International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

An Extensive Survey on Sketch to Photo Synthesis Methods: Trends and Challenges

[1]Jyoti S. Raghatwan, [2]Dr. Sandhya Arora

Submitted: 26/11/2023 **Revised**: 06/01/2024 **Accepted**: 16/01/2024

Abstract: Nowadays Face sketch- photo synthesis has drawn the attention of many researchers. Since the Photo to sketch synthesis or sketch to photo synthesis is used in many different application, detailed review is very much important to analyse state-of-the-art approaches. With this in mind, we offer a thorough analysis of the existing deep learning-based and traditional approaches, which fall into the categories of data-driven and model-driven approaches, in this study. A comparative study of the evaluated methods is conducted by considering several factors like the performance measurements, algorithms, and dataset.

Keywords: *Deep learning, Face sketch, Model driven, Data-driven.*

1. Introduction:

Face sketch to photo synthesis is method to generate photo image from input sketch image. Nowadays many of the researcher proposed different sketch to photo synthesis methods due to their applicability in multiple domains like digital entertainment and low of enforcement [1][2][3]. Because of domain difference between input image (grey scale image) and output image (RGB Coloured image) it is very challenging task.When it comes to law enforcement, eyewitness descriptions sometimes be the sole piece of information that helps identify suspects since photos of them are not always readily available. Unfortunately, the geometric and textural variations between the sketch and the photos makes it challenging to identify the suspect's face directly [4][5]. Face sketch–photo synthesis methods help to bring the face sketches into the same domain[6][7]. Additionally, mobile apps like Facebook, digital entertainment [8], and facial animation [9] have made use of the face sketch–photo synthesis.

As face sketch synthesis is receiving a lot of attention, it is necessary to study existing methods as shown in Fugure 1.. This work focuses on the comprehensive survey and analysis of existing sketch to photo synthesis methods and their challenges.

The existing sketch to photo synthesis methods are divided into two main categories i.e. data-driven methods and model-driven methods. Data driven methods are also known as, exemplar-based methods, Furthermore data - driven method sub- classified into four categories:1. Bayesian inference (BI) based 2. subspace learning (SL) based 3. Combination of BI and SL 4. sparse representation-based.

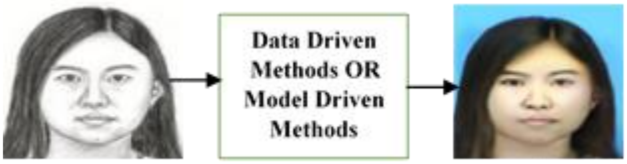

Fig 1 Face Sketch to Photo Synthesis

Mainly similar photo sketch patches from the training set are linearly combined in data-driven approaches. Therefore, patch assembly, weight calculation, neighbor selection, and patch representation are required. The time required for the test is significantly increased due to the similar patch searching process [10][11][12].

Model-driven approach learn the mapping relationship or mathematical function offline from training photo and sketch. Then, by iteratively going through training phase, this learnt mapping is applied during the test phase to convert the test sketch into photo or vice versa. Here for traditional model driven approach Researchers have put a lot of work into investigating hand-crafted features, neighbor finding tactics, and learning procedures [13][14][15]. Many researchers have recently developed deep learning based method for sketch to photo synthesis task. Different CNN architectures and after excellent performance of Generative Adversarial network (GAN) in image-to-image translation task, different GAN architectures are proposed for Sketch to photo synthesis. And the research is still ongoing. Figure 2. represents the taxonomy of sketch to photo synthesis methods.

[1]*Research Scholar, SKN College of Engineering,Pune*
Asst.Prof., RMD Sinhgad School of Engineering
jyotiraghatwan2@gmail.com
[2]*Professor, Cummins college of Engineering, Pune*
sandhya.arora@cumminscollege.in

