A Systematic Review of Face Sketch Recognition System

Nur Nabilah Bahrum, Samsul Setumin*, Mohd Firdaus Abdullah, Mohd Ikmal Fitri Maruzuki, and Adi Izhar Che Ani

Abstract— **Face recognition systems have grown in popularity worldwide due to technological advancements. The use of a sketch to identify a suspect is one of the most common approaches in the field of forensic science when it comes to the application of face recognition algorithms. The difficulties arise when no image capture sources are available. Without an image, it becomes challenging for the police to identify the suspect. In forensic sciences, the identification of face sketch is used when the eyewitness of the crime scene is asked to recall the structures and features of the suspected person, and then the forensic artist will generate a sketch based on the verbal description provided from the eyewitness. Matching these sketches to the corresponding photo is a difficult task. Even though the sketches contain all of the necessary information about the suspect face's spatial topology and geometric details, it is still difficult for manual face recognition to identify and match sketch representations to the corresponding photo. Therefore, an automatic matching method is needed to identify a suspect accurately. This study perform a systematic literature review (SLR) that focuses on discussing of three major aspects of the face sketch recognition system which are the available face sketch database, the recent technique used sketch-photo matching and the contribution of the Generative Adversarial Network (GAN) in face sketch recognition. In this study, 35 different research papers published between 2011 and 2021 are studied and analysed. From this SLR, compared to other methods, GAN had been found to be a great technique for solving a wide variety of general facial recognition problems and also face synthesis.**

Index Terms—**Systematic review, face sketch recognition, deep learning, Generative Advesarial Network (GAN)**

I. Introduction

FACE sketch recognition is the process of matching a given face sketch image to a face photo from a large dataset. Face sketch recognition is a technique commonly used in criminal investigations. Face matching is a complicated process due to the fact that it is obtained from a variety of sources and under a variety of conditions [1].

Progress in biometric technology has provided law enforcement agency with additional tools for identifying criminals. In addition to DNA evidence and circumstantial evidence, if a latent fingerprint is discovered at an investigative scene or if a surveillance camera captures an image of a suspect's face, these cues may assist in determining the criminal's identity through automated biometric identification [2]. However, difficulties arise when no image capture sources are available. Without an image, it becomes extremely difficult for the police to identify the suspect [1].

Many crimes occur when none of this information is available, but an eyewitness explanation of the crime is. In situations like these, it is common for the forensic artist to join efforts in conjunction with the victim to produce a sketch of the suspect's face based on the witness's description.. After the sketch image of the suspect is done, it is will distributed to law enforcement officers and media outlets, hoping that someone knows who the suspect is [2]. Previously, face sketch recognition was done manually by a human observer, which is a process that takes a lot of time and whose accuracy can be affected by the level of human expertise. Fortunately, the technology nowadays could automatically perform the face sketch recognition systems, which could reduce a lot of processing time, reduce error-prone, and increase the accuracy. The face sketch recognition system begins with artists manually drawing a sketch, which is then matched with the photos in the database to determine the most similar image to the sketch and thus identify the person.

The accuracy of face sketch recognition is affected by various factors, including the artists' skills, the matching algorithm, and the types of sketches. The sketches could be divided into three categories which are viewed sketch, semi-forensic sketch and forensic sketch [3]. Related to the fact that the study of face sketch recognition has triggered a lot of interest, with a large number of face recognition solutions being used in a variety of applications such as in automobile security, access control, immigration, education, retail, and healthcare. Indeed, according to information obtained from the scientific indexing service Scopus, the field of research concerned with face sketch recognition has shown a rising level of publication activity from 2011 to 2022, as illustrated in Figure 1.

This manuscript is submitted on 14rd June 2022 and accepted on 11th November 2022. Nur Nabilah Bahrum, , Samsul Setumin, Mohd Firdaus Abdullah, Mohd Ikmal Fitri Maruzuki, Adi Izhar Che Ani are from the College of Engineering, Universiti Teknologi MARA, 13500 Permatang Pauh, Pulau Pinang, Malaysia.

