for the image-to-image GANs, these GANs have been generating remarkable results. To the best of our knowledge, no comparison is done between these two GANs for face aging tasks using a different dataset. In this paper, extensive experiments are conducted which give the comparison analysis between CycleGAN and AttentionGAN to measure the ability to produce plausible photorealistic face images. Some of the applications of face aging or face age progression and its related research fields are shown in figure 1.

The main objectives of this paper are:

1. The comparison between CycleGAN and AttentionGAN models on the face aging application and list their merits and demerits.
2. To evaluate the performance quantitatively and qualitatively, two dataset CelebA-HQ and FFHQ are used.
3. The potential of CycleGAN and AttentionGAN is also shown with their robustness and various perspectives for

   future directions are enumerated.

1.1 *Face aging*

It is a procedure of altering a face image across its different ages which are termed as target age, with natural aging effects or reviving effects on the given face image. The block diagram of the face age progression or face age synthesis is shown in figure 2.

1.1a *Input dataset/target images*: The input dataset and target images are pre-processed so that fed input images are of good quality. At the input, it is also ensured that images are of a face only. Also, the face is cropped which enhances the focus on the face. The number of images used for the input dataset and target dataset is almost equal. These are further trained to learn the pattern for a transition from given input images to the desired target images.

1.1b *Face age synthesis process*: This process includes deep learning from the given database. It extracts the features from the face and memories for the new transition. The larger the database better it will be because the more data gives a lot of variation, repetition to learn and recognize the features which further helps to do the prediction with accuracy. During the training, conversion takes place for every image and testing images give the results from that learned training. Thus, testing images are used to predict how effectively the system is trained on the given dataset. It also depends on the quality and amount of dataset provided to the system. Then the testing results give the transition required for the input image to the target image.

1.1c *Synthesized images*: All the training images may not be able to learn the transition so in the testing phase not all the output results are as desired. The amount of training data, quality of the dataset, pre-processing steps, and algorithm all decide the accuracy of the final output.

Every human aging process is unique. For illustration, the age-progressed faces of Albert Einstein are shown in figure 3. Further, there are four conceptual terms with age. Actual age is also termed the real age of the person. Secondly, the visual appearance of the person is termed as appearance age. Thirdly, perceived age is measured by human subjects from the visual appearance and fourth, age predicted by a machine based on its visual appearance is termed as estimated age.
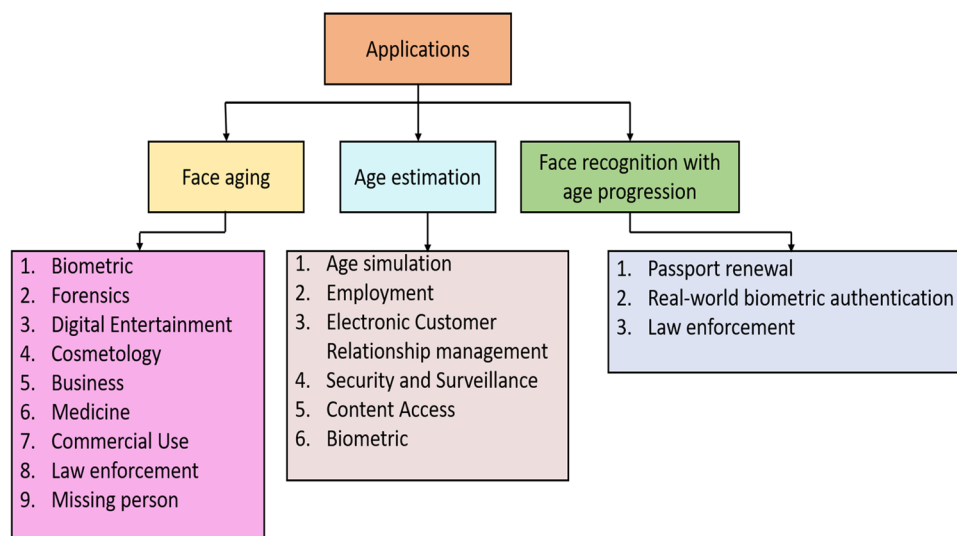

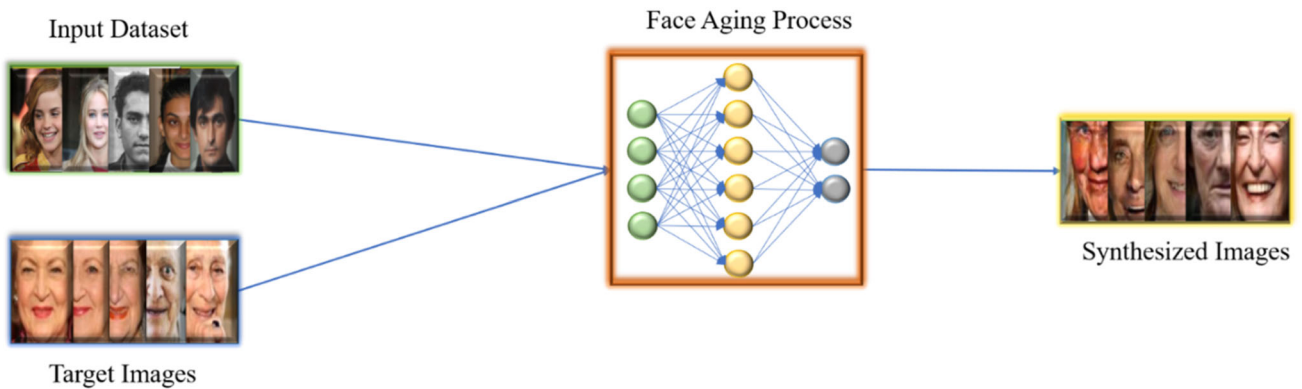**Figure 1.** Various applications of face aging and its related research fields.
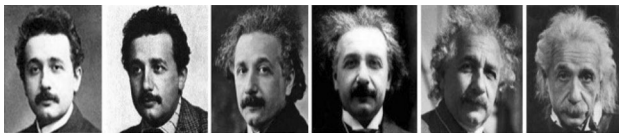
Figure 2. Face age synthesis [4].

Figure 3. Face Aging of Albert Einstein [4].

However, there are many reasons which affect the remarkable changes in human face aging, roughly it is divided into two portions. First is the primary growth i.e., from birth to growth till childhood. This stage shows the changes mainly in a facial curve, facial features (eye, chin, mouth, etc.), and facial features distributions. This is called shape change or craniofacial growth but a slight consequence is on change in color of skin at this stage. Further, the second stage is adult aging i.e., growth from adulthood till old age. In this stage, prominent change is shown in skin texture and color of the skin such as wrinkles appear, facial lines are visible, and minor craniofacial changes, also there is a reduction in muscle strength and elasticity [5–7]. Thus, in face age progression most of the variations are associated with the face texture.

Some signs of facial skin aging are shown in figure 4 with texture changes such as forehead wrinkles, crow's feet wrinkles, or lateral Canthal lines, Glabellar frown lines means a vertical line in between the brows, under-eye bags or full bags, wrinkles under the eye, Nasolabial folds, Paler or yellow skin, Marionette lines that initiate from corner of both side of the mouth to the corner of the chin, wrinkles in the upper lip side. Also, some geometrical changes like brow drop, vertical lower-eyelid length increases, nose elongation, and tip movement, lower face Ptosis, and chin sagging, the small focal accumulation is termed as jowl of fact in the lower cheek overlying the jaw bone [7].

Geometrical transformations include two parts in context to facial feature points. First is the size of components and their distribution on the face such as distance between different facial parts which contributes to face aging. For example, the distance between brows and eyes and distance between nose and lips, etc. The second is face contour or face shape. The face shape change in age progression is mainly jowl variations that occur because of the face skin losses and a noticeable reduction in muscle.

In texture variations, changes in the skin directly or through changes in muscles, fat with age progression. Also, a facial skeleton that can alter the face geometry shows its effect on the texture of the skin. Thus, texture changes on the human face play a vital role in face age progression.

Further, the face pattern is unique to each person and linked with several internal factors like hormonal changes, stress, ethnicity, etc. External factors such as environmental conditions, geographical conditions, lifestyle, etc. affect the face aging results. Some internal and external factors are discussed below.

1.1d *Internal factors*: As ethnicity is described as one particular population in terms of genetic similarities. Face alterations in different ethnic can describe genetic differences which show with face age progression also. Some face age-related factors such as the color of skin, skin thickness and natural moisturizer [8, 9], etc. can be different from one ethic to another. Also, human skin is observed with many gender differences [10]. On average female skin is thinner than male skin. Generally, males have deeper wrinkles than females. The male skin creases are produced at a later stage of age but they are more noticeable when they appear. Moreover, human skin is a maker of hormones [10, 11]. These hormones help in the growth and biological functionality of skin muscles.

1.1e *External factors*: The changes in the human's face are not only because of internal factors. But, external factors play a significant role in face aging methods. Such as geographical area, gravity, pollution, temperature and working environment, etc. However, lifestyle plays a great role which includes habits such as nutrition, sleep, drugs, exercise, exposure to UV rays, and many more.

Figure 4. Facial aging marks on the complete face [7].

Further, face aging categories can be divided in various ways. The input image can be 2D or 3D face images. 3D image representation can provide both the shape and texture information and potentially can produce better results in comparison to 2D images. The changes in the face are observed in terms of texture changes and geometrical or shape changes as a person grows. The methods used for face age progression are prototype-based methods and physical model-based methods that are conventional face aging methods. But, with the advancement in research, deep learning-based approaches can deal with a larger and complex dataset. The deep learning approaches have shown notable results in the face aging field. Figure 5 shows the face age progression with different categories.

2. Related work

The relevant previous work done on GAN for the translations from image to image is provided in this section.

