Journal of Material Science and Mechanical Engineering (JMSME)
p-ISSN: 2393-9095; e-ISSN: 2393-9109; Volume 10, Issue 2; April-June, 2023, pp. 94-103

http://www.krishisanskriti.org/Publication.html

Frontal Face Generation: A Survey

Vivek Bhand[1] and Dr. Tanuja R. Pattanshetti[2]

[1]B.Tech. Computer Engineering, COEP Technological University Pune
[2]Assistant Professor, Computer Engineering, COEP Technological University Pune
E-mail: [1]bhandvv19.comp@coep.ac.in, [2]trp.comp@coep.ac.in

Abstract—*Face recognition is a critical task in various applications, including security systems, surveillance, and social media. However, the performance of face recognition algorithms can be limited when working with side view profile images. To address this challenge, researchers have explored the use of 3D modelling, deep learning techniques such as CNN, generative adversarial networks (GANs) to generate frontal face images from side view profiles. This review paper provides a comprehensive overview of the current state-of-the-art techniques and research papers in this field.*

Firstly, the paper discusses the challenges of face recognition from side view profiles and how generating frontal faces can improve its efficiency. Secondly, the paper provides an overview of the different techniques and how it has been used for face generation. Thirdly, the paper reviews various GAN-based approaches for generating frontal faces, including conditional GANs, progressive GANs, and TP-GANs.

Additionally, the paper provides insights into the evaluation metrics used for measuring the quality of generated images, such as the Fréchet Inception Distance (FID) and Perceptual Path´ Length (PPL). Furthermore, the paper discusses the datasets used for training and testing the GAN-based frontal face generation models, including Multi-PIE, AFLW, and 300W-LP.

Finally, the paper summarizes the current challenges and future research directions in this field, including improving the quality of generated images, addressing variations in facial expressions and lighting conditions, and exploring the potential of using GANs for face recognition tasks. Overall, this review paper provides a comprehensive understanding of the use of GANs for generating frontal faces from side view profiles and its potential implications for improving face recognition systems.

Index Terms: *Frontal face synthesis, Side-view to frontal face conversion, Face recognition, Generative adversarial networks (GANs), Lighting conditions, Evaluation metrics, Dataset specification, Multi-PIE, AFLW.*

1. INTRODUCTION

Facial recognition technology has become an important element of our daily lives. Face recognition is being used in a variety of applications, from unlocking our cell phones to airport security checks. Machine learning has considerably increased the accuracy and reliability of facial recognition systems. Yet, various factors such as stance, illumination, and occlusion continue to limit the efficacy of face recognition systems. Dealing with non-frontal photos is one of the most difficult difficulties in face recognition. Non-frontal photos are those in which the subject's face is not directly facing the camera. Because they capture the face from an unusual angle, these photos might be hard to recognize and identify. Face frontalization has become necessary as a result of this challenge. [21] The technique of converting non-frontal photographs to frontal images is known as frontalization of faces.

This is accomplished through the use of machine learning techniques such as Generative Adversarial Networks (GANs) [14], [20], [23], [26] and Convolutional Neural Networks (CNNs) (CNNs) [1]. By decreasing the effect of position variation, the frontalization procedure helps to improve the performance of face recognition systems.

A. Need of face recognition

Face recognition has shown more growth due to its wide variety of applications. Security, biometrics, law enforcement, and social media all make use of the technology. Face recognition is used in the security domain for access control, identification, and authentication. Face recognition is used in biometrics for identity verification, passport control, and border security. [13] [9] Face recognition is used for crime fighting to identify suspects from CCTV footage, photos collected by bodycams, and forensic analysis. Face recognition systems' precision and dependability are essential in these applications because they can have serious repercussions. Individual misidentification can lead to wrongheaded arrests, invasions of privacy, and potentially the endangering of life.

B. The need for frontalization of faces

The pose of the face in the photograph has a considerable impact on the performance of facial recognition systems. The accuracy of the face recognition system degrades when the face is not directly facing the camera. One of the key issues in face recognition systems is stance variation. Face frontalization is the solution to this problem.

Face frontalization is the process of converting non-frontal photographs to frontal ones. The procedure entails evaluating the attitude of the face in the photograph and converting it to a frontal perspective. Frontalization can aid in reducing the influence of position variation and improving the accuracy of face recognition systems.

C. Example where frontalization of face could have been used:

Frontalization of faces may have been employed in the identification of suspects in law enforcement, for example. Officers frequently encounter non-frontal photographs of suspects when studying CCTV footage or images obtained by body-worn cameras. Traditional facial recognition algorithms may struggle to analyse these photos, resulting in low accuracy and false positives.

The accuracy of facial recognition algorithms can be considerably enhanced by frontalizing these photos using machine learning techniques such as GANs or CNNs. This can result in more accurate suspect identification and help law enforcement authorities solve crimes more successfully.

D. Techniques involved.

Many machines learning techniques, including as GANs and CNNs, are used in the frontalization process. GANs are deep learning architectures made up of two neural networks, one generator and one discriminator. The generator network creates synthetic images, while the discriminator network differentiates between synthetic and genuine images. In GANs, the generator network is utilised to convert the non-frontal image to a frontal view. To assess the quality of the frontalized image, the discriminator network is used. The GANs are trained on a large dataset of non-frontal and frontal images throughout the frontalization process. CNNs are deep neural networks developed to interpret and evaluate picture input. They have been used to learn complicated picture characteristics and patterns, as well as to perform a variety of image processing tasks such as face recognition and frontalization.