*Corresponding author
Email address: samsuls@uitm.edu.my

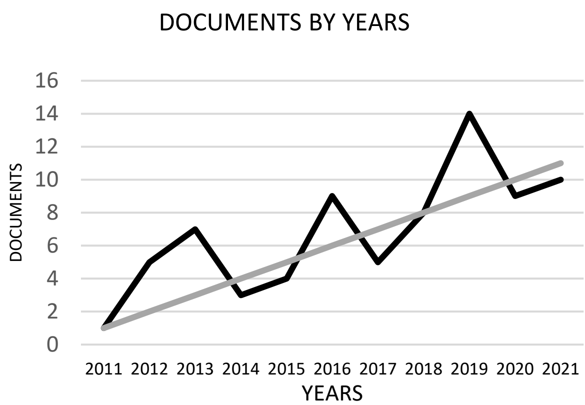

Fig. 1. Count of published studies of face sketch recognition indexed by Scopus from 2011 to 2021. Search string: "Face sketch recognition"

There is a need for secondary studies in this context as new investigations and proposals for original methods (primary studies) emerge. These studies are used to identify scientific gaps and propose taxonomies based on the technical characteristics of preliminary studies. For example, in [4], Kokila et al. presents a study and analysis of the various techniques that have been used for face sketch recognition, as well as the classification of some aspects such as acquisition technologies, face sketch types, databases, and also feature extraction techniques. In this study various difficulties encountered from the past studies during the matching process had been highlighted.

However, Akram et al. [5] had perform a comprehensive comparative study on the representative face sketch synthesis methods. The methods could be classified into two categories which are data-driven and model-driven. As mention in this study the model-driven methods explicitly learn the mapping from face photos to face sketches. Recently, Nikkath et al [6] conduct a survey on the most advanced Artificial Intelligence technique for digitally ascertaining a criminal through a facial recognition system. In this study, the most recent advancement in artificial intelligence is converting a forensic sketch into a real photo using the Deep Convolutional Generative Adversarial Network (DCGAN), which allowed the police to digitally ascertain the criminal's identity.

The past study provides an updated review of face sketch recognition research. By concentrating on elements that have received little attention in secondary studies, such as statistics in this domain, it helps other researchers understand the current state of the field's publications. In contrast, our contributions lies into three major categories which are identifying the types of datasets used in the face sketch recognition, characterization and description of the most frequent techniques used in recent face sketch recognition studies, and the contribution of Generative adversarial networks (GANs) in face sketch recognition. Providing a detailed description of those aspects is intended to guide newcomer researchers in the face sketch recognition field while also drawing attention to weaknesses that the community in future research should address. This study was conducted as a Systematic Literature Review (SLR).

II. PROCESS OF SYSTEMATIC REVIEW

An SLR helps identify, select, and critically evaluate the research to answer a particular question [7]. The systematic review should follow a protocol or plan with clearly stated criteria before start the study. It is a comprehensive, transparent search of multiple databases and grey literature that other researchers can replicate. It requires developing a well-planned search strategy with a specific focus or answering a specific question. The review specifies the categories of data that were gathered, analyzed, and reported within predetermined timeframes. The review needs to include the limits, search terms, and search strategies (like database names, platforms, and search dates) [8]. Figure 2 shows the activities performed at each stage of the systematic review process employed in this study, including planning, conducting, and reporting. This section provides a detailed description of the outcomes of the planning and execution-phase activities.

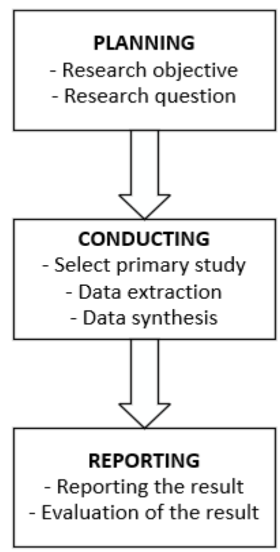

Fig. 2. Systematic literature review process

A. Research Objective and Question

The primary goal of the SLR reported in this study is to investigate the methods that have been established for face sketch recognition from 2011 to 2021, focusing on three major aspects which are identifying the types of datasets used in the face sketch recognition, characterization and description of the most common techniques used in recent face sketch recognition studies, and the contribution of GANs in face sketch recognition face.In order to illustrate both methodological and technical aspects, this study will focus on the factors that are related to the research questions, which are covered in more detail in the following section. The following research questions have also been established to serve as the basis for this review, in addition to the research objective:

- Q1 – What is the available database for face sketch recognition?
- Q2 – What are the most frequently used technique in face sketch recognition based on past study from 2011 to 2021?
- Q3 – What is the contribution of GANs in recent face sketch recognition system?

B. Automated and Manual Search

In order to conduct this study, four platform databases were used which are Google Scholar, IEEE, Scopus and Web of Sciences (WoS). All the database selected is well-known databases that index significant journals and conference papers in the fields of science and technology, as well as based on the authors' analysis of preliminary investigation results. The search was conducted in April 2022 and was limited to studies around 10 years ago that were published between 2011 to 2021, including those available via "early access" in the databases. The language used was only English, and the search string used for each search tool is "face sketch recognition".