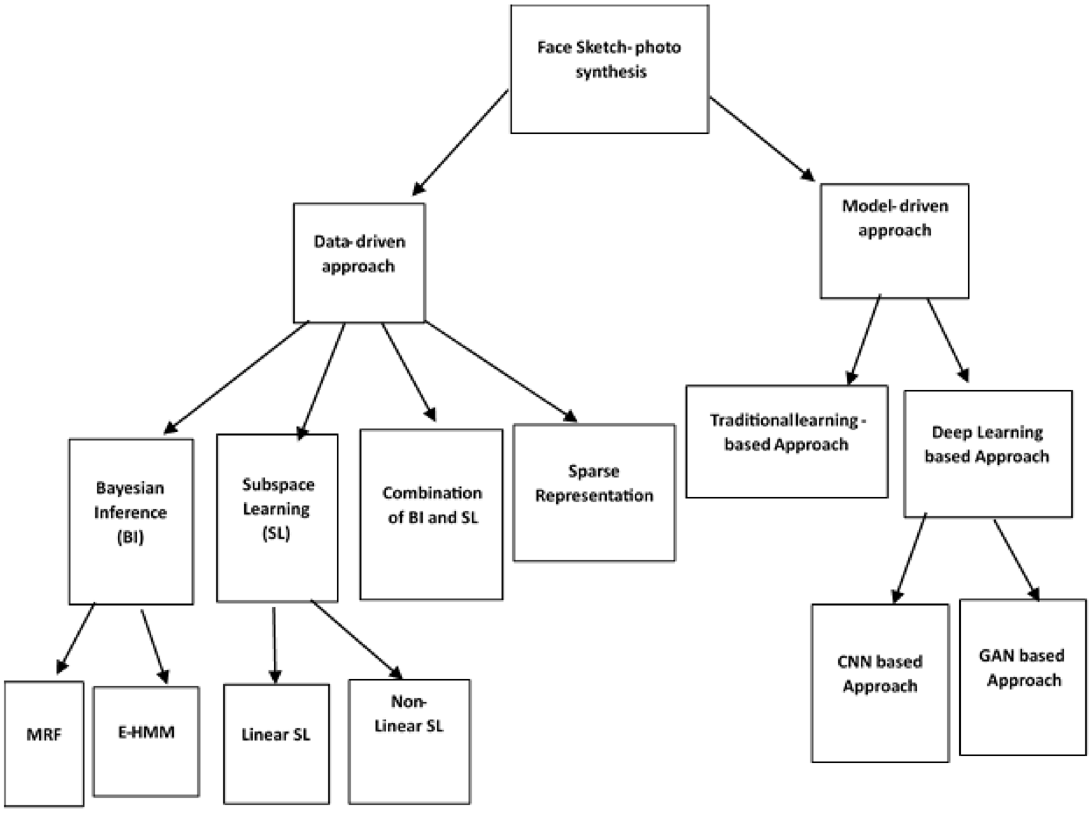

Fig 2. Taxonomy of Sketch to photo synthesis method

This paper is organized as follows in its remaining sections. Data-driven techniques are covered in Section 2. Model-driven techniques are covered in Section 3. We go over some datasets and their features in section 4. Section 5 explains the various evaluation parameters that are employed. Section 6 talks with different challenges. Section 7 concludes and provides guidance for the future.

2. Data-Driven Approach

Typically, data-driven techniques operate at the patch level. Neighbor Selection Model and Neighbor Fusion Model which is also called Weight Computation Model are the two components of these methods. Here we discussed different data driven approach.

A. Bayesian Inference (BI)

Bayesian Inference Framework is based on Bayes' theorem which is given as s $P(A|B) = P(B|A)P(A) /P(B)$, where A and B represent two events in the event space[16]. Bayesian inference-based approach further classified into MRF-based methods the embedded hidden Markov model (EHMM)-based methods. X. Wang et. al. [17] proposed multiscale Markov Random Fields (MRF) model for sketch to photo synthesis and the recognition. In this method for learning, the face region is split into overlapping patches, a compatibility function provided the constraints between the target sketch patch and its surrounding patches, and dependency between the target sketch patch and the input test photo patch is provided by a local evidence function. The scale of local facial structures to be learned is determined by the size of the patches. The faces photo-sketches to be used in this method are in frontal pose, have no occlusions, normal lighting, and a neutral expression. This Methods does not perform well when there are variations of and pose lighting. To handle such variation W. Zhang [18] proposed extended multiscale Markov Random Fields (MRF) model by introducing robust patch matching, shape priors, new compatibility terms. To generate new sketch patches which is not present in training set Zhou et al. [19] proposed a Markov weight fields (MWFs) model. Proposed model is formulated as a convex quadratic programming (QP) problem, for which there is an optimal solution known to exist and cascade decomposition technique (CDM) used to effectively solve it.

Embedded hidden Markov model and selective ensemble strategy-based algorithm proposed by X. Gao et al. [20] for photo-sketch synthesis. Nonlinear relationship between sketch and photos are model by E-HMM. A face image was decomposed into different superstates. To extract the local features in the face image, upstate was decomposed into several embedded states. To achieve the desired outcome, a sequence of synthesized sketches was created and combined utilizing the selective ensemble approach. However, one drawback of the E-HMM is that it is difficult to understand more intricate nonlinear relationships. To reduce overhead of large training data set, B. Xiao et al. [21] proposed model using EHMM, where non-relationship is model by considering face image patch and its neighbouring patches.