2.1 *GAN*

The substantial improvement in deep generative models is the GAN framework presented by Goodfellow *et al* [12]. GAN comprises of two neural nets, the generator and a discriminator that works as a two-player, non-cooperative game competing against each other. The performance of one network comes at the cost of another network. The generator generates fake images and the discriminator distinguishes between actual and fake images. The training is done at the same time for both models. The response from the discriminator aids the generator to progress its performance, the generator cannot access the real data while the discriminator can access the real as well as fake data [13]. The generator and discriminator networks consist of convolutional layers such as Deep Convolutional GAN (DCGAN) or fully connected layers [14]. In a fully-connected GAN, the generator, as well as the discriminator uses a fully-connected network and in a convolutional network GAN, Convolutional Neural Network (CNN) is used in the generator and the discriminator. Although, training both models with CNN is difficult in comparison to
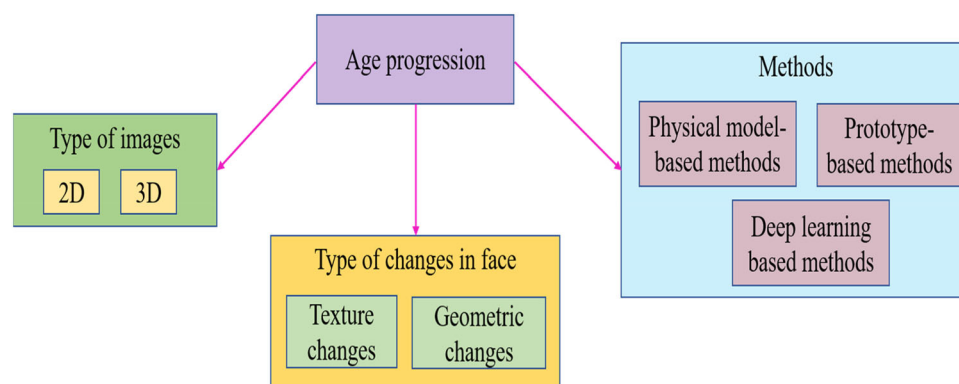

Figure 5.. Face age progression different categories.

fully connected network GAN. The two neural network models play the min-max game which is mathematically expressed in equation (1) as:

$$\min_{G} \max_{D} Y(D, G) = E_{p \sim s_{data}}(p)[\log D(p)] + E_{r \sim s_r(r)}[\log(1 - D(G(r)))], \quad (1)$$

where G, F are generators, D is discriminator, $E_{p \sim s_{data}}(p)[\log D(p)]$ is log probability of D to predict the real-world data is genuine and $E_{r \sim s_r(r)}[\log(1-D(G(r))]$ is the log probability that G's generated data is not genuine. r is random noise as input, p is real data instance, s is a probability distribution.

Alqahtani *et al* 2019 [15] had shown the various GANs with their specific application and provided much border explanation on GAN in various aspects. The paper also addressed the future challenges for GAN in terms of their training and GAN standard evaluation metrics. Goodfellow *et al* [16] in 2020 had shown a wider explanation on GAN. GAN is a type of artificial intelligence algorithm and it is GAN whose generative models can produce highly realistic images. But, still has a lot of challenges to overcome as it is based on game theory, such as it is still difficult to train GANs. With the problem of training GAN and its evaluation, wide space is present in the research community to explore GAN. Guo *et al* 2020 [17] had proposed a GAN called combined GAN (Com-GAN) and done the study using improved GAN for generating fundus images. It had shown high-quality results in comparison to other generative models. Thus, further research can be done with com-GAN for image generation and image translation tasks in various fields.

The previous supervised learning process with deep neural networks for face aging [18–20] needed paired dataset for training, which was quite difficult. To overcome this problem, GANs have been used for face aging to train with the unpaired dataset and its various variants were also used. Conditional GAN (cGANs) [21] had been used for face aging applications with unpaired dataset [22–30], to attain better realistic results than conventional methods like prototype methods [31] and physical model-based methods [32].

2.2 *CycleGAN*

The CycleGAN framework was introduced by Zhu *et al* in 2017 [33] for the image-to-image transformation task without the need for a paired training database. The learning of mapping from the domain P (input) to domain Q (output) and vice versa is achieved with the help of cycle consistency losses. The hypothesis in the paper was that some underlying associations that exist among the two domains, combined the two losses "cyclic losses" with "adversarial losses" on the input and output domain. In figure 6 the framework has two plotting roles G: P to Q and F: Q to P also D_Q and D_P two discriminators. The discriminator, D_Q insists the generator G

which decodes P into new synthesized output images. Similarly, the function of D_P and F is in the opposite direction. Although the cycle-consistency loss in figure 6 is also termed as cyclic loss, the cyclic loss helps to get back the real input image in the same domain from the generated output image in another domain. So, two cyclic losses, one in the forward direction is stated as $(p \rightarrow G(p) \rightarrow F(G(p)) \approx p)$ and another is in the backward direction and expressed as $(q \rightarrow F(q) \rightarrow G(F(q)) \approx q)$. The main objective consists of two terms, one is adversarial losses in that the distribution of synthesized images is matched with the data distribution of images in the target domain. The second is the cycle consistency loss that avoids the learned mappings of G and F to contradict each other. Mathematically, the adversarial loss is shown in equation (2) and cycle consistency loss in equation (3).

For mapping function in G: P $\rightarrow$ Q and D_Q is discriminator. The adversarial loss is expressed as:

$$L_{GAN}(G, D_Q, P, Q) = E_{q \sim s_{data}(q)}[\log D_Q(q)] + E_{p \sim s_{data}(p)}[\log (1 - D_Q(G(p)))], \quad (2)$$

where G tries to generate images G(p) and D_Q. tries to differentiate between synthesized images G(p) and images q (in Q domain). G goals to diminish this objective in comparison to an adversary D which attempts to maximize it, i.e., $min_G max_{D_Q} L_{GAN}(G, D_Q, P, Q)$

Cycle-consistency loss:

$$L_{cyc}(G, F) = E_{p \sim s_{data}(p)}[||F(G(p)) - p||_1] + E_{q \sim s_{data}(q)}[||G(F(q)) - q||_1]. \quad (3)$$

With this, the reconstructed image F(G(p)) closely resemble the input image p.

T objective function for optimization by combining the adversarial loss and cycle-consistency loss is presented in Eqs. (4), (5), and (6) as:

$$L(G, F, D_P, D_Q) = L_{GAN}(G, D_Q, P, Q) + L_{GAN}(F, D_P, Q, P) + \mathcal{M}L_{cyc}(G, F), \quad (4)$$

$$G^* argmin_{G,F} max_{D_P D_Q} L(G, F, D_P, D_Q), \quad (5)$$

$$Loss_{complete} = Loss_{adv} + \mathcal{M}Loss_{cyc}, \quad (6)$$

where $Loss_{adv}$ is adversarial loss, $Loss_{cyc}$ is cyclic loss and $\mathcal{M}$ manage the relative significance of the two objectives.

The CycleGAN model stretches its edges to show the maximum possible setting in unsupervised learning. The generalization of the model is shown with different broader range applications without paired data and had shown the outperforming results. The CycleGAN is best to do changes in texture and color but geometric changes show little progress [33].
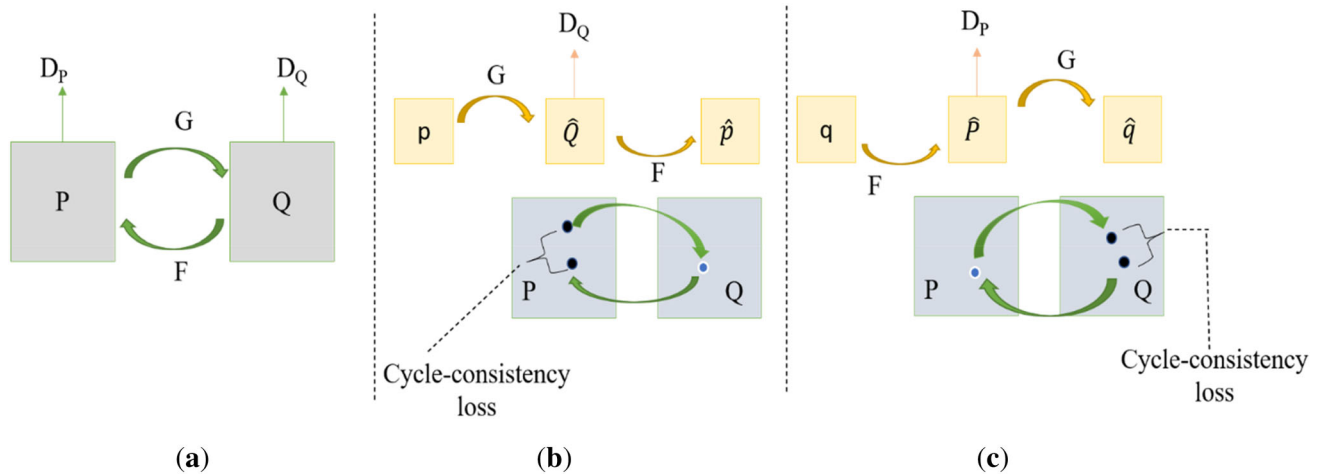
Figure 6. CycleGAN model overview [33].

However, Welander *et al* 2018 [34] presented a comparison among CycleGAN and UNIT using multi-contrast MR images which had shown that CycleGAN produced better results for visually realistic images. Nanavati *et al* in 2020 [35] presented a GANs comparative analysis which included SinGAN, cGAN, CycleGAN, StarGAN and had shown in the mathematical results that CycleGAN had performed well in comparison to various other GANs using several metrics like root mean square error (RMSE), Universal quality measure (UQI), multi-scale structural similarity (MS-SSIM), visual information fidelity (VIF). Burad *et al* 2020 [36] showed a comparative study of CycleGAN and progressive growing GAN and presented that the progressive growing GAN can be good in the case of medical images because they preserve more details and produce a high-resolution image than CycleGAN.

2.3 *AttentionGAN*

AttentionGAN was presented by Tang *et al* [37]. Although various existed methods for the image-to-image translation had produced remarkable results, still it had some visual artifacts in the output images because of the week translation of high-level semantic input images. AttentionGAN identifies the foreground objects and minimizes the changes in the background. As shown in figure 7 (scheme B), the AttentionGAN produces the attention masks also content masks which combine along with the generated output to produce the high-quality target images. The higher intensity level of the attention mask for any particular image indicates the high contribution for better change. The AttentionGAN generalization is shown with the help of various kinds of different application images. The results obtained are sharper and more realistic [37].