For conducting a manual search for primary studies, we consulted the references lists of recent secondary studies on the topic of "facial sketch recognition system". The manual search is not intended to be exhaustive but rather complementary, allowing researchers to incorporate potentially relevant studies that were not found through a specific search string. This is because it is anticipated that a significant number of primary studies will be retrieved from bibliographic databases through automated searches. In order to accomplish this, the surveys from the previos study had been selected as sources of manual search that were made available in 2021 and included face sketch recognition in the discussion.

C. Selection Criteria

Following the automated and manual searches, a manual selection of the necessary documents is carried out in two stages which are preliminary selection and the final selection. The purpose of the preliminary selection is to cut down the number of documents that must be read in the full text which is only specific parts of the documents are evaluated and the documents that are accepted will be through a second round of evaluation by reading in the full text. In order to reduce the bias, two distinct sets of selection criteria have been defined, which are inclusion and exclusion criteria.

- Inclusion criteria:
  - Papers from 2011 to 2021
  - Only journal and conference paper that discuss the face sketch recognition system.
- Exclusion criteria
  - Non-journal and non-conference articles
  - Study is not available in full-text
  - Study not available in English
  - Duplicate study from different database
  - Study does not mention the dataset that had been used.

D. Search and Selection Result

In this part, there are three main stages for the search and selection result of this systematic literature review which are preliminary selection, duplicate removal and final selection. For the preliminary document selection process, the document had been obtained from four database which are Google Scholar, IEEE, Scopus and Web of Science. During this stage, the titles and abstracts of the 319 documents that had been obtained from four databases had been read. Next, In the duplicate removal stage, duplicates are expected to be obtained. For example, the database in Scopus has a duplicate document in Web of Science, meaning one paper exists into 4 databases and since there is also an overlap between documents that are indexed in each database and those that are listed as references of papers that were chosen as sources for the manual search. After removing duplicates there were remained a total of 79 documents. Consequently, a total of 35 documents were accepted for the final selection, each of which was evaluated by reading its full text, and Fig. 3 illustrates the summary search and selection procedures.

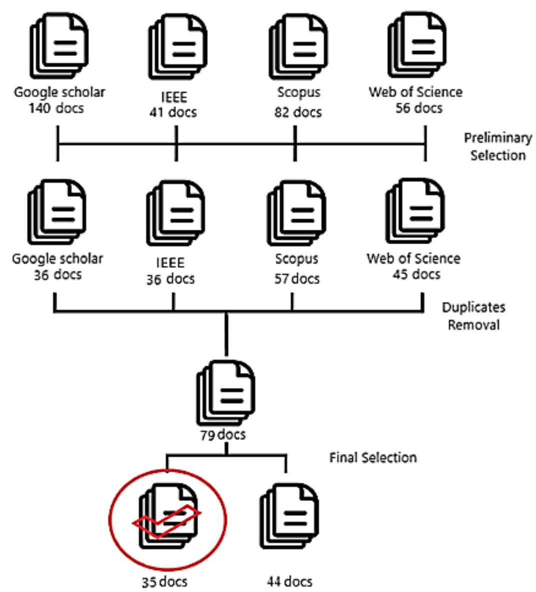

Fig. 3. Summary of the documents' search and selection procedures

III. FACE SKETCH DATABASE

In this section, the publicly available database containing face sketches with their corresponding photo that are widely used in the reviewed paper will be discussed. In 2017 Kokila et al. [4] had conduct a study and analysis of various techniques to match sketches to Mugshot photos, in this study the Autor had provided the information about the available face sketch database that had been used in the literature review, however this study is focusing more on the technique that had been used in the past study and based on the table 1 in [4], our study will elaborate and discuss more about the available face sketch database based on the past study from 2011 to 2021.

The face sketch recognition system begins with artists drawing a face sketch, which is then compared to various photos in the database to determine the image that is the closest match to the sketch and thus helps identify the person. Face sketch recognition accuracy is affected by various factors, including the artists' skills, the matching algorithm, and the type of sketches. Sketches can be classified into three types which are viewed sketches, semi-forensic sketches, and forensic sketches [3].

For the viewed sketches, the artist can either look at a photo of the subject or directly at the subject when creating the sketch. As a result, these sketches contain a great deal of information about the original subject, which leads to improved accuracy.