B. Subspace Learning (SL)

Subspace Learning approach classified into two categories: 1. Linear Subspace Learning 2. Non-linear Subspace Learning. Linear Subspace Learning methods are based on principal component analysis (PCA) [22]. X. Tang and X. Wang [6], [23] proposed face sketch synthesis method by considering it as linear process and Utilizing PCA, developed the eigen transform Method. Linear combination coefficients are achieved by projecting input photo onto the training dataset. The corresponding training sketches and the previously acquired projection coefficients were then combined linearly to create the target sketch. After that they proposed method [7] to separate out shape and texture and then apply eigen transformation on shape and texture. By combining the synthesized shape and the texture, target sketch is obtained. But linear subspace learning is not effective for entire face photos and affect recognition performance. To overcome that Liu et al. [1] proposed a nonlinear subspace learning approach which is based on local linear embedding. Instead of using the photo space, the suggested method conducted eigen analysis on a hybrid space made up of training sketches and training photos. Projecting the query sketch on the sketch space spanned by the columns of the eigen sketch matrix yields the projective coefficients, which are then generated by splitting the hybrid projection matrix produced from the eigen-analysis into two linked matrixes: an eigen photo matrix and an eigen sketch matrix. Lastly, the linear combination of eigen photos weighted by the determined coefficients is used to generate the pseudo-photo.

C. Combination of BI and SL

Some methods are proposed using combination of Bayesian Inference and Subspace Learning. W. Liu [24] proposed two step method where in first step initial estimate are generated using LLE-based methods. In second step by using proposed tensor models learn the inter-space dependencies photo patch space and the sketch patch space. The Bayesian MAP framework integrates tensor modelling and statistical optimization. Tensor modelling is combined with statistical optimization in the Bayesian MAP framework.

D. Sparse Representation

In sketch-to-photo synthesis techniques, sparse representation techniques use a small amount of data to represent the sketch, and then they create a realistic photo based on that representation. Sparse-based representation techniques are advantageous because they let the synthesis model concentrate on the important features of the sketch rather than extraneous details. It can be used for several kinds of computer vision tasks. Sparse representation-based face sketch synthesis method is proposed by L. chang et al. [25]. Utilizing Lasso [1], they solve l1- norm minimization problem to synthesise sketch image from given photo. They divide training photos and sketches into overlapped region and with same sparse coefficients face photo patch and the subsequent sketch patch decomposed on the photo patch and the sketch patch dictionary [26]. The sparse representation coefficient of each image patch in the test photo can be calculated with respect to the photo elements in the coupled dictionary. With the same coefficient and the sketch elements in a coupled dictionary, the sketch patch can be retrieved. Finally, by combining the acquired sketch patches, a target sketch can be produced. Sparse representation technique also utilized by Gao et al. [27] to proposed two -step framework for sketch-photo synthesis and improve the results. Initial estimation of pseudo-photo or pseudo-sketch achieved through sparse neighbor selection method to adaptively find closely related neighbors. Then further the quality of image improved by sparse-representation-based enhancement (SRE) techniques. Also, J. Zhang [28] suggested a two-step method that took the image's perceptual quality into account. They proposed support vector regression techniques to get high frequency information. To obtain the first estimate of the synthesized image, they apply an existing technique. Initial estimate and the SVR based high frequency information is combined to obtain final resultant image. Multi-dictionary sparse representation method proposed by N. Wang.[29] to improve the synthesized result by focusing on detailed and high frequency information. They utilized LLE based technique for initial estimation.Table 1 summarizes data driven methods.

Table 1. Summary of Data driven methods for sketch to photo synthesis methods

Data driven Methods	Algorithm used	Paper	Dataset used	Performance Metrics Used	Compared with
Bayesian Inference	MRF	[17]	CUHK Face Sketch (CUFS) dataset	Cumulative Match Score (CMS)	Eigen face, EGM
	Multiscale (MRF)	[18]	CUHK Face Sketch (CUFS) dataset	Structural Similarity Index Measure (SSIM) Feature Similarity Index Metric. (FSIM)	LLE MRF LWF Trans SSD
	Markov weight fields (MWF)	[19]	CUHK Face Sketch (CUFS) dataset	Visual Perception (VP)	MRF Multiscale MRF
	E-HMM	[20]	FERET, AR	Universal Image Quality Index (UIQI)	Non-linear methods
Subspace Learning	Linear Subspace Leaning (Eigen transform)	[23]	FERET	Cumulative Match Score (CMS)	Geometry Method Eigenface Method
		[7]	CUHK Face Sketch (CUFS) dataset	Cumulative Match Score (CMS)	EGM Eigenface Method
	Non-Linear Subspace Leaning	[1]	CUHK Face Sketch (CUFS) dataset	Visual Perception (VP)	Eigenface Method
Combination of BI and SL	LLE and Bayesian MAP	[24]	CUHK Face Sketch (CUFS) dataset	RMSE	local geometry preserving (LGP) eigen transform
Sparse Representation	Sparse Representation	[25]	CUHK Face Sketch (CUFS) dataset	Visual Perception (VP)	MRF Linear and Non-Linear Subspace Leaning
	Sparse-representation-based enhancement (SRE)	[27]	CUHK Face Sketch (CUFS) dataset VIPSL database	Mean opinion score (MOS)	Non-Linear Subspace Leaning
	Support vector regression techniques	[28]	CUHK Face Sketch (CUFS) dataset VIPSL database	VP	LLE E-HMM