Technique	Advantages	Limitations
GANs	High-quality results, can handle large variations in pose and illumination	Requires large amounts of training data, can suffer from mode collapse
CNNs	Fast and efficient, can handle real-time applications	Limited by the complexity of the model, may not generalize well to new. faces
3D Morphable Models	Accurate and realistic, can handle large variations in pose and expression	Requires 3D facial models, computationally expensive

Overall, face frontalization is an important method to improve the efficiency of face recognition systems, especially in practical circumstances where non-frontal photos are frequent. Methods based on machine learning like GANs and CNNs have shown significant potential in generating higher quality frontal photographs from non-frontal images. In the sections that follow, we will examine the existing literature on the use of various techniques for face frontalization and evaluate their performance and limitations.

Given the significance of face recognition technology in today's society, it is critical to investigate and develop strategies to improve its performance. Frontalization of faces is one such approach, which entails converting a non-frontal facial picture to a frontal perspective. This procedure is especially significant because many face recognition algorithms rely on frontal face photos to identify people.

Frontalization has been accomplished using a variety of methodologies, including geometry-based methods, deep learning-based methods, and 3D modelling-based methods [18]. Yet, deep learning approaches such as Generative Adversarial Networks (GANs) [] and Convolutional Neural Networks (CNNs) have demonstrated promising results in frontalization of faces in recent years.

CNNs have been used to learn sophisticated translations between non-frontal and frontal face images, while GANs have been used to generate realistic frontal face images from non-frontal face images. These methods may help to improve the accuracy and reliability of face recognition algorithms, especially when circumstances involving non-frontal face photos, such as surveillance footage, social media images, and forensic investigations.

The identification of the Boston Marathon bombers in 2013 is one such real-world case where frontalization of faces could have been beneficial. The two suspects were captured in non-frontal perspectives on surveillance footage, making it more difficult for law enforcement to identify them. It would have been easier to identify the perpetrators and maybe prevent the fatal tragedy if the footage had been frontalized.

To conclude, there is an increasing demand for effective facial recognition technology, and the adoption of frontalization techniques can substantially enhance the precision and robustness of these systems. The idea of achieving efficient frontalization of faces has evolved into reality with the advancement of deep learning techniques such as GANs and CNNs. The major points presented in this introduction are summarised in the following table:

Topic	Description
Importance of facial recognition tech	Discussing the significance of facial recognition technology in modern times
Need for frontalization of faces	Highlighting the need to transform non-frontal faces to frontal views
Techniques used for frontalization	Outlining the different methods used for achieving frontalization
Deep learning-based techniques	Discussing GANs and CNNs as promising techniques for efficient frontalization
Real-world example	Illustrating the potential benefits of frontalization in the Boston bombings
Conclusion	Summarizing the need and potential benefits of frontalization techniques

In the parts that follow, we will look deeper into the approaches utilised for face frontalization, their strengths and drawbacks, and their applicability in many sectors.

2. IMPORTANCE OF FACE RECOGNITION

Facial recognition technology has evolved as a rapidly evolving discipline with numerous uses ranging from security to education. Based on biometric identification principles, the technology analyses an individual's physical or behavioural features using image processing, pattern recognition, and machine learning algorithms. The application of facial recognition in security systems, law enforcement, marketing, healthcare, and education has changed the landscape of these industries, delivering efficient, accurate, and safe answers to longstanding challenges. It aids in the prevention of criminal conduct in security systems by identifying individuals of interest and regulating access to secure sites. It aids in the identification of suspects, missing persons, and victims of crime in law enforcement, resulting in effective case resolutions and lower crime rates in various cities. In marketing and advertising, facial recognition is being used to personalise the buying experience, track consumer behaviour, and target advertisements to certain demographics. [13] [9] [12] [24] Its growing usage in healthcare for patient identification, access control, and monitoring is especially beneficial in cases when patients may be unable to identify themselves. Facial recognition is being used in education to take attendance, monitor student behaviour, and improve campus security. [24] Nonetheless, the growth of face recognition technology has raised ethical, legal, and societal concerns, notably over problems of accuracy, prejudice, and privacy. The research community is currently investigating these difficulties and working to build more accurate and impartial facial recognition systems. Facial recognition technology is based on the principles of biometric identification, which involves the analysis of physical or behavioural characteristics of an individual's face. The technology employs a combination of hardware and software components to capture, analyse, and match facial features.

A. Face recognition methods

Facial recognition methods can be broadly classified into three categories: local, holistic (subspace), and hybrid approaches.

- Local approaches focus on analysing specific features or regions of the face, such as the eyes, nose, and mouth, and using these features to identify individuals. Local approaches are often used in conjunction with machine learning algorithms to recognize patterns in the facial data.
- Holistic (subspace) approaches, on the other hand, analyse the entire face as a single entity and use techniques such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) to extract features that are relevant for recognition. These approaches are based on the idea that faces can be represented in a lower-dimensional subspace, and that variations in facial appearance can be captured by a small number of basis vectors.
- Hybrid approaches combine both local and holistic techniques to improve the accuracy of facial recognition. For example, a hybrid approach might use local techniques to detect specific facial features, and then use subspace techniques to analyse the overall structure of the face.