However, for the semi-forensic sketch method, the artist is allowed to view a picture of the subject for a brief period of time and then be asked to draw the corresponding sketch of the photo from his memory. This method of sketches is also used in academic research, it includes the memory component and is considered more relevant to real-world challenges than viewed sketches [9]. Next, forensic sketches are drawn by the artist based on the explanation provided by an eyewitness. These are real-life examples obtained primarily from law enforcement agencies (and therefore, very few image sets are available for research purposes). The most challenging problem is recognising forensic sketches because the eyewitness may have only seen a face for a brief moment, usually in a stressful situation, and sketch formation depends on the eyewitness's description and the sketch artist's expertise [9].

According to the SLR, the IIIT-D database contains all three types of sketches which are viewed, semi-forensic sketches, and forensic sketches. The IIIT-D Viewed Sketch Database is comprised of 238 sketch-digital image pairs that were created by a professional sketch artist using digital images gathered from various sources. The sketches in the IIIT-Delhi Semi-forensic Sketch Database were created using the artists' memories rather than eyewitness descriptions and a sketch artist also draws the forensic sketches for the Forensic Sketch Database based on the description of an eyewitness and their recollection of the crime scene. However, the IIIT-D database is no longer available, and for the past ten years, only one researcher has used this database to perform a study of face sketch recognition.

From the SLR, most of the study for face sketch synthesis and recognition uses the Chinese University of Hong Kong (CUHK) Face Sketch database (CUFS). It contains 188 faces from CUHK student database, 123 faces from the AR database and 295 faces from the XM2VTS database. There are a total of 606 faces. The type of sketches from this database is in viewed sketches categories because an artist has created this sketch based on a photo taken in a frontal pose and with a neutral expression. Other than that, CUHK also provides a large dataset that consists 1194 of face sketches that are also be creating using the viewed method, and this dataset has been known as CUHK Face Sketch FERET Database (CUFSF) [10].

On the other hand, there is also a study of face sketch recognition using software-generated sketches that can be accessed from the E-PRIP database. The E-PRIP has 123 pairs of composite sketches and photos made from the AR Face dataset's photos. The composite sketches were divided into four groups created with a different subject. One set was created by an American artist using the FACES software, two sets by Asian artists using both the FACES and Identi-Kit tools, and one set by an Indian artist using the FACES software [10].

Based on the SLR, in order to measure the face sketch recognition system performance more accurately, nine group of researchers which are [15], [21], [25], [26], [35], [37], [38], [39], [40] use the forensic sketches types for their dataset that can be accessed from Pattern Recognition and Image Processing Hand-Drawn Composite (PRIP-HDC) database. PRIP-HDC database is very different from other public-domain sketch databases, like CUHK's, which are drawn by looking at a picture. These hand-drawn composites are drawn according to the eyewitness or victim description, and the release of this dataset contains 47 pairs of hand-drawn composites-mugshot pairs. However, UoM-SGFS also provides forensic sketch types but uses a software-generated method containing 300 subjects in the Color-FERET database, which were created using the EFIT-V software. A qualified forensic scientist trained the operator of the EFIT-V from the Malta Police Force to ensure that real-world practices were followed in the creation of the UoM-SGFS database. Based on the SLR, there are five group of the researchers are using the UoM-SGFS database in their study while only four group of researchers had been used the PRIP-HDC database in their study. Table 1 below shows the types of the database used in the past study based on the SLR.

TABLE I
TYPES OF DATABASES USED IN PAST STUDY BASED ON SLR

Database		Type of sketch	Method of sketch	Used in
CUFS	CUHK	Viewed	Hand-drawn	[2] [11]][13] [14] [15] [16] [17] [18] [19] [20] [21] [22] [23] [24] [25] [26] [27] [28] [29] [30]
	AR Database	Viewed	Hand-drawn	
	XM2VTS	Viewed	Hand-drawn	
CUFSF		Viewed	Hand-drawn	[11] [14] [31] [32] [18] [20] [23] [33] [26] [27]
E-PRIP		Viewed	Software-generated	[9] [34] [35] [25] [26] [36]
IIITD		Viewed	Hand-drawn	[27]
		Semi forensic	Hand-drawn	
		Forensic	Hand-drawn	
PRIP-HDC		Forensic	Hand-drawn	[15] [37] [38] [21]
UoM-SGFS		Forensic	Software-generated	[35] [39] [40] [25] [26]

A. *Discussion of face sketch database*

Based on the SLR, there are five types of databases that had been used in the past study of face sketch recognition from 2011 to 2021 which are CUFS, CUFS, E-PRIP, IITD, UoM-SGFS and PRIP-HDC. Sketches can be classified into three types which are viewed sketches, semi-forensic sketches, and forensic sketches [3]. From the table 1, IITD databased is consist all three types of sketches which are viewed, semi-forensic and forensic. Other than that, CUFS, CUFS, and E-PRIP databases consist of viewed sketches, while PRIP-HDC and UoM-SGFS consist of forensic sketches.