AttentionGAN had proposed two schemes: scheme A and scheme B. In scheme A, the generator focuses on the

specific sections of the image which are responsible for producing good expression at output such as eyes, mouth, and other parts remain untouched like hairs, glasses, clothes. Thus, scheme A can change the output image better where there is a greater overlapping similarity between the two domains such as facial expression to another facial expression translation job. To overcome the disadvantages of scheme A, scheme B was proposed. It has two generators G, F composes of two sub-nets for producing several intermediate attention masks and content masks that help to eliminate the drawback of scheme A. Thus, the accomplishment of scheme B with the generation of both foreground and background attention masks, lets the model alter the foreground and instantaneously preserve the background of a given face image. Mathematically, content and attention masks get multiplied with the generated image from the generator to produce a realistic image as shown in equation (7) as:

$$G(p) = \sum_{f=1}^{n-1}(C_q^f * A_q^f) + p * A_q^b, \qquad (7)$$

where $(\{A_q^f\}_{f=1}^{n-1}, A_q^b)$, k attention masks are produced, $G(p)$ generated target face aged image, P and Q are two domains, p is the input image, C_q^f signifies a content mask and q are images in another domain. Also, to produce a reconstructed image from G(p), another generator F, with a similar has the structure to G is used. With the help of content masks, attention masks, and G(p), recreate the original image p and expressed mathematically in equation (8) as:

$$F(G(p)) = \sum_{f-1}^{n-1}(C_p^f * A_p^f) + G(p) * A_p^b, \qquad (8)$$

where $F(G(p))$ is the recreated image that is close to the real image p. Further, C_p^f is a content mask, G(p) is a
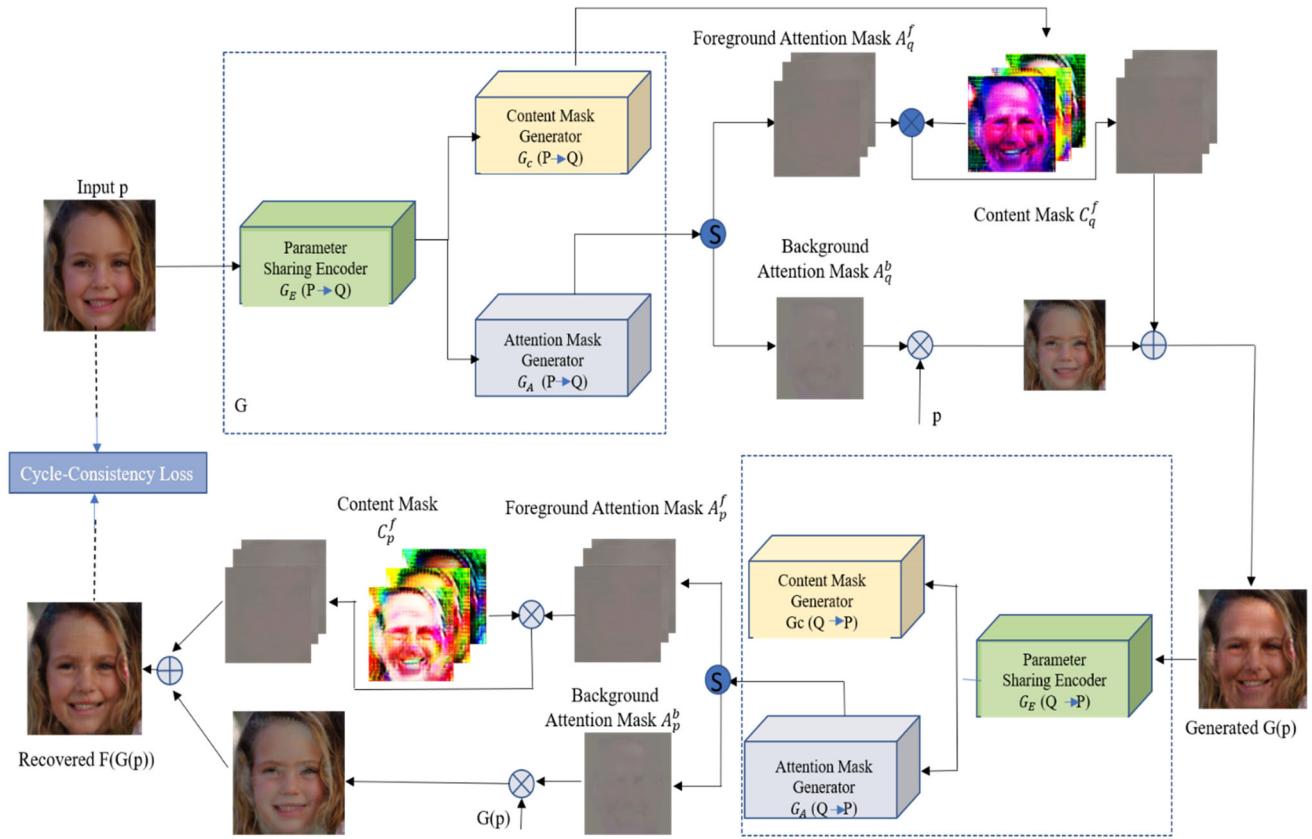
Figure 7. AttentionGAN framework [37] applied on face aging application.

synthesized image, A_p^b, A_p^f are background attention mask and foreground attention mask respectively. Thus, the aim of the image-to-image conversion is achieved.

This paper provides the comparison using AttentionGAN (scheme B) on face aging application for image-to-image translation GAN. Mathematically objective of optimization in AttentionGAN scheme B is expressed in equation (9).

$$L = L_{GAN} + \epsilon_{cycle} * L_{cycle} + \epsilon_{id} * L_{id}, \qquad (9)$$

where, L_{GAN} represents GAN loss, L_{cycle} is termed as cyclic loss and L_{id} represents identity preserving loss. Also, ϵ_{cycle} and ϵ_{id} are the parameters to manage each term relation.

Also, the min-max game of AttentionGAN works as expressed in equation (10) as:

$$L_{AGAN}(G, D_{QA}) = E_{q \sim s_{data}}(q)[\log D_{QA}([A_q, q])$$
$$+ E_{p \sim s_{data}}(p)[\log(1 - D_{QA}([A_q, G(p)]))], \qquad (10)$$

where D_{QA} is attention-guided discriminator, G is a generator, $[A_q, G(p)]$ is fake image pairs and $[A_q, q]$ is real image pairs. In this equation, D_{QA} tries to discriminate between the generated image pair $[A_q, G(p)]$ and the real image pairs $[A_q, q]$. Similarly, from the other domain $L_{AGAN}(F, D_{PA})$ where F is a generator and D_{PA} is a discriminator which attempts to

differentiate among fake image pairs $[A_p, F(q)]$ and actual image pairs $[A_p, p]$. Thus, the discriminator focuses on the main content and overlooks the unrelated content. This is used for scheme A only, in the case of scheme B, the generator is itself very effective to learn main content from the source and target domain images.

2.4 Face aging GAN variants

Image synthesis turns out to be an interesting area in the computer vision field. Human faces have been thoroughly studied in multimedia fields, computer graphics, and computer vision [4]. In the last few years, a growing amount of research on face age progression and its associated applications have been described such as face age estimation, cross-age face study, and entertaining field, etc. Some current and existing impressive face age progression approaches using GANs have been discussed below in table 1.

3. The proposed evaluation method

The comparison assessment among the two image-to-image transformation methods provides a result that which model can produce more realistic output images, the robustness of

models i.e., is how the model behaves with different kinds of input images and which model can preserve the identity better. Therefore, qualitative and quantitative measures are performed for evaluating CycleGAN and AttentionGAN frameworks.

3.1 *Dataset*

- CelebA-HQ: The dataset was introduced in the paper progressive growing GANs for improved quality by Karras *et al* [51]. It contains 30,000 high-quality face images out of CelebA dataset. The dataset is largely diverse with images that have larger variations in pose, annotations, and background clutter. A total of 1000 images are trained from the CelebA-HQ dataset. For the face aging task, two groups of images are taken, one younger age (age up to 18 years) and another older age (age above 55 years).
- FFHQ: Flicker-Faces-HQ dataset contains high-quality human faces with 70,000 PNG format images. The dataset contains large variations in terms of the image background, age, ethnicity, and images with accessories like eyeglasses, sunglasses, caps etcetera. A total of 1000 images are trained for the two frameworks with two groups of younger age and older age.

Some images are presented in figure 8 from CelebA-HQ and FFHQ dataset.

3.2 *Training and implementation details*

Both dataset are split into a training, testing ratio of 70-30%. The training images are 706 and the testing images are 300 for each experiment are used. System architecture with Nvidia Geforce GTX with one GPU, 1660 Ti is used. The training time was approximately 10 hours for each age group in AttentionGAN. For CycleGAN, it took approximately 12 hours for each age group. Both the models completed 210 epochs. For a fair comparison, the same epochs are used in both models. The optimum training epochs in GAN are dependent on the size of the dataset, dataset type, and application for which GAN has been used [52]. As shown in figures 9(a) and (b) for the FFHQ dataset and figures 9(c), (d) for the CelebA-HQ dataset training loss graph is presented. The training loss plots (figure 9) representing that the loss of G and D should remain consistent throughout the training. That is, the G loss should be greater than the D loss and should not diminish. Diminishing the generator (G) loses infer that the model has become good in generating images while diminishing the discriminator (D) loses indicate that the generator has either become good or the discriminator has not improved in differentiating real or fake face images. Either of these conditions hinders the learning process. Since the requirement is to neither increase nor decrease the G and D losses too much,

therefore, obtaining a training loss graph that looks the same throughout the model's training is required. So, created images in training can be used to select a model and additional training epochs may not essentially mean better quality synthesized images. Finally, the models are trained for 210 epochs. Based on constant monitoring of these training results it is observed that in the beginning, outputs are blurry images that become more and more realistic as training progresses. After the 40th epoch, the synthesized images are visibly improving and age-progressed face features are taking meaningful aging signs. That also should signify that both models have learned well enough.