3. Model-Driven Approach’

Using all nearby training photos as the test patch, data-driven techniques create each target sketch patch. These methods restrict the synthesis's efficiency. Instead, model driven approach learn mapping from sketch and photo patches.

A. Traditional learning-based Approach

Peng et al. [30] proposed multiple representations-based face sketch–photo synthesis (MrFSPS) method. Face image patch is defined by multiple representation, which is attained by the application of three filters: Gaussian smoothing, center-surround divisive normalization (CSDN), and DoG filters. The combination weights of several representations can be adaptively learned using Markov networks. S. Zhanget al. [31] propose Sparse Representation-Based Greedy Search model to synthesise photo using prior knowledge and similarity between different image patches. They initially use a learned dictionary to get the sparse coefficient for each photo patch. After that the candidate photo patches selected from the training photo patches set using the greedy search method and the sparse coding information, which contains the dimension selection order and the related sparse coefficient. Markov network model utilize for final sketch synthesis. Further they also proposed model [32] to synthesize arbitrarily stylistic sketches. N. Wang et al. [33] proposed transductive face sketch-photo synthesis method in which they built probabilistic framework to model the process of sketch-photo synthesis. Reconstruction fidelity of the input and output image optimizes using alternating optimization method.

B. Deep Learning based Approach

Deep learning approaches demonstrate their superiority in computer vision tasks. The nonlinear spatial transform between the input and output domains can be learned using these methods. A six-layered fully convolutional network was proposed by Zhang et al. [30] to generate photo sketches. People's identities are preserved in this suggested architecture through the use of a novel joint generative-discriminative formulation in the optimization of the objective function. Many cross-modality applications have effectively employed deep convolutional neural networks (DCNNs) [31][32]. A model to convert image style into random sketch images was proposed by Li et al. [33]. For texture synthesis, they use generative Markov random field (MRF) models in conjunction with deep neural networks, which are known for their discriminative power. By learning feature representation at several CNN layers, Gatys et al. [33] developed the first neural style transfer system that applies the style of a reference image to any input image.

CNNs' primary objective function in image synthesis applications has been to minimize the Euclidean distance between the pixels in the predicted image and the pixels in the ground truth image. On the other hand, a network that is trained using this objective function often produces result with blurry effect [34]. In recent times, there has been a notable advancement in image generating tasks

through the use of deep convolutional generative adversarial networks (GANs) [35], which choose a different architecture and loss function to produce better quality, realistic images. Generative adversarial network consists of two subnetwork generator and discriminator. Generator generate the images and tries to fool discriminator which differentiates between generated and the real images. .Lu Y. et al. proposed contextual GAN architecture [36] to learn joint distribution between photo and sketch. contextual loss and perceptual Loss is incorporated to improve the result. However, this method's drawback is that face synthesis from input photographs cannot be guaranteed to preserve identity, and some attributes are absent from the resulting images. This issue is address by attribute guided GAN architecture [37] using skip connection approach with reconstruction loss and adversarial loss function. Multiple photos with specific attribute from single sketch are generated by quality aware GAN architecture [38] using Hybrid discriminator approach Sketch with pixel-wise face labels as input is provided to the composition-aided generative adversarial network (CA-GAN) [39], which focuses on hard-generated components and delicate facial structures with additional compositional reconstruction loss function. GAN architecture with deep residual U-Net as the generator and Patch-GAN with residual blocks as the discriminator [40] proposed for high-fidelity face sketch-photo synthesis. Further the quality of generated image is enhanced by Encoder Guided Generative Adversarial Network (EGGAN) [41] where feature encoder utilize to direct training phase and skip connections approach in cycle-consistent GAN. Identity information is preserved through the use of feature loss and feature consistency loss function. High resolution photo images are generated by CycleGAN [42] by utilizing multi-adversarial networks and Cycle consistency loss. Feature based condition is handle using a conditional cycle GAN [43], it does not require photo-sketch pair data for training. This framework maintains the facial style even though the generated photo changes some properties. Identity-Aware Cycle GAN[44] enhances the recognition of generated faces by taking significant facial features like the nose and eyes into account. They used perceptual loss function. Table 2. summarised GAN based sketch to photo synthesis method based on different features.