The method used to generate the sketches is divided into two categories which are hand-drawn and software generated. CUFS, CUFSF, IIITD and PRIP-HDC databases used the hand-drawn method, while E-PRIP and UoM-SGFS used a software-generated method to produce the sketches. Figure 4 below shows the graph of the database that had been used in the past study based on the SLR. The graph shows that most of the researchers used the viewed sketches database from CUFS, CUFSF and E-PRIP compared to the forensic and semi forensic sketches database. This is because the forensic sketches database (PRIP-HDC and UoM-SGFS) and semi forensic sketches (IIITD) has fewer sketches than the viewed sketches from CUFS, CUFSF and E-PRIP database, and most of the studies required a large number of the dataset to achieve a high-performance accuracy.

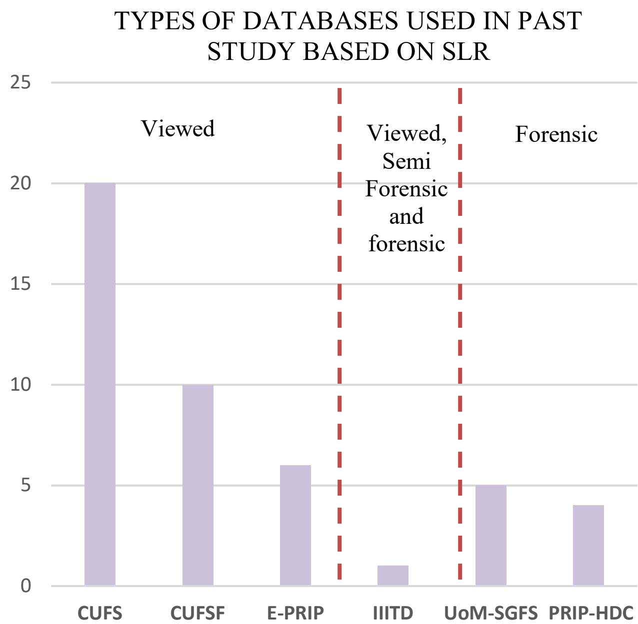

Fig.4. Graph of the database types used in past study based on SLR

IV. FACE SKETCH RECOGNITION TECHNIQUES

A. *Hand-Crafted Features*

Generally, face sketch recognition techniques can be categorized as either component-based or holistic-based techniques. The eyes, nose, and mouth are some of the facial features from the sketch that can be extracted using component-based approaches, which are primarily designed for this purpose. Then the sketch is matched by comparing its components to the photo in the database. However, the Holistic-based approaches are concerned with incorporating the entire face region into the recognition process. Tang and Wang [41] had performed early work on methods based on a holistic approach. In this paper, Eigen transformation was implemented, and the reported results indicate that this method had achieved a 71% recognition accuracy using the CUHK database. In addition to the low recognition rate reported, this study required a large number of training samples.

The further work for face sketch recognition is by using Local Binary Patterns (LBP), which had been discussed in [11], [14], [17]. In 2012, Kiani et al. [11] performed a study based on the LBP method by using the Local Radon Binary Pattern (LRBP). The reported results indicate that LRBP outperforms previous LBP-based face descriptors in recognising face sketches using CUFS and CUFSF datasets. To improve the system accuracy by using the same dataset, Alex et al. [14], had developed the Local difference of the gaussian binary pattern (LDoGBP). The proposed method provides a higher recognition rate despite having a very low computational complexity, proving that LDoGBP performs better than most LBP-based descriptors with the accuracy achieved 100% at rank 5 for the CUFS database and 98.9% for the CUFS database at rank 10.

Other than that, further work was performed by using Histogram oriented gradient (HOG), which had been used in [13], [31], [40]. In 2012 Galoogahi et al. [13] had performed a study using Histogram of Averaged Oriented Gradients (HAOG), by developing a new face descriptor based on gradient orientations, this study is focusing on reducing the modality difference during the feature extraction stage. As mention in this study, HAOG is different from the Histogram oriented gradient (HOG). This is because HOG is computed entirely on images with both fine and coarse textures, but HAOG is only extracted from coarse textures, which are less sensitive to differences in modality. In 2020, Xu et al.[40] developed a face sketch recognition algorithm based on high-level semantic attributes and multi-scale HOG features. In this study, the structure and detail features are represented by the global HOG features of the face and the local HOG features of each face component. The candidate matching list at the score level is created by combining the contour and detail features. The proposed method achieves rank 10 identification accuracy of 86.6 and 96.7 percent, respectively, when tested on the PRIP-VSGC and UoM-SGFS databases..