4. Simulation results

Extensive experiments are performed for the evaluation, to compare the face aging output results of CycleGAN and AttentionGAN.

4.1 *Qualitative evaluation*

4.1.1 *Face aging:*   Figures 10 and figure 11 present the aged face images from the CelebA-HQ (figure 10(a)) and FFHQ (figure 10(b)) dataset which are generated by CycleGAN and AttentionGAN frameworks. The images are generated for the age group of 55+ age group. Though input face images consist of an extensive choice of the people with gender, expression, makeup, pose, and race. The synthesized aged face images are photo-realistic that retain original details of the face such as wrinkles, skin texture, muscles, etc. For illustration, the hair goes grey, and the skin wrinkles appear. Also, all face images can preserve their original identities. While hair color usually changes into grey as the human face ages, it differs from person to person and is based upon the various internal and external factors and the training images. This clarifies why some of the synthesized face images in figures 10 and 11 show few face aging effects. From comparison in figure 10(a), CycleGAN shows a better performance than AttentionGAN. The opposite is however illustrated in figure 10(b) where AttentionGAN presents the aged faces better than CycleGAN. Thus, the dataset also plays a major role along with the framework to generate quality images. Besides, the individual results generated from the CycleGAN and AttentionGAN are shown in figure 11 where AttentionGAN outperforms CycleGAN by generating sharper, realistic aged faces and visually appealing. Since the key contribution of AttentionGAN is that it learns the foreground and preserves the background of an image simultaneously. It depicts that each method independently generates significant results. Besides, increasing the number of images for training and performing the training twice or more can improve the results further.

4.2 *Robustness*

Figure 12 shows the robustness of CycleGAN and AttentionGAN in terms of profile face images, expression, and occlusion. The age-progressed face images are still photo-realistic and true to given inputs. The images are obtained for the 55+ age group. Thus, both the models are robust to pose, expression and occlusion. The CycleGAN and AttentionGAN model takes the input for the whole face and generates the realistic age-progressed images including hair although the previous methods still work with cropped faces without including hair aging [23, 39, 53]. The performance of CycleGAN and AttentionGAN in terms of robustness are similar with little exception, AttentionGAN synthesized face images have minor artifacts.

4.3 *Aging signs*

Figure 13 shows aging signs that occur with age progression in CycleGAN and AttentionGAN. It is illustrated in figures 13(a) and (b) for CycleGAN that it synthesizes smooth aging variations with high fidelity for different parts like showing the lower half part of the face in figure 13(a). In this figure, it is shown that changes in skin texture with deeper wrinkles as age progress also lips become thinner. However, with the appearance of significant aging signs, identity is well preserved. In figure 13(b) the half-face is illustrated to display the performance globally as nasolabial fold appears and are prominent with aging.

Besides, in figures 13(c) and (d) for AttentionGAN, the effect on eye orbit includes wrinkles around the eyes, wide eye orbits are present, and eyebrows become thin in figure 13(c). It also clearly determines the change of age pattern. In figure 13 (d) hair whitening occurs as hair is varied in shape, color, and texture therefore it is tough to model. With age progression expectations, the hair grows wispy and thin. This is also shown in the age progression simulation. It validates the ability to preserve the necessary facial details while aging.

Thus, images are visually realistic in CycleGAN and AttentionGAN. Both CycleGAN and AttentionGAN accomplish aged face significantly.

4.4 *Quantitative evaluation*

4.4.1 *Identity preservation:* Following the convention [25, 27, 29, 41–43] to estimate identity preservation objectively, the online face analysis tool Face++ is used. In this paper, images from FFHQ and CelebA-HQ dataset are used to evaluate the confidence score with the Face++ tool [54] to measure the similarity among the synthesized aged face image and real face image. The high score above the threshold value (76.5), signifies the higher similarity between the two face images [54]. Table 1 shows the confidence score for CycleGAN and AttentionGAN for the

proposed work and graphical representation is presented in figure 14. It is shown in table 2 that the confidence score for CycleGAN and AttentionGAN have values above a threshold which means each framework preserves the identity well in the aged face images. Besides as shown in table 2, values for CycleGAN are better in comparison to AttentionGAN.

4.5 *Image quality assessment*

Fidelity is a significant part to evaluate image generation tasks. Following [35, 37, 55, 56] for the extensive quantitative evaluation, several image quality assessment metrics used are Frechet Inception Distance (FID) [41], Peak Signal to Noise Ratio (PSNR), Structural Similarity Index (SSIM), Universal Quality Image Index (UQI), Visual Information Fidelity (VIF). These are the most frequently used metrics to assess the quality of samples of GANs [35].

FID score tells how real the generated face images are in comparison to real face images. Assuming, (N_t, P_t) and (N_g, P_g) are the mean and covariance of the true images and generated face images features respectively. Thus, mathematically it is expressed as:

$$\text{FID} = \left|N_t - N_g\right|^2 + \text{ tr } \left(P_t + P_g - 2\left(P_t P_g\right)^{1/2}\right). \quad (11)$$

The lower value represents the better quality of the image. The size of the dataset should be larger to compute the FID score [45]. So, following the [45], taking references value for FID score by computing the FID score between the 2 splits of actual image dataset, thus providing the baseline FID score of 87.1 for CelebA-HQ and 132.6 for the FFHQ dataset. The outcomes are illustrated in table 3 by computing the FID score among real and synthesized face images. It is observed that in CelebA-HQ, AttentionGAN has a better FID score and for the FFHQ dataset, CycleGAN has shown a better FID score. The graphical representation for FID values is shown in figure 15(b). As shown, model performance varies with the type of dataset used.

Besides this, quality assessment is evaluated with PSNR. It has been extensively used in numerous digital image measurements. The PSNR is used before SSIM and it is easy. Also, it has been considered tested and valid. Higher PSNR value is better Mathematically, PSNR is expressed in equation (12):

$$PSNR = 10 log10\left(\frac{max^2}{MSE}\right), \quad (12)$$

where max is the highest scale value of an image, MSE is the logarithm function of the mean square error of a given image. Figure 15(a) shows the PSNR values (in dB) for CycleGAN and AttentionGAN and its graphical presentation is shown in figure 15(c). From numerical

Table 1. Literature of GANs based on face aging methods.

GANs	Author/ year	Pros	Cons
CAAE [23]	Zhang Z et al 2017	The face age progression task was performed without the need for paired samples. At the first time, this method achieved face age progression and regression in a general method. It has the potential to assist as a general method in the case of face-age-related tasks. Thus, it can state whether the input face matches a specific age, that is the exact motive of face age estimation.	In this method, the generated output shows rough wrinkles on account of inadequate discriminative and generative capability.
Age-cGAN [22]	Antipov et al 2017	Introduced "Identity-preserving" latent vector optimization method that preserves the unique individual's identity in the reconstruction. It is a universal method means it can be used to preserve identity not only for face aging but also in other face modifications tasks such as the addition of sunglasses, a beard, etc.	Age-cGAN losses the unique individual's identity even before age progression/regression in almost 20% of cases, makes it impossible for practical application to progress in cross-age face verification tasks.
CycleGAN [33]	Zhu et al 2017	Works with an unpaired dataset. Being unsupervised, image quality is good in many translations of the images. Also, it works well with texture and color change.	But geometric changes show little progress.
IPcGAN [24]	Wang et al 2018	IPcGANs can be useful for multi-attribute generation tasks, such as facial expressions, hair colors, etc. It can be used for imbalanced data classification scenes. If the conditional part is removed, this framework can be used for the image translation tasks.	With a data augmentation method, aged faces generated in this method were not able to improve face recognition performance.
Age-DCGAN [38]	Liu et al 2018	Perceptual similarity loss substituted adversarial loss in GANs as the objective function.	The result on age regression was not so good, especially for grown-up males.
Contextual GAN [39]	Liu et al 2018	This method had achieved very well when the given image and target age groups were adjoining. It was because of the proposed Transition Pattern Discriminative network.	This method was slightly weak for synthesizing face images of 60+ age from children's face images.

Table 1 continued

GANs	Author/ year	Pros	Cons
DualGAN [26]	Song *et al* 2018	The primal conditional GAN transforms an input face image to another age, based on the age condition. However, the dual conditional GAN learned to invert the task.	FT demo result outperformed Dual cGANs by 4%.
Wavelet GLCA-GAN [28]	Li *et al* 2019	With the introduction of the frequency domain information, the generated images were stronger also sensitive to facial texture.	The model had difficulty in learning to change the hair color.
A3GAN [40]	Liu Y *et al* 2019	With the use of WPT (Wavelet Packet Transform), the computational cost was significantly reduced. Thus, also improve visual fidelity.	Outcomes attained with setting 'wWPT' still suffered from incorrect facial attributes.
S2GAN [41]	He *et al* 2019	As S2 -module was orthogonal to some methods, thus reduces the computational consumption and enables continuous aging.	Whole personalized aging factors using several images for each individual at diverse ages could be established.
Triple-GAN [42]	Fang *et al* 2020	The triple translation helped in learning age patterns independently. Because it made the correlations among age domains.	Triple translation loss needed supervision otherwise, the distance of domains among input face image and output face image was still so far, results in messy and lost face images.
PFA-GAN [43]	Huang *et al* 2020	Introduced an aging smoothness metric and new age estimation loss. The PFA-GAN can be optimized in an end-to-end manner to eliminate the accumulative error.	The networks needed the source image age label at the input to attain the aging process. Also, splitting into age groups in face aging, made it difficult for end-to-end training. Therefore, the transformation between two adjoining age groups will become less clear.
AGR-GAN [44]	Yadav *et al* 2020	This method included age gap loss and identity loss among the input and the synthesized face images.	The synthesized face images may appear over-smoothed in some cases. This may be credited to the existence of the L1 term in the loss function that has been observed in other methods as well [37]. Another cause for it could be the inadequate amount of training images for different age groups, specifically the young and old age groups.