In recent years, deep learning approaches such as convolutional neural networks (CNNs) have also been used for facial recognition, achieving state-of-the-art performance on several benchmarks. These approaches are capable of learning.

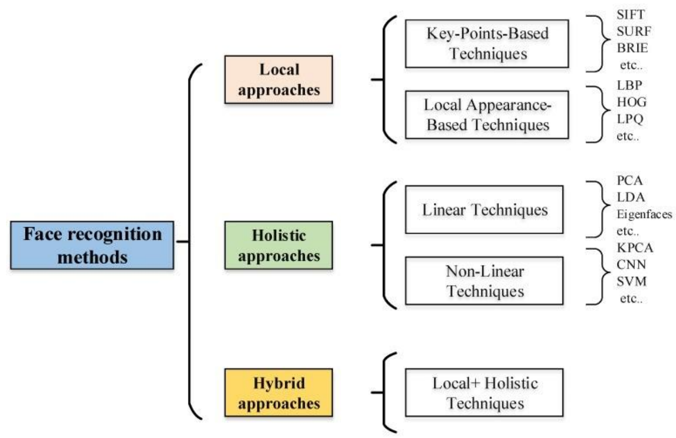

Fig. 1: Face recognition methods

complex, hierarchical representations of facial features, and can be trained end-to-end to optimize recognition performance.

B. Techniques involved.

The hardware components of a facial recognition system typically include cameras, which capture the images of faces, and sensors, which measure the depth, texture, and temperature of the face. The software components consist of algorithms that extract facial features from the images captured by the cameras, and then match these features against a database of known faces. [5] [8]

The algorithms used in facial recognition technology employ a variety of techniques, including image processing, pattern recognition, and machine learning. Image processing techniques involve enhancing the quality of the facial image by correcting for factors such as lighting, distortion, and pose. Pattern recognition techniques identify key features of the face, such as the position and size of the eyes, nose, and mouth, and use these features to create a unique facial signature. Machine learning techniques are used to train the algorithms to recognize patterns in the facial data and to improve the accuracy of the recognition over time. [25] [11] [18] [3] [12]

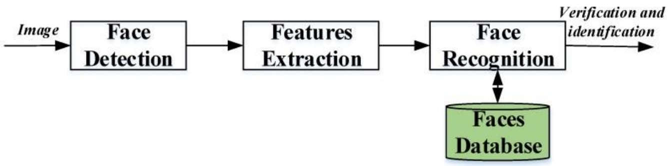

Fig. 2: Flow of Face Recognition

C. Challenges in face recognition

One of the major challenges facing facial recognition technology is pose variation. [22] [21] Changes in head pose, lighting conditions, and other environmental factors can significantly affect the performance of facial recognition systems, making it difficult to accurately identify individuals across different poses. Pose variation can also introduce significant variations in the appearance of facial features, making it challenging for recognition algorithms to extract relevant information from the images. Researchers are exploring a variety of techniques to address this challenge, including 3D modelling of facial features, pose normalization, and deep learning approaches that are capable of learning more robust representations of facial features. Despite these efforts, pose variation remains a significant obstacle for facial recognition technology, and further research is needed to develop more accurate and reliable recognition systems that are robust to changes in pose and other environmental factors. [22]

3. NEED FOR FRONTALIZATION OF FACES

The evolution of facial recognition technology from constrained to unconstrained images has resulted in a significant drop in recognition rates due to the many new challenges that unconstrained images pose. These challenges include changing expressions, occlusions, varying lighting, and non-frontal poses. [16] [1] []

Facial pose variations, in particular, have been a major challenge in unconstrained face recognition as they can significantly affect the accuracy and reliability of recognition algorithms. Researchers have designed representations that pool information over large image regions to account for possible misalignments due to pose changes. They have also improved 2D face alignment accuracy and used massive face collections to learn pose-robust representations.

Despite these efforts, unconstrained face recognition remains a challenging problem due to the significant variations in poses that can occur in real-world scenarios. To address this challenge, researchers have proposed the technique of" frontalization" to simplify unconstrained face recognition by reducing it, at least in terms of pose variations, to the simpler, constrained settings. [1] [16]

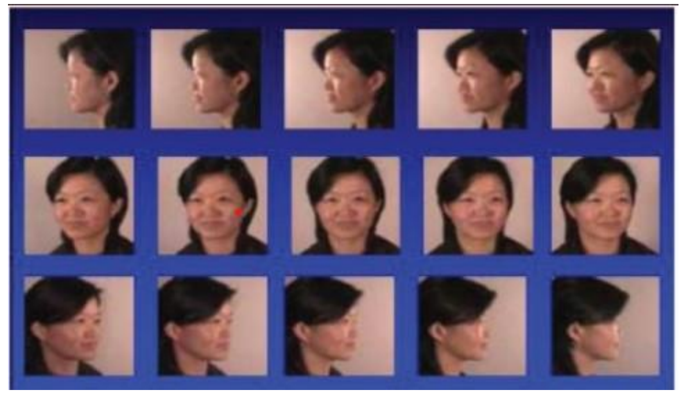

Fig. 3. Challenge of Pose variation [17]