Next, another researcher used a scale-invariant feature transform (SIFT) method for face sketch recognition which had been discussed in [2] [15], [17], [42]. In 2011, Brendan et al. [2] proposed a study that describes a reliable method for matching composite sketches to large mug shot (image) databases kept by law enforcement agencies. This paper presented a framework that had been called local feature-based discriminant analysis (LFDA). In LFDA, SIFT and Multiscale Local Binary Pattern (MLBP) descriptors on each patch. For minimum distance matching, multiple discriminant projections are applied to partitioned vectors of the feature-based representation. Both viewed and composite sketches are used in this study. For viewed sketch matching, the CUHK dataset is used. A total of 159 composite sketches are matched against 10,159 mugs shot images in the gallery. The size of the target gallery is reduced during the recognition process by using race and gender information to improve the matching performance. This study shows that the proposed framework achieves cutting-edge accuracy when matching viewed sketches.

Next in 2016, Aziz et al. [15] conducted a study using SIFT feature and Euclidean distance in the face recognition process on CUHK viewed-sketch and PRIP-HDC forensic sketch database. The accuracy of this study was achieved 96% using the CUHK dataset, and for PRIP-HDC dataset showed improvement in the recognition performance from 34% to 57% at rank 50.

Tharwat et al. [17] conducted a study to compare the LBP and SURF methods in the face sketch recognition system. In this study, a Minimum distance and Support Vector Machine (SVM) classifiers are used to match the features of an unknown sketch with photos. The reported result indicates that the SIFT algorithm achieved a better recognition task performance than the LBP algorithm. However, In 2017, Kokila et al. [42]. had proposed a study on a novel feature-based approach for comparing sketches and mugshot photos by combining the SIFT features extraction with Speed Up Robust Features (SURF) algorithms, and the resulting features are matched using the nearest neighbour algorithm. The accuracy performance of this study was archived 95.45% at rank 50.

B. Learned Features

Deep Learning is being used to solve difficult problems in digital image processing, for example, image colourization, classification, segmentation and detection. With the help of big data and abundant computing resources, deep learning methods such as Convolutional Neural Networks have improved prediction performance while also testing the bounds of what was previously possible. By using deep learning, unsolvable problems in digital image processing are now being solved with superhuman precision [43]. Deep learning techniques are now frequently used to replace traditional techniques, especially in face recognition and classification tasks. Deep learning approaches have been utilized to address the problem in sketch-photo recognition by learning the relationship between the two modalities of the sketch and photo..

However, from the standpoint of deep learning, the sketch recognition problem is more difficult than the traditional face recognition problem. This is due to the diversity of sketch and photo modalities, as well as the limited availability of large databases to avoid over-fitting and local minima. Most datasets, for instance, only contain one sketch per subject, making it extremely difficult for a deep model to learn robust features [37]. Because of that, many deep techniques employ shallow models or train the network exclusively on the photo modality [9]. The researchers in [9], [20], [25], [26], [27], [28], [29], [32], [35], [36], [37], [38], [39] had been conducted a study on facial sketch recognition using deep learning approach.

One of the first methods that used deep learning for face sketch recognition was proposed in [9], Mittal et al. develop a novel algorithm for matching composite sketches with photographs, which combines transfer learning with deep learning representation in order to achieve a good result. In the proposed algorithm of this study, the deep learning architecture-based facial representation is first learned from a large face database of photos, and then the representation is updated from a small problem-specific training database. In this study, the autoencoder and Deep Belief Network (DBN) were used to reduce the obtain dimensionality representation used for recognition. Next, in 2017, Galea et al. [37] conducted a study using deep learning for forensic sketch recognition. The proposed method was found to be effective for actual forensic sketches. The result reported that this proposed method could reduce the error rate by 80.7% for viewed sketches and lowers the mean retrieval rank by 32.5% for real-world forensic sketches.

Other than that, the further work was done by Patil et al. [38] this study used Multiscale Local Binary Patterns (MLBP), Tchebichef Moments, and Multiscale Circular Weber Local Descriptor (MCWLD) to extract texture features. Principal Component Analysis (PCA) was used to fuse the extracted features, and Deep Convolutional Neural Network (DCNN) was used to recognize the face. This study demonstrates that multiple feature extraction methods can reduce complexity and improve precisionIn this study, the system accuracy was archived 97% by using the PRIP-HDC dataset. Next, Wan et al. [20] conducted a study on feature representation for face sketch recognition using the VGG-Face network. This study becomes the first research that transforms the gallery photos into sketches to decrease the modality difference, and this method outperforms others on multiple face photo-sketch datasets.