Table 1 continued

GANs	Author/ year	Pros	Cons
AMGAN [45]	Despois et al 2020	The patch-based method allowed conditional generative adversarial networks to be trained on huge face images though keeping a large batch size. This method was applied to several problems and it could be used to tackle high-resolution difficulties with limited computation resources.	A major drawback of patch-based training was that small patches may look similar such as cheek and forehead. As yet they must be aged differently. For example, horizontal and vertical wrinkles respectively.
Attention GAN [37]	Tang et al 2020	Works with the unpaired dataset, Image quality was good at the output and generates realistic and sharper images. As generators can learn on the foreground of the target domain and preserve the background of the source domain efficiently.	AttentionGAN limitation was shown by scheme A where the model could not handle a complex task for translation.
InterFace GAN [46]	Shen et al 2020	Like many methods manipulating the age, gender, presence of eyeglasses, and expression, this method can even modify the face image pose and fix the GANs artifacts which were accidentally produced.	This method may fail for long-distance manipulation due to the linear assumption.
EigenGAN [47]	He et al 2021	EigenGAN embeds one linear subspace with an orthogonal basis into each generator layer.	Discovered semantic attributes are not always the same at different training times in two cases: in gender and pose learning. Sometimes the model can discover a specific attribute but sometimes cannot, such as eyeglasses.
MTLFace [48]	Huang et al 2021	MTLFace (multi-task learning framework), able to focus on age-invariant identity-related representation and achieves notable face synthesis.	Since in general, the GANs still face the training problem.
CFA-GAN [49]	Jeon et al 2021	The new loss function for identity preservation maximizes the cosine similarity among the given input image and the synthesized identity source features.	It is observed that the age errors are comparatively high with the target ages larger than 30. It can be because of data imbalance.
pixel2style2pixel (pSp) [50]	Richardson et al 2021	The encoder can straight embed actual images into W+, with no added optimization.	The high-quality face images that are produced with the use of pre-trained StyleGAN come with a cost. This technique is limited to images that can be synthesized by StyleGAN.
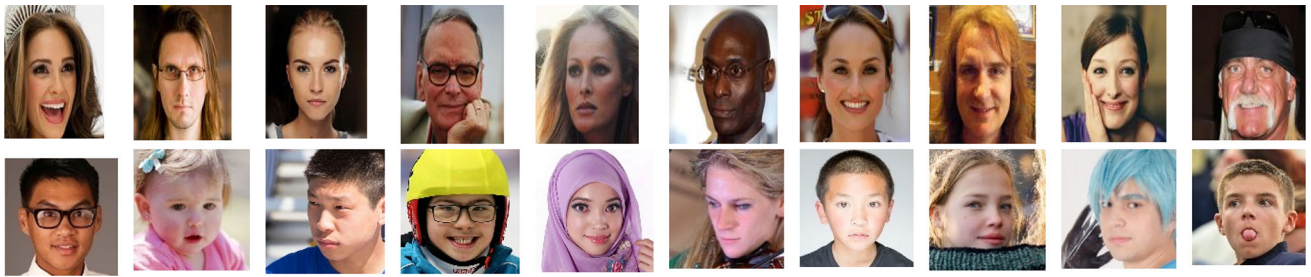
Figure 8. Represents some images from the CelebA-HQ dataset in the upper row and the lower row presents some images from FFHQ dataset.
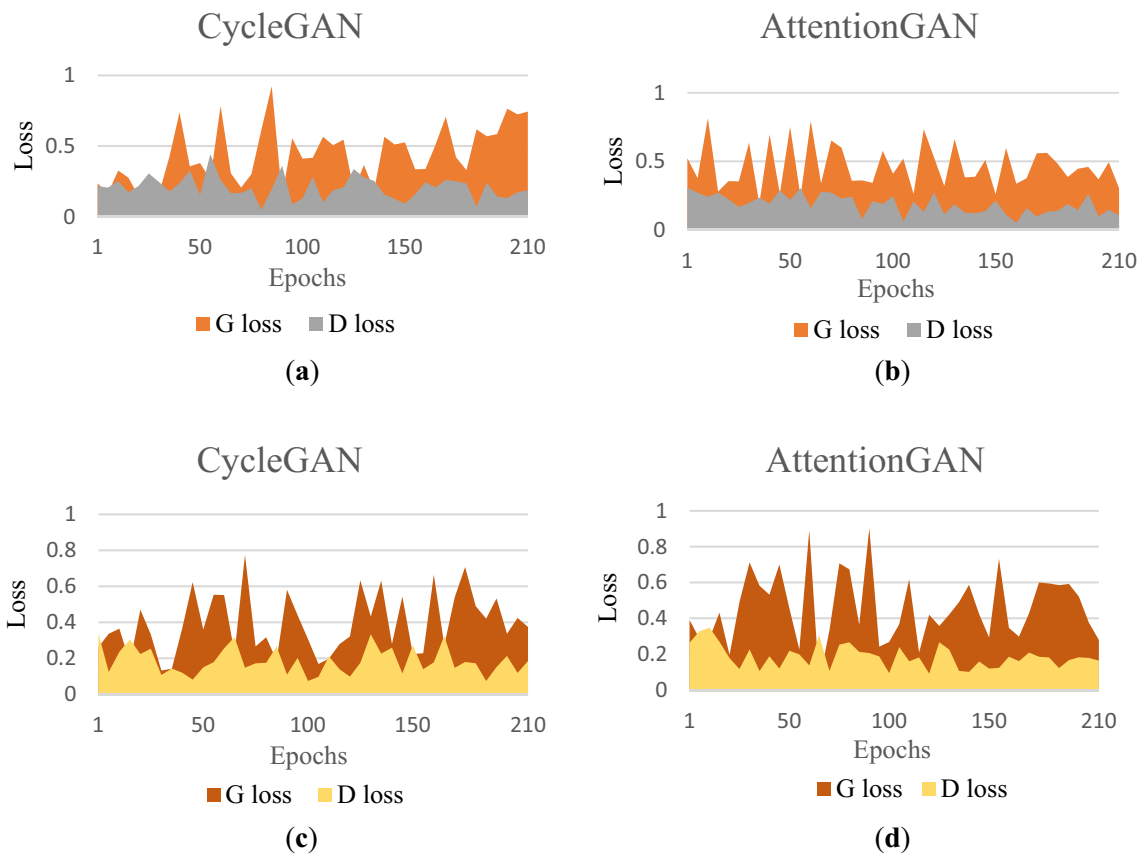
Figure 9. Training loss graph (**a**) CycleGAN, (**b**) AttentionGAN with the FFHQ dataset, (**c**) CycleGAN and (**d**) AttentionGAN with the CelebA-HQ dataset.

simulations, it is clear that AttentionGAN performance is better than CycleGAN. This shows the quality of synthesized images is better (as presented in figure 11). Since, the AttentionGAN generator is the same as CycleGAN with some modifications i.e., built-in generators (figure 7) which help to learn the foreground and preserve the background details of an image while the conversion process.

Further, SSIM measurement is considered depending upon three factors i.e., contrast (C), luminance (L), and structure (S) (equation 13) to be acceptable with the working of the human visual system [57]. The SSIM values range between 0 and 1, in which 1 means a perfect match of the generated image with the original image. The mathematical expression of SSIM is expressed in equations (13) and (14) as:

(a)
Synthesized results obtained from CelebA-HQ

(b)
Synthesized results obtained from FFHQ
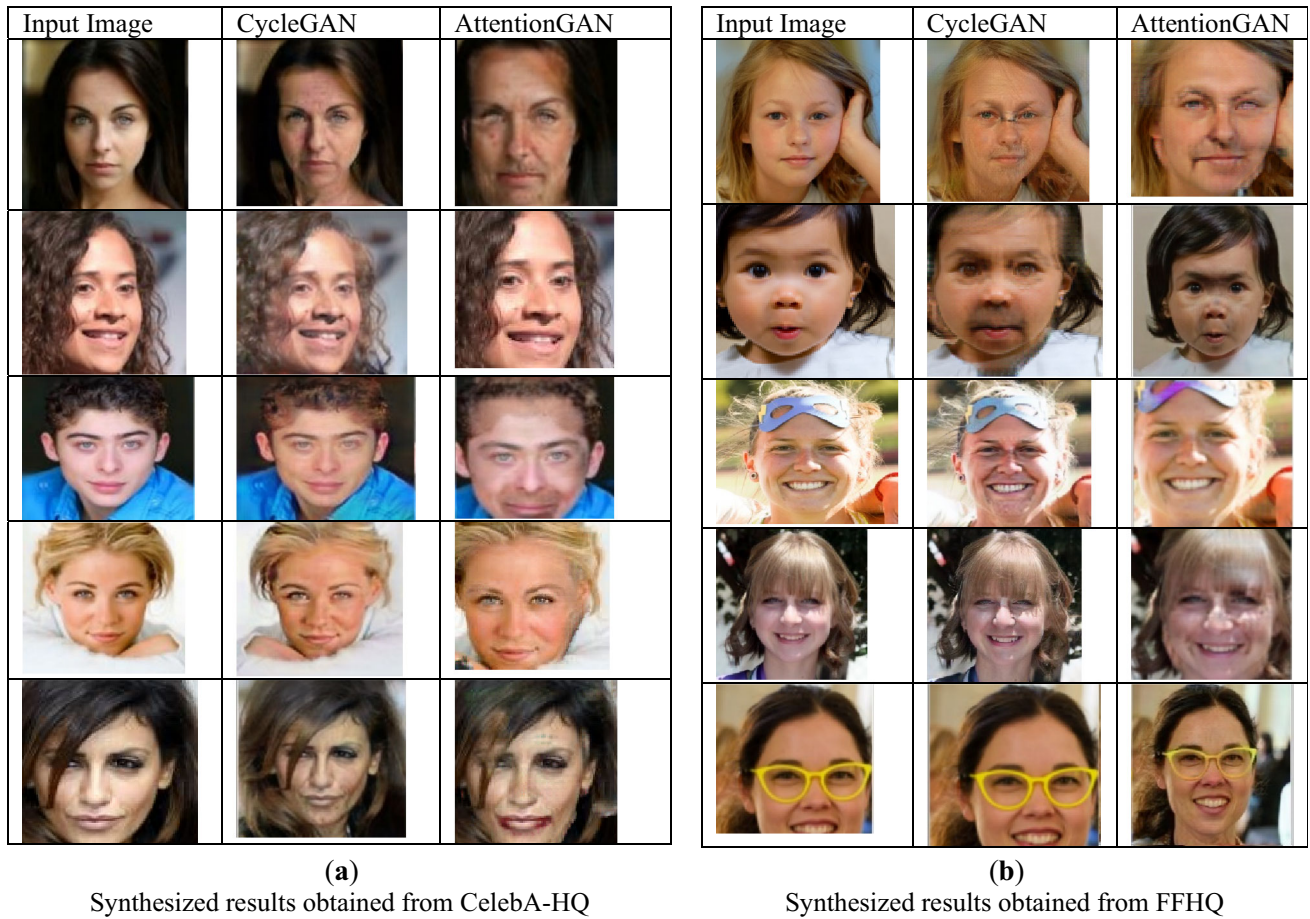
Figure 10. Comparison of images generated from AttentionGAN and CycleGAN.