Table 2. Summary of GAN based sketch to photo synthesis methods

GAN architecture used	Dataset used	Performance Metrics Used	Result Achieved	Compare with
Composition-aided generative adversarial network (CA-GAN) [39]	CUHK Face Sketch (CUFS) dataset CUFSF dataset	Fréchet inception distance	FID=30.5	cGAN CA-GAN stack-cGAN SCA-GAN
Identity-Aware CycleGAN [44]	CUHK Face Sketch (CUFS) dataset CUFSF dataset	Structural similarity index measure Feature Similarity Index Metric	SSIM=0.6495 FSIM=0.7652	cGAN CycleGAN
GAN with hybrid discriminator and a multi-stage-generator.[38]	CelebA CelebA-HQ WVU Multi-modal CUHK, IIT-D FERET datasets	Fréchet inception distance	FID-34.1	BP-GAN C-GAN CA-GAN SCA-GAN
Deep residual U-Net as generator and a Patch-GAN with residual blocks as discriminator. [40]	CUHK Face Sketch (CUFS) dataset	Fréchet inception distance	FID-60.233	Ground Truth Pix2Pix
Standard generated adversarial net-work (GAN)[35]	CelebA dataset	Peak signal-to-noise ratio structural similarity index measure	PSNR=16.3069 SSIM =0.5790	Pix2Pix
Encoder Guided Generative Adversarial Network (EGGAN) [41]	CUHK Face Sketch (CUFS) dataset CUFSF dataset AR dataset XM2VTS dataset	Structural similarity index measure	SSIM=0.6531	Pix2Pix , CycleGAN DualGAN, CSGAN GAN.
Conditional Cycle - GAN[43]	CUHK Face Sketch (CUFS) dataset CUFSF dataset	Structural similarity index measure	Accuracy =65.53%	Ground Truth
Multi-Adversarial Networks PS2-MAN [45]	CUHK Face Sketch (CUFS) dataset CUFSF dataset	Structural similarity index measure Feature Similarity Index Metric	SSIM=0.7915 FSIM=0.8062	Pix2Pix, DualGAN, CycleGAN
Contextual GAN[36]	CelebA CelebA-HQ WVU Multi-modal CUHK, IIT-D FERET datasets	Structural similarity index measure	SSIM=0.8856	Pix2Pix

4. Dataset

In this section we discuss different datasets consisting of photo-sketch pairs and utilize by different researchers for sketch-photo synthesis task

CUHK Face Sketch Database (CUFS): CUFS[17] is oldest and most widely utilized datasets by many researchers. It consists of total 606 photo-sketch pairs where 123 AR face database [],188 from the CUHK student database [] and 295 XM2VTS database , For every sample, there is an artist's sketch and matching photo. Every Photo has a neutral expression and is capture in frontal pose in regular lighting with solid backgrounds. This dataset contains only a few limited-style sketches that were made by the same artist.

CUFSF dataset: CUFSF[46] dataset also most widely used in evaluating performance of sketch-photo synthesis methods. It consists of total 1194 photo-sketch pair from FERET database. Since the photographs in the dataset change illumination, each face has low contrast with the background, and each sketch has exaggerated outlines, it is more difficult than the CUFS dataset.

VIPS dataset: VIPS [47]consist of 200 face photos gathered from Indian face databases, FRAV2D, FERET. Photo Capturing conditions are same as CUFS, but VIPSL contains five sketches for each face, made by five artists of various styles.

IITD Dataset: IIIT-D [48] contain three sketch databases, forensic sketch database, a semi-forensic sketch database, viewed sketch database. Forensic sketch database contains 190 sketches drawn by a sketch artist based on an eyewitness' description of a crime scene. Semi-forensic sketch database contains 140 sketch-photo pair where all the sketches are drawn based on memory after the artist has seen the relevant photograph. Viewed sketch database contains 238 pairs image where all sketches drawn by the professional artist based on a provided photo.

APDrawing dataset [49]: It consists of 140 high-resolution photo-sketch pair drawn by only one artist and demonstrates multiple styles.