Frontalization is a process that involves using facial landmarks to determine transforms between a facial image and a template, and it has become a popular technique for improving the accuracy of facial recognition algorithms. In the past decade, a variety of landmarking techniques have been developed, many of which rely on handcrafted features. However, with the advent of deep learning, researchers have been able to use this technology for landmark training and regression, resulting in significant improvements in landmarking accuracy. Current algorithms provide landmark sets of varying sizes, typically ranging from 7 to 194 points. To improve comparative analysis between algorithms and across different datasets, land markers have recently begun to conform to a 68-point.

standard. [1] [5] [8]

Frontalization involves automatic synthesis of new, frontal facing views of faces, which can help to mitigate the impact of variations in pose and lighting conditions and improve the accuracy and reliability of facial recognition algorithms. Although there are still some challenges to be overcome, the gradual improvement of recognition performance has led to claims being made for super-human face recognition capabilities. [16]

4. WAYS OF FRONTALIZATION

A. 3D Face model

One approach to address this challenge is through the use of 3D face models, which provide a more complete representation of the face than 2D images. These models can be constructed from a set of 2D images captured from different viewpoints using techniques such as photometric stereo or multi-view stereo. Once a 3D model is constructed, it can be used to synthesize new views of the face at different poses, which can be used to train and test recognition algorithms. [10] [18] [2] There are several types of 3D face models that can be used for frontalization and recognition tasks. One common approach is to use a generic 3D model, which is constructed from a large dataset of faces and is not specific to any individual. This type of model can be used for tasks such as pose estimation and expression recognition but may not be as accurate for tasks such as identity recognition.

Another approach is to use a personalized 3D model, which is constructed from multiple images of a specific individual. This type of model can provide a more accurate representation of the individual's face but requires more data and processing time to construct.

Overall, the use of 3D face models has shown promising results in addressing the challenge of pose variation in facial recognition. However, there are still limitations and challenges in constructing accurate and efficient models, as well as in scaling these models to large datasets.

1) 3D Morphable Model: he 3D Morphable Model (3DMM) is one of the most widely used subspace-based methods, which represents faces as a linear combination of shape and texture basis functions learned from a set of 3D face scans. [2]

To perform face frontalization using the 3DMM-based approach, the 3DMM is first fitted to the input face image to estimate its 3D shape and texture. Then, the estimated 3D shape and texture are utilized to synthesize a frontal view of the face by rendering the estimated 3D model from a desired viewpoint. The rendered frontal view is then utilized as the output of the frontalization process. This method has demonstrated the ability to produce high-quality frontalized faces with realistic textures and shapes.

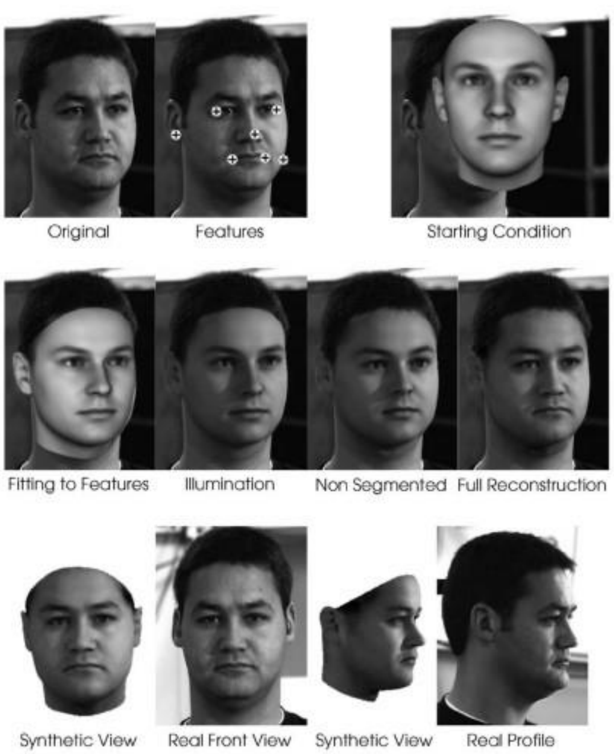

Fig. 4: 3D Morphable Model

Despite its effectiveness, the 3DMM-based methods are limited by the requirement of 3D face scans for training, which limits their practicality in real-world scenarios. Furthermore, the 3DMM-based methods can be computationally expensive and require sophisticated optimization procedures to fit the 3D model to the input face image. Recent research has focused on developing more efficient and scalable subspace-based methods for face frontalization, such as the Subspace Aggregation and Representation Learning (SARL) approach, which utilizes a hierarchical subspace representation to efficiently capture the shape and texture variations of face images and synthesize frontal views.

B. 2D face models

Another approach to frontalization is to use 2D warping techniques to align non-frontal face images with a frontal template. This method involves identifying a set of fiducial points, or landmarks, on the non-frontal face image and then warping the image to match the corresponding points on the frontal template. This technique is computationally efficient and does not require 3D models of the face, but it can result in lower-quality frontal images compared to 3D-based methods. [16]

The 2D warping method involves identifying landmarks on the non-frontal face image, such as the eyes, nose, and mouth, and then mapping these points onto a corresponding frontal template. The non-frontal face image is then warped to match the corresponding landmarks on the frontal template, resulting in a frontal view. This approach has been widely used in face recognition and detection systems, and various landmark detection techniques have been developed to improve the accuracy of the warping process.