Figure 11. Synthesized results from the CycleGAN and the AttentionGAN with CelebA-HQ and FFHQ dataset.
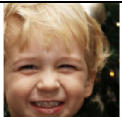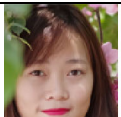
	CycleGAN				AttentionGAN			
Side pose								
Expression								
Occlusion								

Figure 12. Robustness of CycleGAN and AttentionGAN with side pose, expression, and occlusion.

(a)   (c)

(b)   (d)

CycleGAN   AttentionGAN

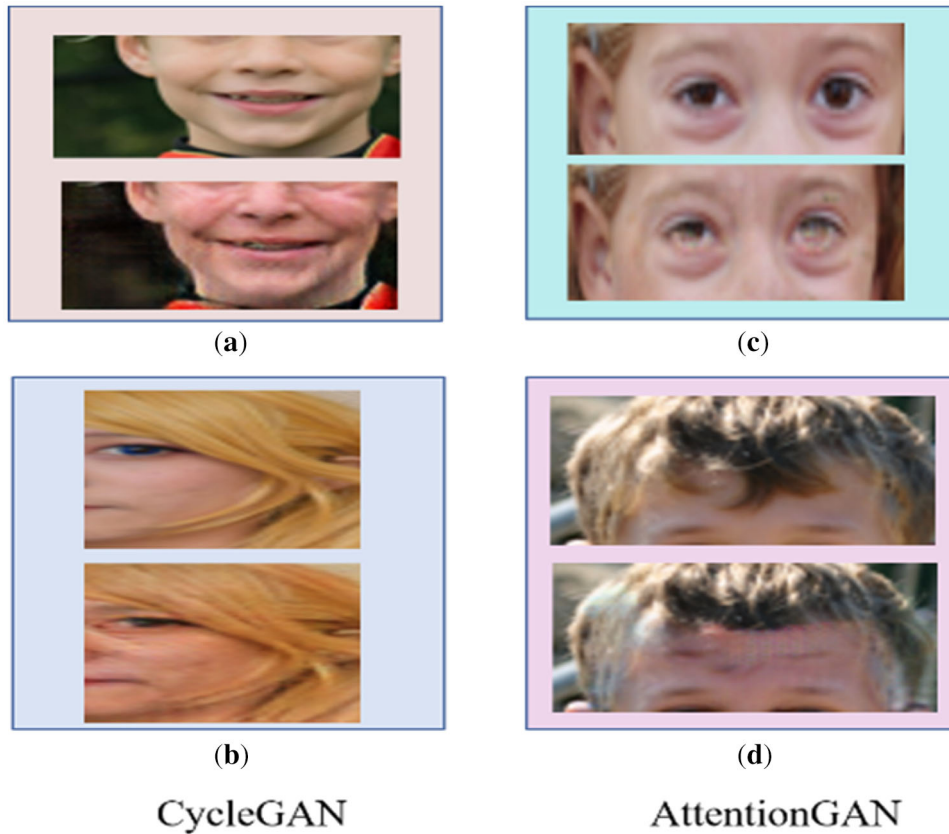**Figure 13.** Face aging signs: In CycleGAN (**a**) Cheeks wrinkles and skin texture change, (**b**) Nasolabial fold appears with aging. In AttentionGAN (**c**) Wide deep orbit, eyebrows thin with age progression and (**d**) Hair whitening with aging.

Confidence Score

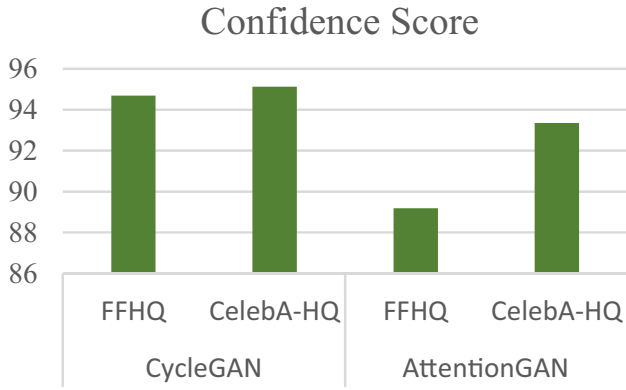**Figure 14.** Graphical representation of confidence score for CycleGAN and AttentionGAN using FFHQ and CelebA-HQ dataset.

$$SSIM(q,\ q') = L(q,\ q')C(q,q')S(q,q'), \qquad (13)$$

where L, C, S are functions that compare the image q and image q' [(q, q')] for luminance, contrast, and structure.

$$SSIM(q,q') = \frac{\left(2u_q u_{q'} + C_1\right)\left(2\sigma_{qq'} + C_2\right)}{\left(u_q^2 + u_{q'}^2 + C_1\right)\left(\sigma_q^2 + \sigma_{q'}^2 + C_2\right)}, \qquad (14)$$

where q is real image intensity value and q' is generated image intensity value with a respected mean ($u_q u_{q'}$) and variances ($\sigma_{qq'}$). C1, C2, and C3 are constant values used to avoid the zero denominators [57]. Figures 15 (a)-(d) represent the CycleGAN and AttentionGAN with SSIM values and their graphical representation respectively. Here, CycleGAN SSIM values are better than the AttentionGAN for both the dataset. As the advantage of CycleGAN, it works well for texture and color changes [33].

VIF (visual information fidelity), points out that the results were more similar to the targeted output and had a recognizable structure and visually pleasing images. Mathematically, VIF is expressed as:

$$VIF = \frac{\sum_{j \in sub\,bands} I(\vec{C}^{R,k}; \vec{F}^{R,k} | S^{R,k})}{\sum_{j \in sub\,bands} I(\vec{C}^{R,k}; \vec{E}^{R,k} | S^{R,k})}, \qquad (15)$$

where the numerator and denominator are the information extracted from the given image and generated face

Table 3. Evaluation with Frechet Inception Distance (FID), lower is better.

Method	FFHQ	CelebA-HQ
Real data	132.6	87.1
CycleGAN	**50.4**	60.8
AttentionGAN	66.7	**55.1**

images respectively. $\vec{C}^{R,k}$ represents R elements of the RF (random field) and C_k describes the coefficients from subband k. Therefore, VIF will deliver the image quality measure.

For VIF, a higher value is better, figures 15 (a)-(e) show the CycleGAN and AttentionGAN VIF values. It depicts that CycleGAN and AttentionGAN have shown better results with the FFHQ dataset. Thus, model performance depends upon the type of dataset and the quality of the dataset at the input.

UQI (universal quality image index) was considered to model any image distortion with the grouping of three factors: loss of correlation, contrast distortion, and luminance distortion. Mathematically, UQI (Q) is expressed as:

$$Q = \frac{\sigma_{qq'}}{\sigma_q \sigma_{q'}} \frac{2\overline{q}\overline{q'}}{(\overline{q})^2 (\overline{q'})^2} \frac{2q\sigma_{q'}}{\sigma_q^2 + \sigma_{q'}^2}. \qquad (16)$$

This provides the map of Qs, further average value of this map gives the quality measure of Q which is expressed as:

$$Q = \frac{1}{M} \sum_{j=1}^{M} Q_j, \qquad (17)$$

where M is steps depending on the size of an image. As SSIM was built from UQI, it is noticed that the result presented in figure 15(a) by UQI is somewhat closer to 1 than SSIM.

Higher value signifies better quality. Figures 15(a)-(f) show the CycleGAN and AttentionGAN UQI values. It shows that both the models have small differences in output values. But for the overall performance of UQI, AttentionGAN is better than CycleGAN.

Table 2. The confidence score for CycleGAN and AttentionGAN.