FS2K dataset [50]: It consists of 2104 photo-sketch pairs. It is a high-quality dataset due to the many backgrounds, lighting settings, and sketch styles. Furthermore, it has other attributes such as hair condition, gender, and grin that make it apart from other datasets**.**
Summary of the different datasets is illustrated in Table 3.

Table 3. Summary of different dataset (√) represent publicly available (X) unavailable.

Dataset	Sample Images	Total Size	Train	Test	Public	Image Size	Lighting Condition	Year
CUFS		606	306	300	√	200×250	Constant	2009
CUFSF		1194	500	694	√	200 × 250	Varying	2011
VIPS		1000	100	900	√	220 × 220	Constant	2011
IITD		231	58	173	x	320 × 440	Constant	2010
APDrawing		140	70	70	√	1024 × 512	Varying	2020
FS2K		2104	1058	1046	√	250 × 250	Varying	2022

5. Evaluation Parameter

For Sketch to photo synthesis methods or vice versa, subjective or quantitative evaluation metrics can be used. mean opinion score (MOS) or visual perception can be used for subjective quality evaluation. The MOS is the average of the quality ratings that observers provide, which range from 1 to 5. Visual perception is based on the observers' perception, without using a numerical quantification. Though it is direct and accurate metric to represent an individual's perception, it is expensive and require man power. So, to overcome that automated metrics for quantitative quality assessment have been proposed. For a general quality assessment, these include the cross-correlation, peak signal-to-noise ratio (PSNR), mean square error (MSE) or root mean square error (RMSE), and structural similarity index measure (SSIM) methods. Recently proposed deep learning based model used quantitative metrics like Inception Score, Fréchet inception distance (FID), Feature Similarity Index Metric (FSIM) to evaluate performance of their model.

6. Challenges

i. Existing image-sketch databases are typically small and lack diversity due to the high cost of collecting professional sketches. It affects the performance of deep learning methods as it requires large amount of dataset.

ii. Evaluation of sketch-photo synthesis method is challenging task, Facial sketches differ significantly from RGB-based facial images, making it challenging to apply the current evaluation metrics to them.

iii. For every given shape, the sketch incorporates shape deformation since it exaggerates a few certain face traits, much like a caricature.

iv. Lack of training data, robust style transfer technique, optimization of sketch-photo synthesis process, are challenging task in the field of photo-sketch synthesis task.

7. Conclusion

In this paper we have presented every aspect of the facial sketch to photo synthesis problem in detail. After conducting a thorough survey, methodologies for converting sketches into photos were divided into two categories: model-driven and data-driven. The researchers' performance comparison of different approaches on benchmark datasets is tabulated, and the authors have made it apparent that GAN-based methods beat other state-of-the-art methods. We think that the readers will gain a thorough understanding of face sketch synthesis and identification from this comparison study.

Reference:

[1] Q. Liu, X. Tang, H. Jin, H. Lu, and S. Ma, "A nonlinear approach for face sketch synthesis and recognition," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., San Diego, CA, USA, Jun. 2005, pp. 1005–1010

[2] H. Zhou, Z. Kuang, and K.-Y. K. Wong, "Markov weight fields for face sketch synthesis," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Providence, RI, USA, Jun. 2012, pp. 1091–1097.

[3] N. Wang, D. Tao, X. Gao, X. Li, and J. Li, "Transductive face sketchphoto synthesis," IEEE Trans. Neural Netw. Learn. Syst., vol. 24, no. 9, pp. 1364–1376, Sep. 2013

[4] R. G. Uhl, Jr., N. da Vitoria Lobo, and Y. H. Kwon, "Recognizing a facial image from a police sketch," in Proc. 2nd IEEE Workshop Appl. Comput. Vis., Sarasota, FL, USA, Dec. 1994, pp. 129–137

[5] R. G. Uhl, Jr., and N. da Vitoria Lobo, "A framework for recognizing a facial image from a police sketch," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., San Francisco, CA, USA, Jun. 1996, pp. 586–593.

[6] X. Tang and X. Wang, "Face sketch recognition," IEEE Trans. Circuits Syst. Video Technol., vol. 14, no. 1, pp. 50–57, Jan. 2004

[7] X. Tang and X. Wang, "Face sketch synthesis and recognition," in Proc. 9th IEEE Int. Conf. Comput. Vis., Nice, France, Oct. 2003, pp. 687–694

[8] H. Koshimizu, M. Tominaga, T. Fujiwara, K. Murakami. On KANSEI facial image processing for computerized facial caricaturing system PICASSO. In Proceedings of IEEE International Conference on Systems, Man, and Cybernetics,IEEE, Tokyo, Japan, pp.294–299, 1999. DOI: 10.1109/ICSMC.1999.816567.