While 2D warping techniques are computationally efficient, they have some limitations. One limitation is that the accuracy of the warping process depends on the quality of the landmark detection, which can be affected by factors such as pose variation and occlusion. Additionally, the warping process can result in distortion and loss of information in the image, which can affect the quality of the frontalized image.

To address these limitations, researchers have proposed hybrid methods that combine 2D warping with 3D-based approaches, such as the 3DMM-based method discussed earlier. These hybrid methods aim to leverage the advantages of both approaches to produce high-quality frontalized images. For instance, a recent study proposed a hybrid method that uses 2D warping to initialize the 3DMM-based frontalization process, resulting in more accurate and efficient frontalization compared to using 3DMM alone.

Based on the research paper [16], The proposed method for face frontalization takes an input profile view face, and database, D, containing a wide range of profile-frontal pose pairs for different individuals. The method involves two steps to frontalize the face. First, facial landmark detection is performed on the input face using a state-of-the-art method [19]. Then, the most similarly posed face and its corresponding frontal view face are retrieved from the exemplar database D. By matching the profile views of two faces, there is a high likelihood that the two individuals have similar facial structures. This property is exploited to obtain the geometrical transformations required for frontalization of the input face.

The frontal view of image is obtained using the affine transformations between image and the retrieved frontal view face, If. Unlike a recently proposed state-of-the-art method that uses a generic 3D model for computing this transformation, the proposed method preserves the discriminative structural information unique to an individual face by finding the nearest profile exemplar and its corresponding frontal view face. Therefore, the proposed method provides a more individualized approach to face frontalization.

This method has shown promising results in preserving the unique structural information of an individual face, which is crucial for face recognition and identification tasks. However, it has limitations in cases where the exemplar database does not contain a suitable match for the input face, and in scenarios where the input face has extreme pose variations or occlusions. Further research is needed to improve the robustness and scalability of this method for real-world applications.

C. Deep Learning Techniques

The third approach to frontalization employs the use of deep learning techniques that have shown great potential in generating realistic frontal views of the face. This technique involves training a neural network on a large dataset of face images with corresponding frontal and non-frontal views, and then using the trained network to generate frontal views from non-frontal input images. The deep learning approach has the advantage of being highly flexible and capable of handling a wide range of pose and expression variations, making it an attractive solution for practical applications such as face recognition and analysis.

Recent research in deep learning-based frontalization methods has demonstrated promising results, achieving high accuracy in generating realistic frontal views of the face even under challenging conditions such as large pose and illumination variations. For instance, the method proposed by [27] uses a multi-scale convolutional neural network (CNN) to estimate the pose-invariant facial geometry and synthesize the corresponding frontal view of the face. Similarly, the method proposed by [6] [7] employs a generative adversarial network (GAN) to synthesize frontal views of the face by minimizing the difference between the generated frontal view and the ground truth frontal view.

Despite the advantages of deep learning-based frontalization methods, they require a large amount of training data and can be computationally expensive. Furthermore, the generated frontal views may lack certain structural details and result in artifacts, highlighting the need for continued research in this area. Overall, deep learning-based frontalization methods have shown great potential for generating realistic frontal views of the face, but further improvements are needed to make them more practical for real-world applications.

The deep learning which can be used for the task at hand are as follows.

1) Generative Adversarial Networks (GANs): GANs have been successfully used to generate realistic frontal face images from non-frontal images by minimizing the difference between the generated frontal view and the ground truth frontal view.
2) Autoencoders: Autoencoders can be used for front face generation by learning a compressed representation of a non-frontal image and then reconstructing it as a frontal image.
3) Variational Autoencoders (VAEs): VAEs are similar to autoencoders but add a constraint on the latent space, resulting in more structured and continuous representations. VAEs have been used for front face generation by learning the distribution of non-frontal images in the latent space and generating new frontal views.
4) Convolutional Neural Networks (CNNs): CNNs can be used for front face generation by learning a mapping from non-frontal to frontal images through multiple convolutional layers. These networks can be trained on a large dataset of frontal and non-frontal images.
5) Deep Convolutional Generative Adversarial Networks (DCGANs): DCGANs are a type of GAN that use convolutional neural networks in both the generator and discriminator networks. DCGANs have been used for front face generation by learning a mapping from non-frontal to frontal images through multiple convolutional layers.

5. USE OF GANS

Generative Adversarial Networks (GANs) have emerged as a powerful deep learning technique for synthesizing high quality and realistic images. GANs consist of two deep neural networks, a generator and a discriminator, which are trained together in an adversarial manner. The generator learns to create images that are similar to the real images in a given dataset, while the discriminator learns to distinguish between real and generated images. The training process of GANs results in the generator producing images that are increasingly similar to the real images, while the discriminator becomes increasingly better at distinguishing between real and generated images.

In the context of frontal face generation, GANs have been used to synthesize frontal views of faces from non-frontal views or profile views. The generator network is trained to produce a frontal view of a face given a non-frontal view as input, while the discriminator network is trained to distinguish between real frontal views and generated frontal views. [14], [20], [26]

The use of GANs for frontal face generation has shown promising results and has been applied in various applications such as face recognition, emotion recognition, and facial expression analysis. However, the training process of GANs can be computationally expensive and requires a large amount of training data. Moreover, the generated images may suffer from artifacts and distortions, and the quality of the generated images can be highly dependent on the quality and diversity of the training data.