However, Mendez et al. [39] perform a study that used a local deep learning approch for the recognition of composite face sketch images. The structures and features from intermediate layers of deep learning models that have already been trained are extracted using a component-based approach, and LDA metric learning is applied to improve the discrimination between different types of sketches when compared to real photographs. In this study, two deep models were proposed which are DEEPS had been used to train for sketch recognition, and ResNet-Dlib had been used trained for face recognition. This study was performed using UoM-SGFS and PRIP-VSGC datasets, and it was discovered that the proposal based on local features from the ResNet-Dlib model achieves the highest recognition rates in general.

Other than that, the researchers in [35], [26] and [29] have conducted a study on face sketch recognition using Siamese Network. The Siamese neural network is consists of two or more identical subnetworks. A configuration with the same parameters and weights is referred to as identical in this context. Neural networks based on Siamese neural networks are commonly used in tasks that require the comparison of two similar items. A pair of inputs are needed for a Siamese neural network to learn the distance margin, such as a positive pair of similar or nearly similar images and a negative pair of dissimilar images. It is difficult to build a robust model from a single sample when the image has a lot of variations.

In 2019, Fan et al. [35] conducted a study on face sketch recognition using the Siamese Convolution Network (SCNN). But, due to the used of extracted features that cannot eliminated the effect of different image modalities, these methods cannot achieve a higher recognition rate for all face sketch datasets. To improve the system accuracy, the same researcher [26] conducted a study using Siamese Graph Convolution Network (SGCN) for face sketch recognition. This study reported that SGCN could perform well on some face photo-sketch datasets. The model performance of this study based on the graph structure representation of the data using the SGCN is more stable than an SCNN model. However, in 2021, Sabri et al [29] performed a study using SCNN and evaluated system performance with four different activation functions. The reported result of this study shows that the sigmoid activation function had achieved an accuracy of 100% as early as 300 learning iterations for 10-way one-shot learning using the CUHK dataset.

C. *Discussion of Face Sketch Recognition Techniques*

From the SLR, the rank-based method had been used in evaluated the system performance of the study. For example, using Cumulative Match Curve (CMC). CMC will perform a cumulative measurement across the ranks to determine the percentage of correct identification. The accuracy in rank-1 simply refers to the percentage of correct matches that are made solely on the basis of the shortest distance, which is similar to the retrieval rate. The retrieval rate of the correct match during the first k smallest distances is given by rank-k accuracy. For example, if the rank-k percentage is 100%, which indicates that the method used in the study is capable of shortlisting face candidates without an error.

In this SLR, 35 selected papers were divided into two categories based on the study's technique. Figure 5 shows the graph of face recognition techniques that have been used from 2011 to 2021. From the graph, 19 studies had used the hand-crafted features for the face sketch recognition system, while the other 16 studies used the learned features approach. The number of papers using the hand-crafted features is higher compared to the learned features. This is because learned features, especially deep learning is a new study approach and started popular in face sketch recognition system in 2015.

But recently, learned features methods have replaced traditional techniques based on hand-engineered features, which provided the higher accuracy compared to traditional method. Due to the substantial gains in system accuracy over other approaches, deep learning-based become a standard approach in face recognition system. In addition, it is simple to scale up these systems to achieve even higher accuracy by increasing the size of the training datasets or the network capacity.

However, there has been limited use of deep learning for face sketch recognition. The sketch is restricted and limited due to privacy concerns and the time-consuming human labeling efforts involved in sketch creation. Deep neural networks have a difficult time detecting robust features in images because there is only one sketch per subject in the available face sketch database. In addition, there is generally only one sketch per subject in the available face sketch database, making it challenging for a deep neural network to identify the image's robust features.

Lately with the development of deep learning, GAN has been introduced widely. GAN are algorithmic architectures that employ two neural networks, pitting one against the other to generate new, synthetic data instances that can pass for actual data. They are generally used in generating images, videos, and voices. In image generating, GAN could produce a great outcome, even having a less number of images per subject. This is because, from the SLR, GAN could generate new faces or new identity by using the dataset that consist one image per person or subject such as CUHK dataset. In past 10 years, three group of researchers have used GAN in face sketch synthesis and recognition. This is due to the GAN was utilised to enhance images by concentrating on realistic textures rather than pixel accuracy, which resulted in higher image quality when viewed at high magnification. Thus, using the images generated from GAN could help increase the accuracy of the face sketch recognition system.