[9] J. Yu, M. Wang, and D. Tao, "Semisupervised multiview distance metric learning for cartoon synthesis," IEEE Trans. Image Process., vol. 21, no. 11, pp. 4636–4648, Nov. 2012.

[10] X. Gao, N. Wang, D. Tao, and X. Li, "Face sketchphoto synthesis and retrieval using sparse representation," IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 8, pp. 1213–1226, 2012.

[11] Y. Song, L. Bao, Q. Yang, and M. H. Yang, "Real-time exemplar-based face sketch synthesis," in European Conference on Computer Vision, 2014, pp. 800–813.

[12] Q. Pan, Y. Liang, L. Zhang, and S. Wang, "Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis," in Computer Vision and Pattern Recognition, 2012, pp. 2216– 2223.

[13] S. Zhang, X. Gao, N. Wang, J. Li, and M. Zhang, "Face sketch synthesis via sparse representation-based greedy search." IEEE Transactions on Image Processing, vol. 24, no. 8, pp. 2466–77, 2015.

[14] S. Zhang, X. Gao, N. Wang, and J. Li, "Robust face sketch style synthesis," IEEE Transactions on Image Processing, vol. 25, no. 1, p. 220, 2016.

[15] N. Wang, D. Tao, X. Gao, X. Li, and J. Li, "Transductive face sketchphoto synthesis," IEEE

Transactions on Neural Networks and Learning Systems, vol. 24, no. 9, pp. 1364–1376, 2013.
[16] Gelman, A., Carlin, H., Stern, J., & Rubin, D. (2003). Bayesian data analysis. Boca Raton: Chapman & Hall/CRC
[17] X. Wang and X. Tang, “Face photo-sketch synthesis and recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 11, pp. 1955–1967, Nov. 2009.
[18] W. Zhang, X. Wang, and X. Tang, “Lighting and pose robust face sketch synthesis,” in Proc. 11th Eur. Conf. Comput. Vis., Heraklion, Greece, Sep. 2010, pp. 420–433
[19] H. Zhou, Z. Kuang, and K.-Y. K. Wong, “Markov weight fields for face sketch synthesis,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Providence, RI, USA, Jun. 2012, pp. 1091–1097.
[20] X. Gao, J. Zhong, J. Li, and C. Tian, “Face sketch synthesis algorithm based on E-HMM and selective ensemble,” IEEE Trans. Circuits Syst. Video Technol., vol. 18, no. 4, pp. 487–496, Apr. 2008.
[21] B. Xiao, X. Gao, D. Tao, and X. Li, “A new approach for face recognition by sketches in photos,” Signal Process., vol. 89, no. 8, pp. 1576–1588, 2009
[22] I. T. Jolliffe, Principal Component Analysis. Hoboken, NJ, USA: Wiley, 2002
[23] X. Tang and X. Wang, “Face photo recognition using sketch,” in Proc. IEEE Int. Conf. Image Process., Rochester, NY, USA, Sep. 2002, pp. I-257–I-260.
[24] Liu, W., Tang, X., & Liu, J. (2007b). Bayesian tensor inference for sketch-based face photo hallucination. In Proceedings of international joint conference on artificial intelligence (pp. 2141– 2146).
[25] M. Chang, L. Zhou, Y. Han, and X. Deng, “Face sketch synthesis via sparse representation,” in Proc. 20th Int. Conf. Pattern Recognit., Istanbul, Turkey, Aug. 2010, pp. 2146–2149.
[26] H. Lee, A. Battle, R. Raina, and A. Y. Ng, “Efficient sparse coding algorithms,” in Proc. 21st Annu. Conf. Adv. Neural Inf. Process. Syst., Vancouver, BC, Canada, Dec. 2007, pp. 801–808
[27] X. Gao, N. Wang, D. Tao, and X. Li, “Face sketch–photo synthesis and retrieval using sparse representation,” IEEE Trans. Circuits Syst. Video Technol., vol. 22, no. 8, pp. 1213–1226, Aug. 2012.
[28] J. Zhang, N. Wang, X. Gao, D. Tao, and X. Li, “Face sketch-photo synthesis based on support vector regression,” in Proc. 18th IEEE Int. Conf. Image Process., Brussels, Belgium, Sep. 2011, pp. 1125–1128.
[29] N. Wang, X. Gao, D. Tao, and X. Li, “Face sketch-photo synthesis under multi-dictionary sparse representation framework,” in Proc. 6th Int. Conf. Image Graph., Hefei, China, Aug. 2011, pp. 82–87.
[30] L. Zhang, L. Lin, X.Wu, S. Ding, and L. Zhang. “End-to-end photo-sketch generation via fully convolutional representation learning.,” in *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval*, pages 627–634. ACM, 2015.
[31] R. Socher, M. Ganjoo, C. D. Manning, and A. Ng. “Zero-shot learning through cross-modal transfer,” in *Advances in neural information processing systems*, pages 935–943, 2013.
[32] M. Iranmanesh, A. Dabouei, H. Kazemi, and N. M. Nasrabadi. “Deep cross polarimetric thermal-to-visible face recognition,” in *Biometrics (ICB), 2018 International Conference on*. IEEE, 2018.
[33] C. Li and M. Wand, “Combining markov random fields and convolutional neural networks for image synthesis,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2479–2486, 2016.
[34] L. A. Gatys, A. S. Ecker, and M. Bethge, “Image style transfer using convolutional neural networks”, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2414–2423, 2016.
[35] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets”, in Advances in neural information processing systems pages 2672–2680, 2014.
[36] Lu Y., Wu S., Tai YW., Tang CK, “Image Generation from Sketch Constraint Using Contextual GAN”, in: Ferrari V., Hebert M., Sminchisescu C., Weiss Y. (eds) Computer Vision – ECCV 2018. ECCV 2018. Lecture Notes in Computer Science, vol 11220. Springer, Cham.
[37] J. Zhao, X. Xie, L. Wang, M. Cao and M. Zhang, "Generating Photographic Faces from the Sketch Guided by Attribute Using GAN," in IEEE Access, vol. 7, pp. 23844-23851, 2019
[38] Uche Osahor, Hadi Kazemi, Ali Dabouei, Nasser Nasrabadi,(2020), “Quality Guided Sketch-to-Photo Image Synthesis”, In: Computer Vision And Pattern Recognition (CVPR).
[39] Yu et al., "Toward Realistic Face Photo-Sketch Synthesis via Composition-Aided GANs," in IEEE Transactions on Cybernetics, doi: 10.1109/TCYB.2020
[40] Chao, Wentao et al. “High-Fidelity Face Sketch-To-Photo Synthesis Using Generative Adversarial Network,” 2019 IEEE International Conference on Image Processing (ICIP) (2019): 4699-4703
[41] J. Zheng, W. Song, Y. Wu, R. Xu and F. Liu, (2019), "Feature Encoder Guided Generative Adversarial Network for Face Photo-Sketch Synthesis", IEEE Access, vol. 7, pp. 154971-154985
[42] Wang, V. Sindagi and V. Patel, "High-Quality Facial Photo-Sketch Synthesis Using Multi-Adversarial Networks," 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), Xi'an, 2018, pp. 83-90.