Several variations and improvements of GANs have been proposed in recent years, such as conditional GANs, CycleGANs, and StyleGANs, [14], [20], [26] that address some of these limitations and challenges. These techniques allow for more control and flexibility in the generation of images and have shown to produce highly realistic and diverse images.

Several recent studies have demonstrated the effectiveness of GANs for frontal face generation. For example, a study [23] used a GAN to generate realistic frontal views of faces from profile views. The study showed that the GAN was able to generate high-quality frontal images that were comparable to those generated by other state-of-the-art methods.

Another study used a GAN to improve the features of the frontal views of faces. The study showed that the GAN was able to generate highly realistic images with a wide range of features, including those that were not present in the training data. [20]

Compared to other methods for frontal face generation, GANs have several advantages. First, they are highly flexible and can generate images with a wide range of pose and expression variations. Second, GANs are able to learn the underlying distribution of the training data, which allows them to generate images that are both realistic and diverse. Finally, GANs can be trained with relatively small amounts of data, which makes them attractive for tasks where large datasets are not available.

6. EXAMPLES OF GANS USED

1) Profile To Frontal Coupled GAN

PF-cpGAN is a novel framework designed for profile to frontal face recognition. The main goal of this framework is to match profile face images with a gallery of frontal face images that were not seen during training, in a common embedding subspace. The framework is comprised of two modules - a profile GAN module and a frontal GAN module. Both modules use a generator and a discriminator, as well as a perceptual network based on VGG-16. The generators are designed using a U-Net auto-encoder architecture that is able to extract global features and generate images by leveraging this overall information, which is useful for global shape transformation tasks such as profile to frontal image conversion. Patch-based discriminators are used, which are trained iteratively along with the respective generators. A contrastive loss function is used to find the hidden relationship between the profile face features and frontal face features in a latent common embedding subspace. This framework was evaluated on several standard datasets, including Celebrities in Frontal-Profile, Labelled Faces in the Wild (LFW), and the Chinese Face Patch (CFP) dataset. The results demonstrated that PFcpGAN outperformed other state-of-the-art algorithms for profile to frontal face verification. For instance, under extreme pose of ±90◦, PF-cpGAN achieved improvements of approximately 11% when compared to state-of-the-art methods for the CMU-MultiPIE dataset. Additionally, the performance of PF-cpGAN was compared to that of two other similar implementations, coupled CNN (cpCNN) and domain adaptation network (ADDA), and was shown to perform much better than these two implementations. The frontal image reconstruction performance of PF-cpGAN was also evaluated. Finally, an ablation study investigated the improvement achieved by different losses, including perceptual and GAN losses, in the proposed algorithm.

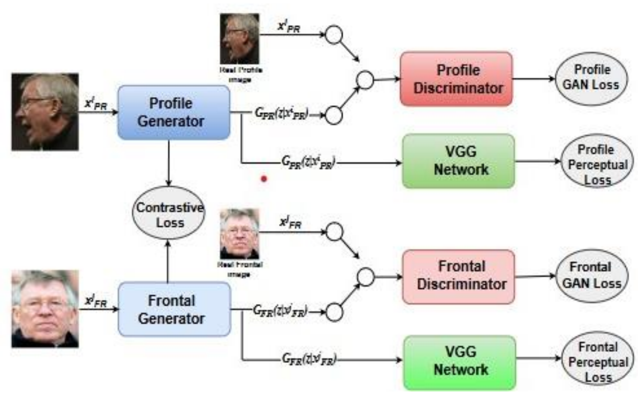

Fig. 5. Caption

2) Two Pathway GAN The Two Pathway Generative Adversarial Network (TP-GAN) proposed in this study is designed for frontal view synthesis from a single image. Unlike previous methods that use a single network to model the synthesis function, TP-GAN's generator (G_{θ_G}) has two pathways, with one global network (G_{θ_g}) processing the global structure and four landmark located patch networks (G_{θ_i}), $i \in 0,1,2,3$ attending to local textures around four facial landmarks. This two-pathway structure is based on the popular routine for 2D/3D local texture warping methods that divide the normalization of faces into two steps. The first step aligns the face globally with a 2D or 3D model, and the second step warps or renders local textures to the global structure.

To handle the highly non-linear transformation of synthesizing a frontal face from a profile image, TPGAN employs a global parametric model combined with a local non-parametric model. TP-GAN argues that using only a global network cannot learn filters that are suitable for both rotating a face and precisely recovering local details. The success of the two-pathway structure in traditional methods is transferred to a deep learning-based framework, resulting in the humanlike two-pathway generator for frontal view synthesis. The proposed (G_{θ_g}) is composed of a down sampling Encoder ($G_{\theta_g^E}$) and an up-sampling Decoder ($G_{\theta_g^D}$), and extra skip layers are introduced for multi-scale feature fusion. The bottleneck layer in the middle outputs a 256dimensional feature vector vid, which is used for identity classification to allow for identity-preserving synthesis. TP-GAN also concatenates a 100-dim Gaussian random noise to vid to model variations other than pose and identity.