Model	CycleGAN		AttentionGAN	
Dataset	FFHQ	CelebA-HQ	FFHQ	CelebA-HQ
Confidence Score	94.69 ±1.68	95.13±0.42	89.19 ±4.68	93.35 ± 1.03

Metrics	Dataset	FFHQ	CelebA-HQ
PSNR (dB)	CycleGAN	25.22	22.97
	AttentionGAN	**26.82**	**28.31**
SSIM	CycleGAN	**0.89**	**0.86**
	AttentionGAN	0.82	0.82
VIF	CycleGAN	**0.52**	0.45
	AttentionGAN	**0.52**	0.49
UQI	CycleGAN	**0.96**	0.92
	AttentionGAN	**0.99**	0.94

(a)
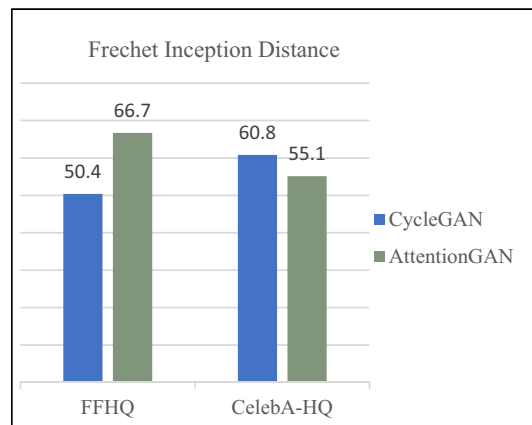
(b)
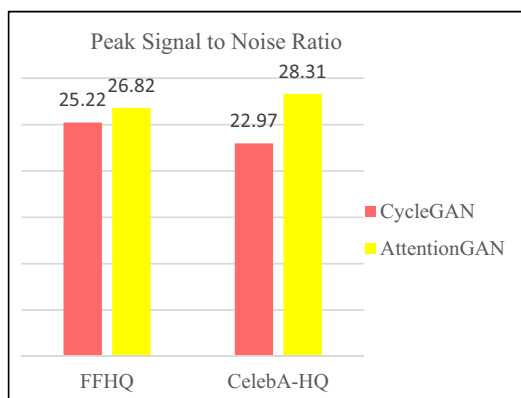
(c)
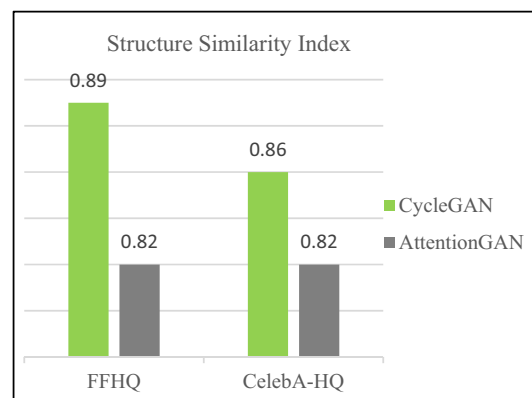
(d)
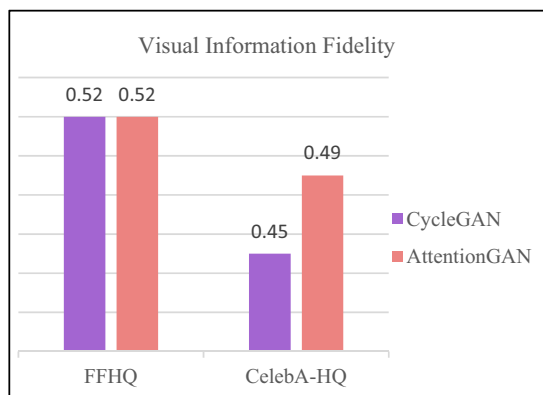
(e)
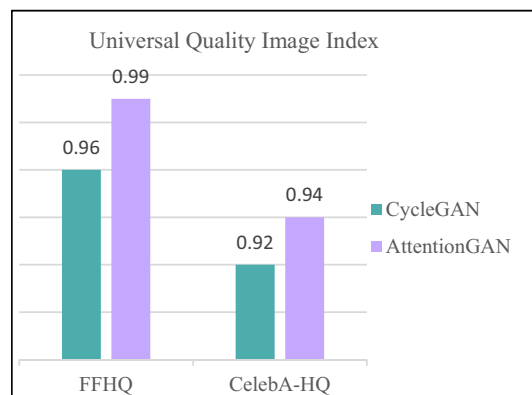
(f)

Figure 15. (**a**) Quantitative results with image quality assessment score, (**b**) FID, (**c**) PSNR, (**d**) SSIM, (**e**) VIF, (**f**) UQI are its graphical representation.

4.5.1 *Subjective evaluation:* For evaluation of the results, following the work [23–25, 41, 42] a user study was performed. Randomly selected 10 original face images with their corresponding generated images were used for the age group of 55+. Then, 20 evaluators had assessed the aged face images from CycleGAN and AttentionGAN regarding reference to the input face image. Finally, volunteers were asked to select realistic aged face images from CycleGAN and AttentionGAN having natural aging effects and lesser artifacts. The faces with lesser artifacts and more natural aging effects voted with 57% for CycleGAN, 36% AttentionGAN, and 7% for voted that both the model has shown overall equal results. In few output

images, AttentionGAN synthesized images are better than CycleGAN. Further, the model performance depends greatly upon the number of training images also large dataset can better yield the results, also the type of dataset matters most.

5. Pros and Cons of CycleGAN and AttentionGAN

GANs	Pros/Cons
CycleGAN	*Pros:*
	CycleGAN's biggest advantage is that it can produce remarkable results without the need for a paired dataset. Moreover, it works well for texture and color changes [33]. For identity preservation (Table 2), FID score (Table 3, figure 15(b)) and SSIM (figures 15(a)-(d)) CycleGAN is a better performer than AttentionGAN in this paper.
	Cons:
	In this paper, the training time required by CycleGAN is more for the same number of epochs and training images. Because this paper focuses on face age progression only, CycleGAN performs bidirectional translation simultaneously (converting the images for progression as well as for regression process.)
AttentionGAN	*Pros:*
	AttentionGAN works for the unpaired dataset. As illustrated in figure 11, individual results synthesized by AttentionGAN are better than CycleGAN. Individually, it has produced more realistic and sharper aged face images.
	Cons:
	AttentionGAN produces artifacts in some images. CycleGAN images are better than AttentionGAN (figure 10).

6. Conclusion and future scope

The image-to-image conversion process with GANs had done exponential development. The unsupervised algorithms can produce remarkable results in comparison to supervised learning without paired data. In this paper, it is shown that the results generated for the comparison between the CycleGAN and AttentionGAN for face aging task vary for each dataset. Visually realistic and significant aging signs are shown by both the CycleGAN and AttentionGAN models. However, some results in AttentionGAN show artifacts in comparison to CycleGAN. But the

individual results obtained from the AttentionGAN are better than CycleGAN. So, it is hard to claim strongly why one or the other performs slightly better than the other one, as model performance depends on several factors. Overall, CycleGAN performance the better than AttentionGAN in this paper.

Some face aging problems that lead to further improvements. As research is a continuous process, the proposed work can be further continued with the different dataset for various applications, like ethnicity-based face aging. Mostly the dataset used in face aging methods is a mixture of ethnic groups so there is a difference between the face and its various feature in shape across ethnicities. Thus, collecting the more specific dataset like ethnic-based can be helpful for more detailed study in the face aging application. Moreover, the comparison experiment on the different variants of GANs can be helpful for the research community further. The different dataset shows the varying output results on the same model. The proposed work is done on 2D images. So, it can be further extended to 3D images. But it may require a large processing time and increase the memory usage because 3D face images require the 3D face scanner to obtain the 3D images of the face [11]. As the 3D faces can describe the face structure and associated texture clearly that is why it can generate more photo-realistic and precise results. The objective function for evaluating the performance of GANs is still an open problem [58]. Besides, human face aging depends on internal and external factors. Thus, due to emotions of human some face muscles contract and expand which cause the development of lines and wrinkles. These lines and wrinkles with age progression become permanent even when the face relaxes. Thus, a study of wrinkles on the individual faces with facial emotion expression changes can be supportive for the face aging process. A healthy lifestyle with a healthy diet and exercise slows the aging process in comparison to the smoker and alcoholic persons. Thus, to precisely predict the face in the future of the person there is also a necessity to link with external factors and tune them in automatic face age progression method to produce more realistic results. A fusion can be also helpful for extracting rich information from the face of the person for recognizing which part of the face aging faster.

References

[1] Aggarwal A, Mittal M and Battineni G 2021 Generative adversarial network: An overview of theory and applications. *Int. J. Inf. Manag. Data Insights*. https://doi.org/10.1016/j.jjimei.2020.100004

[2] Wang M, Chen Z, Wu Q M J and Jian M 2020 Improved face super-resolution generative adversarial networks. *Mach. Vis. Appl.* 31(4): 1–12. https://doi.org/10.1007/s00138-020-01073-6

[3] Tolosana R, Vera-Rodriguez R, Fierrez J, Morales A and Ortega-Garcia J 2020 Deepfakes and beyond: A Survey of face manipulation and fake detection. *Inf. Fus.* 64: 131–148. https://doi.org/10.1016/j.inffus.2020.06.014

[4] Fu Y, Guo G and Huang T S 2010 Age synthesis and estimation via faces: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* 32(11): 1955–1976. https://doi.org/10.1109/TPAMI.2010.36

[5] Farage M A, Miller K W, Elsner P and Maibach H I 2008 Intrinsic and extrinsic factors in skin ageing: A review. *Int. J. Cosmet. Sci.* 30(2): 87–95. https://doi.org/10.1111/j.1468-2494.2007.00415.x

[6] Ramanathan N and Chellappa R 2008 Modeling shape and textural variations in aging faces. In: *2008 8th IEEE International Conference on Automatic Face and Gesture Recognition*:1-8. https://doi.org/10.1109/AFGR.2008.4813337

[7] Farazdaghi E 2017 *Facial ageing and rejuvenation modeling including lifestyle behaviours, using biometrics-based approaches.* PhD diss, Signal and Image Processing. Université Paris-Est, 2017. English. NNT: 2017PESC1236ff. tel-01760426

[8] Rawlings A V 2006 Ethnic skin types: are there differences in skin structure and function? *Int. J. Cosmet. Sci.* 28(2): 79–93. https://doi.org/10.1111/J.1467-2494.2006.00302.X

[9] Vashi N A, Maymone M B de C and Kundu R V 2016 Aging Differences in Ethnic Skin. *J. Clinic. Aesthet. Dermatol.* 9(1):31–38. PMID:26962390, PMCID: PMC4756870

[10] Zouboulis C C 2000 Human skin: An independent peripheral endocrine organ. *Hormone Res. Paediatr.* 54(5–6): 230–242. https://doi.org/10.1159/000053265

[11] Dayan N 2008 Skin Aging Handbook: An Integrated Approach to Biochemistry and Product Development. *William Andrew.* ISBN 978-0-8155-1584-5

[12] Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al 2014 Generative Adversarial Nets. *Adv. Neural Inf. Process. Syst.* 27

[13] Pieters M and Wiering M 2018 Comparing Generative Adversarial Network Techniques for Image Creation and Modification. arXiv:1803.09093

[14] Creswell A, White T, Dumoulin V, Arulkumaran K, Sengupta B and Bharath A A 2018 Generative adversarial networks: An overview. *IEEE Signal Process. Mag.* 35(1): 53–65. https://doi.org/10.1109/MSP.2017.2765202