[43] Hadi Kazemi et al. “Facial Attributes Guided Deep Sketch-to-Photo Synthesis”,2018 IEEE Winter Conference on Applications of Computer Vision Workshops20
[44] Yuke Fang Yuke Fang “Identity-aware CycleGAN for face photo-sketch synthesis and recognition,” Pattern Recognition,Volume 102,2020,107249,ISSN 0031-3203.
[45] Wang, V. Sindagi and V. Patel.,"High-Quality Facial Photo-Sketch Synthesis Using Multi-Adversarial Networks", 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), Xi'an, 2018, pp. 83-90.
[46] W. Zhang, X. Wang, and X. Tang, “Coupled information-theoretic encoding for face photo-sketch recognition,” in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2011, pp. 513– 520.
[47] X. B. Gao, N. N. Wang, D. C. Tao, X. L. Li., “Face sketch photo synthesis and retrieval using sparse representation,” IEEE Transactions on Circuits and Systems for Video Technology, vol.22, no.8, pp.1213–1226, 2012. DOI: 10. 1109/TCSVT.2012.2198090.
[48] H. S. Bhatt, S. Bharadwaj, R. Singh, M. Vatsa, “On matching sketches with digital face images.,” in Proceedings of the 4th IEEE International Conference on Biometrics: Theory, Applications and Systems, IEEE, Washington DC, USA, 2010. DOI: 10.1109/BTAS.2010.5634 507.
[49] R. Yi, M. F. Xia, Y. J. Liu, Y. K. Lai, P. L. Rosin, “Line drawings for face portraits from photos using global and local structure based GANs”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.43, no.10, pp.3462–3475, 2021. DOI: 10.1109/TPAMI.2020.2987 931
[50] Fan, Deng-Ping, et al. "Facial-sketch synthesis: a new challenge." Machine Intelligence Research 19.4 (2022): 257-287.