To make the ill-posed synthesis problem well constrained, TP-GAN introduces adversarial loss, symmetry loss, and identity-preserving loss in the training process. Adversarial loss guides the synthesis to reside in the data distribution of frontal faces. Symmetry loss explicitly exploits the symmetry prior to ease the effect of self-occlusion of large pose cases. Identity-preserving loss ensures that the generated synthesis results are not only visually appealing but also readily applicable to accurate face recognition.

Experimental results demonstrate that TP-GAN not only presents compelling perceptual results but also outperforms state-of-the-art results on large pose face recognition. The global and local perception GAN framework, containing two separate pathways, models the out of plane rotation of the global structure and non-linear transformation of the local texture, respectively, and provides a promising approach for frontal view synthesis from a single image.

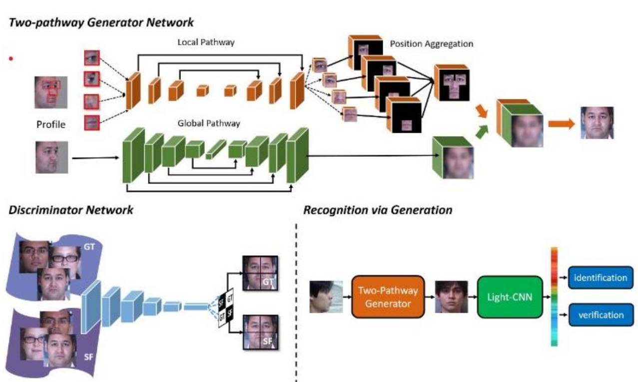

Fig. 6: Caption

7. COMMONLY USED DATASETS

The task of front face generation from the side view profiles is a challenging problem in the field of computer vision and image processing. In order to address this problem, researchers have utilized various datasets with diverse characteristics and specifications.

1) One of the most widely used datasets in this domain is the multi-PIE dataset, which was introduced by Carnegie Mellon University. It contains over 750,000 images of 337 subjects captured under various illumination conditions, facial expressions, and poses. The MultiPIE dataset is frequently utilized for face recognition and synthesis tasks due to its high level of variability and complexity. [4], [15], [20], [23]
2) Another notable dataset is the 300W-LP dataset, which is a large-scale face alignment dataset containing 122,450 face images with significant variations in pose, expression, and occlusion. The dataset includes both 2D and 3D annotations, enabling researchers to use it for face alignment and face reconstruction tasks with increased accuracy.
3) The CFP dataset is a relatively new and challenging dataset that contains over 1,000 frontal-profile pairs of celebrity faces with significant variations in pose, expression, and lighting conditions. The dataset is designed to evaluate the performance of algorithms in generating frontal faces from side view profiles. Each frontal-profile pair in the dataset is manually aligned and normalized, allowing researchers to focus solely on the generation task without the need for additional preprocessing. [15], [20], [23]
4) The LFW dataset is a widely used benchmark dataset for face recognition tasks, containing over 13,000 face images of 5,000 subjects captured under various lighting conditions, poses, and expressions. The dataset includes a subset of 1,680 face pairs labelled as" same" or" different", which are commonly used for evaluating face verification models. However, the dataset's variability and complexity make it a challenging benchmark for the frontalization of side view profile faces. [4], [13], [20], [25]

Finally, the CelebA dataset is a large-scale face attributes dataset that contains over 200,000 images of celebrities annotated with facial attributes such as gender, age, and hair colour. The dataset is commonly used for face attribute recognition and synthesis tasks due to its high level of variability and detailed annotations.

Researchers often employ combinations of these datasets to train their models, augmenting them with additional images or creating custom datasets using specialized equipment and protocols. The selection of dataset(s) depends on the specific research problem and objectives of the study, as well as the complexity and diversity required for the task at hand.

8. EVALUATION METRICS

The task of generating front-facing faces from side-view profiles is a challenging problem in the field of image generation. There are several evaluation metrics used to measure the quality of generated images, including Fréchet Inception´ Distance (FID) and Perceptual Path Length (PPL).

FID is a commonly used evaluation metric that measures the distance between the distributions of the real and generated images in a feature space obtained from a pre-trained Inception network. Lower FID values indicate that the generated images are closer to the real images in terms of visual quality and diversity. However, FID has been criticized for not always correlating well with human perception.

PPL is another evaluation metric that measures the perceptual quality of generated images. It calculates the average distance between the representations of consecutive images in a perceptual space and aims to encourage smooth, continuous changes between —images. Lower PPL values indicate that the generated images are more perceptually coherent and visually plausible.

For the specific task of generating front-facing faces from side-view profiles, additional metrics can be used, such as the degree of facial symmetry and realism of the generated faces. These metrics can be subjective and may require human evaluation.

It is important to note that evaluation metrics should be used in conjunction with human Below is the table summarizing some of the evaluation metrics used for measuring the quality of generated images, including those specifically used for generating front-facing faces from side-view profiles:

Metric	Description	Advantage	Limitations
Fréchet Inception Distance (FID)	Measures the distance between the distributions of real and generated images in a feature space obtained from a pre-trained Inception network	Computationally efficient, widely used, correlates well with visual quality	May not always correlate well with human perception
Perceptual Path Length (PPL)	Measures the average distance between the representations of consecutive images in a perceptual space to encourage smooth, continuous changes between images	Encourages perceptual coherence and visual plausibility	Can be computationally expensive
Symmetry Score	Measures the degree of facial symmetry in generated faces compared to real faces	Provides insight into the realism of generated faces	May not fully capture the visual quality of generated faces
Human Evaluation	Involves having humans rate the quality of generated images	Provides a comprehensive understanding of visual quality	Subjective, time-consuming, expensive

It is important to note that these metrics are not exhaustive and that there may be other evaluation metrics used in the field of image generation. Additionally, the selection of evaluation metrics should be tailored to the specific task at hand and should be used in conjunction with human evaluation for a comprehensive understanding of visual quality.