[15] Alqahtani H, Kavakli-Thorne M and Kumar G 2021 Applications of Generative Adversarial Networks (GANs): An updated review. *Arch. Comput. Methods Eng.* 28(2): 525–552. https://doi.org/10.1007/S11831-019-09388-Y

[16] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D and Ozair S *et al.* 2020 Generative adversarial networks. *Commun. ACM* 63(11): 139–144. https://doi.org/10.1145/3422622

[17] Guo J, Pang Z, Yang F, Shen J and Zhang J 2020 Study on the method of fundus image generation based on improved GAN. *Math. Probl. Eng.* 2020 article ID 6309596. https://doi.org/10.1155/2020/6309596

[18] Duong C N, Luu K, Quach K G, and Bui T D 2016 Longitudinal face modeling via temporal deep restricted Boltzmann machines. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*: 5772–5780. https://doi.org/10.1109/CVPR.2016.622

[19] Wang W, Cui Z, Yan Y, Feng J, Yan S, Shu X and Sebe N 2016 Recurrent Face Aging. In: *IEEE Conference on Computer Vision and Pattern Recognition*:2378-2386. https://doi.org/10.1109/CVPR.2016.261

[20] Duong C N, Quach K G, Luu K, Le N and Savvides M 2017 Temporal Non-Volume Preserving Approach to Facial Age-Progression and Age-Invariant Face Recognition. In: *Proceedings of the IEEE International Conference on Computer Vision*: 3735–3743. https://doi.org/10.1109/ICCV.2017.403

[21] Mirza M and Osindero S 2014 Conditional Generative Adversarial Nets. arXiv:1411.1784

[22] Antipov G, Baccouche M and Dugelay J-L 2017 Face aging with conditional generative adversarial networks. In: *Proceedings of IEEE International Conference on Image Processing, (ICIP):*2089–2093. https://doi.org/10.1109/ICIP.2017.8296650

[23] Zhang Z, Song Y and Qi H 2017 Age progression/regression by conditional adversarial autoencoder. In: *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR):*4352-4360. https://doi.org/10.1109/CVPR.2017.463

[24] Wang Z, Tang X, Luo W and Gao S 2018 Face Aging with Identity-Preserved Conditional Generative Adversarial Networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*:7939–7947. https://doi.org/10.1109/CVPR.2018.00828

[25] Yang H, Huang D, Wang Y and Jain A K 2018 Learning Face Age Progression: A Pyramid Architecture of GANs. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition:*31–39. doi: https://doi.org/10.1109/CVPR.2018.00011

[26] Song J, Zhang J, Gao L, Liu X and Shen H T 2018 Dual Conditional GANs for Face Aging and Rejuvenation. In: *Proceedings of the Twenty-seventh International Joint Conference on Artificial Intelligence (IJCAI):*899–905. https://doi.org/10.24963/ijcai.2018/125

[27] Liu Y, Li Q and Sun Z 2019 Attribute-aware face aging with wavelet-based generative adversarial networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition:* 11869–11878. https://doi.org/10.1109/CVPR.2019.01215

[28] Li P, Hu Y, He R, and Sun Z 2019 Global and Local Consistent Wavelet-Domain Age Synthesis. In: *IEEE Transactions on Information Forensics and Security.* 14(11):2943–2957. https://doi.org/10.1109/TIFS.2019.2907973

[29] Li Q, Liu Y and Sun Z 2020 Age Progression and Regression with Spatial Attention Modules. In: *Proceedings of the AAAI Conference on Artificial Intelligence.* 34(7):11378–11385

[30] Sun Y, Tang J, Shu X, Sun Z and Tistarelli M 2020 Facial age synthesis with label distribution-guided generative adversarial network. *IEEE Trans. Inf. Forens. Secur.* 15: 2679–2691. https://doi.org/10.1109/TIFS.2020.2975921

[31] Kemelmacher-Shlizerman I, Suwajanakorn S and Seitz S M 2014 Illumination-aware age progression. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*: 3334–3341. https://doi.org/10.1109/CVPR.2014.426

[32] Suo J, Zhu S-C, Shan S and Chen X 2009 A compositional and dynamic model for face aging. *IEEE Trans. Pattern*

Anal. Mach. Intell. 32(3): 385–401. https://doi.org/10.1109/TPAMI.2009.39

[33] Zhu J Y, Park T, Isola P and Efros A A 2017 Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. In: *Proceedings of the IEEE International Conference on Computer Vision:* 2242–2251 https://doi.org/10.1109/ICCV.2017.244

[34] Welander P, Karlsson S and Eklund A 2018 Generative Adversarial Networks for Image-to-Image Translation on Multi-Contrast MR Images - A Comparison of CycleGAN and UNIT. arXiv:1806.07777

[35] Nanavati T, Modi H, Patel D, Parikh V and Gupta J 2020 Generative adversarial networks: A comparative analysis generative adversarial networks: a comparative analysis. *Int. J. Adv. Res. Comput. Eng. Technology (IJARCET).* 9(4): 2278–1323

[36] Burad Y and Burad K 2020 A comparative study of CycleGAN and Progressive growing GAN for synthetic data generation. *Int. J. Eng. Applied. Sci. Technol.* 5(3) ISSN no. 2455-2143:657-660

[37] Tang H, Liu H, Xu D, Torr P H S and Sebe N 2021 AttentionGAN: Unpaired image-to-image translation using attention-guided generative adversarial networks. *IEEE Trans. Neural Netw. Learn. Syst..* https://doi.org/10.1109/TNNLS.2021.3105725

[38] Liu X, Xie C, Kuang H and Ma X 2018 Face Aging Simulation with Deep Convolutional Generative Adversarial Networks. In: *Proceedings of the 10th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA) IEEE*:220–224 https://doi.org/10.1109/ICMTMA.2018.00060

[39] Liu S, Sun Y, Zhu D, Bao R, Wang W, Shu X and Yan S 2017 Face aging with contextual generative adversarial nets. In: *MM 2017 - Proceedings of the 25th ACM International Conference on Multimedia*:82–90. https://doi.org/10.1145/3123266.3123431

[40] Liu Y, Li Q, Sun Z and Tan T 2021 A3GAN: An attribute-aware attentive generative adversarial network for face aging. *IEEE Trans. Inf. Forens. Secur.* 16: 2776–2790. https://doi.org/10.1109/TIFS.2021.3065499

[41] He Z, Kan M, Shan S and Chen X 2019 S2GAN: Share aging factors across ages and share aging trends among individuals. *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* :9439–9448 https://doi.org/10.1109/ICCV.2019.00953

[42] Fang H, Deng W, Zhong Y and Hu J 2020 Triple-GAN: Progressive face aging with triple translation loss. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*:3500-3509. https://doi.org/10.1109/CVPRW50498.2020.00410

[43] Huang Z, Chen S, Zhang J and Shan H 2020 PFA-GAN: Progressive Face Aging with Generative Adversarial Network. *IEEE Trans. Inf. Forens. Secur.* 16: 2031–2045. https://doi.org/10.1109/TIFS.2020.3047753

[44] Yadav D, Kohli N, Vatsa M, Singh R and Noore A 2021 Age Gap Reducer-GAN for Recognizing Age-Separated Faces.

2020 25th International Conference on Pattern Recognition (ICPR) IEEE:10090-10097. doi: https://doi.org/10.1109/ICPR48806.2021.9412078

[45] Despois J, Flament F and Perrot M 2020 AgingMapGAN (AMGAN): High-resolution controllable face aging with spatially-aware conditional GANs. In: *European Conference on Computer Vison*:613-628. Springer, Cham

[46] Shen Y, Yang C, Tang X and Zhou B 2020 InterFaceGAN: Interpreting the disentangled face representation learned by GANs. *IEEE Trans. Pattern Anal. Mach. Intell..* https://doi.org/10.1109/TPAMI.2020.3034267

[47] He Z, Kan M and Shan S 2021 EigenGAN: Layer-Wise Eigen-Learning for GANs. arXiv:2104.12476

[48] Huang Z, Zhang J and Shan H 2021 When age-invariant face recognition meets face age synthesis: a multi-task learning framework. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition:*7282-7291

[49] Jeon S, Lee P, Hong K and Byun H 2021 Continuous Face Aging Generative Adversarial Networks. *International Conference on Acoustic, Speech and Signal Processing (ICASSP):*1995–1999. https://doi.org/10.1109/icassp39728.2021.9414429

[50] Richardson E, Alaluf Y, Patashnik O, Nitzan Y, Azar Y, Shapiro S and Cohen-Or D 2021 Encoding in Style: A StyleGAN Encoder for Image-to-Image Translation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*:2287–2296

[51] Karras T, Aila T, Laine S and Lehtinen J 2017 Progressive growing of gans for improved quality, stability and variation. arXiv:1710.10196

[52] https://github.com/junyanz/pytorch-CycleGAN-and-pix2pix/issues/1127

[53] Elmahmudi A and Ugail H 2021 A framework for facial age progression and regression using exemplar face templates. *Vis. Comput.* 37(7): 2023–2038. https://doi.org/10.1007/s00371-020-01960-z

[54] Megvii Inc. 2013 Face++ research toolkit. http://www.faceplusplus.com/

[55] Zhang H, Riggan B S, Hu S, Short N J and Patel V M 2019 Synthesis of high-quality visible faces from polarimetric thermal faces using generative adversarial networks. *Int. J. Comput. Vis.* 127(6): 845–862

[56] Khan A, Jin W, Haider A, Rahman M and Wang D 2021 Adversarial gaussian denoiser for multiple-level image denoising. *Sensors* 21(9): 2998. https://doi.org/10.3390/s21092998

[57] Setiadi D R I M 2021 PSNR vs SSIM: imperceptibility quality assessment for image steganography. *Multimed. Tools Appl.* 80: 8423–8444. https://doi.org/10.1007/s11042-020-10035-z

[58] Alqahtani H, Kavakli-Thorne M and Kumar G 2019 An analysis of evaluation metrics of GANs. In: *International Conference on Information Technology and Applications (ICITA).* 7