REFERENCES

[1] Sandipan Banerjee, Joel Brogan, Janez Krizaj, Aparna Bharati, Brandon Richard Webster, Vitomir Struc, Patrick J. Flynn, and Walter J. Scheirer. To frontalize or not to frontalize: Do we really need elaborate pre-processing to improve face recognition? In *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 20–29. IEEE.

[2] V. Blanz and T. Vetter. Face recognition based on fitting a 3d morphable model. 25(9):1063–1074.

[3] Zhang Dehai, Ding Da, Li Jin, and Liu Qing. A PCA-based face recognition method by applying fast fourier transform in pre-processing:

[4] Geonmo Gu, Seong Tae Kim, Kihyun Kim, Wissam J Baddar, and Yong Man Ro. Differential generative adversarial networks: Synthesizing non-linear facial variations with limited number of training data.

[5] Peiyun Hu and Deva Ramanan. Finding tiny faces.

[6] Rui Huang, Shu Zhang, Tianyu Li, and Ran He. Beyond face rotation: Global and local perception GAN for photorealistic and identity preserving frontal view synthesis. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2458–2467. IEEE. correlation research: a tutorial. 9(1):1.

[7] Rui Huang, Shu Zhang, Tianyu Li, and Ran He. Beyond face rotation: Global and local perception gan for photorealistic and identity preserving frontal view synthesis, 2017.

[8] Yu Yin, Songyao Jiang, Joseph P. Robinson, and Yun Fu. Dual-attention GAN for large-pose face frontalization. In *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*, pages 249–256. IEEE.

[9] Huaizu Jiang and Erik Learned-Miller. Face detection with the faster r-CNN.

[10] Maher Jridi, Thibault Napoleon, and Ayman Alfalou. One lens optical´ correlation: application to face recognition. 57(9):2087.

[11] Zhiqi Kang, Mostafa Sadeghi, and Radu Horaud. Face frontalization based on robustly fitting a deformable shape model to 3d landmarks.

[12] Yassin Kortli, Maher Jridi, Ayman Al Falou, and Mohamed Atri. A comparative study of CFs, LBP, HOG, SIFT, SURF, and BRIEF for security and face recognition. In Ayman Al Falou, editor, *Advanced Secure Optical Image Processing for Communications*, pages 13–1–13– 22. IOP Publishing.

[13] Yassine Kortli, Maher Jridi, Ahmed A Falou, and Mohammed Atri. Face recognition systems: A survey. Sensors, 20(2):342, 2020.

[14] Shengcai Liao, Anil K. Jain, and Stan Z. Li. Partial face recognition: Alignment-free approach. 35(5):1193–1205.

[15] Jinjin Liu, Qingbao Li, Ming Liu, and Tongxin Wei. CP-GAN: A crosspose profile face frontalization boosting pose-invariant face recognition. 8:198659–198667.

[16] Xiao Luan, Hongmin Geng, Linghui Liu, Weisheng Li, Yuanyuan Zhao, and Min Ren. Geometry structure preserving based GAN for multi-pose face frontalization and recognition. 8:104676–104687.

[17] Mallikarjun B R, Visesh Chari, and C.V. Jawahar. Efficient face frontalization in unconstrained images. In 2015 Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), pages 1–4. IEEE.

[18] Rohan Naik and Kalpesh Lad. Face recognition from multi angled images. 4(2).

[19] Thibault Napoleon and Ayman Alfalou. Pose invariant face recognition: ´ 3d model from single photo. 89:150–161.

[20] Avhi Poddar, Surabhi Gawade, Prasad Varpe, and Sumedha Bhagwat. Frontal face landmark generation using GAN. In 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC), pages 1172–1177. IEEE.

[21] Changle Rong, Xingming Zhang, and Yubei Lin. Feature-improving generative adversarial network for face frontalization. 8:68842–68851.

[22] Shilpi Singh and S.V.A.V. Prasad. Techniques and challenges of face recognition: A critical review. 143:536–543.

[23] Merrin Mary Solomon, Mahendra Singh Meena, and Jagandeep Kaur. CHALLENGES IN FACE RECOGNITION SYSTEMS. 6(2).

[24] Fariborz Taherkhani, Veeru Talreja, Jeremy Dawson, Matthew C. Valenti, and Nasser M. Nasrabadi. Profile to frontal face recognition in the wild using coupled conditional GAN.

[25] Q. Wang, A. Alfalou, and C. Brosseau. New perspectives in face

[26] Meng Xi, Liang Chen, Desanka Polajnar, and Weiyang Tong. Local binary pattern network: A deep learning approach for face recognition. In *2016 IEEE International Conference on Image Processing (ICIP)*, pages 3224–3228. IEEE.

[27] Xiangyu Zhu, Zhen Lei, Xiaoming Liu, Hailin Shi, and Stan Z. Li. Face alignment across large poses: A 3d solution. 41(1):78–92.