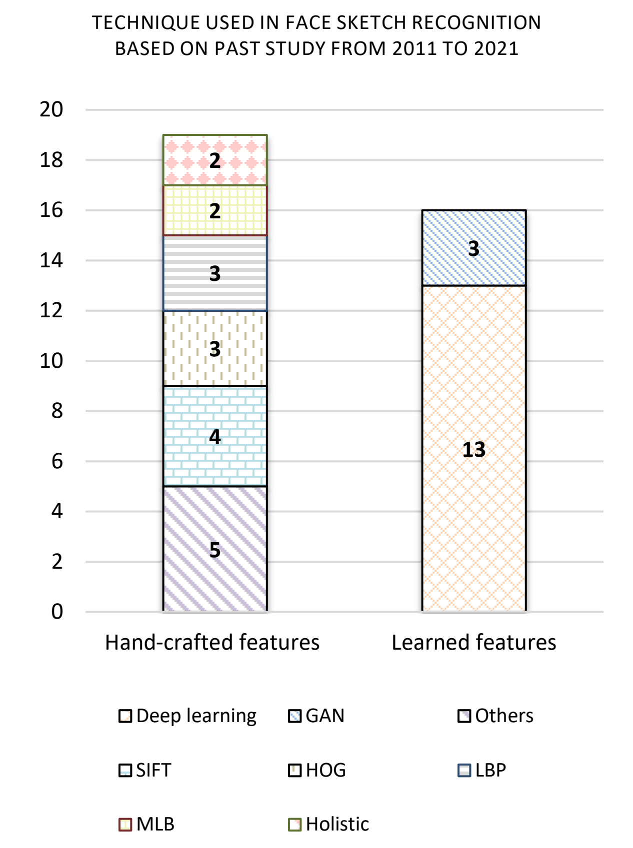

Fig. 5. Graph of face sketch recognition technique that had been used in past study from 2011 to 2021

V. GENERATIVE ADVERSARIAL NETWORKS (GAN) IN FACE SKETCH RECOGNITION

In 2014, Goodfellow et al. introduced GANs, which consist of a generative model and a discriminative model jointly trained in an adversarial way [44]. The generative model generates data with a distribution similar to the training data, whereas the discriminative model estimates the probability that a sample comes from the training data as opposed to the generative model. In general, both models are multilayer perceptrons, and their abilities are poor at the start of training. The competition between the two models improves their abilities, allowing the generative model to generate data with a similar distribution to the training data. GANs produce impressive outcomes without requiring a complicated procedure.

In 2018, with the advancement of deep learning, GAN became popular in generating high-quality realistic photos from sketches and sketches from photos. GAN-based techniques have demonstrated outstanding performance on image-to-image translation problems and, in particular, photo-to-sketch synthesis. Yu et al. [33] proposed a conditional CycleGAN to generate a pseudo image. This study specifically designed a feature-level loss to induce the network to produce high-resolution and high-quality pseudo photos. Moreover, by combining the advantages of CycleGAN and conditional GANs, the synthetic images were suitable for face recognition against an image gallery.

Next, Wang et al. [23] proposed a Photo-Sketch Synthesis utilizing Multi-Adversarial Networks (PS2-MAN). This novel synthesis framework generates low-resolution to high-resolution images in an adversarial style using an iterative process. The hidden layers of the generator are supervised to produce lower-resolution images. Then the network is refined to produce higher-resolution images. In addition, since photo sketch synthesis is a paired translation problem, this study use the CycleGAN framework to leverage the pair information. In this study, image quality evaluation and photo-sketch matching were conducted to demonstrate that the proposed framework outperforms the existing state-of-the-art method.

In contrast, Sannidhan et al. conducted research on the generation of face sketches using generative adversarial networks. This study emphasized the use of a conditional Generative Adversarial Network (cGAN) to generate color photo images from facial sketches, allowing for efficient classification while avoiding cross-domain problems in sketch identification. The quality of the sketches used to train the adversarial network determines the accuracy of the image generated by GAN. So, in order to improve the training process, this study proposed combining two different deep learning systems which are a trained convolution neural network is used for sketch generation, and a conditional generative adversarial network's pix2pix model is used for colour photogeneration. The result of this study shows that the quality of photos generated from trained CNN sketches is better compared to the actual sketches in the dataset.

A. Discussion of GAN in Face Sketch Recognition

GAN-based methods have significantly improved image synthesis based on SLR, but they do not explicitly consider the purpose of recognition. GANs have gained popularity and made significant progress in image-to-image translation tasks in recent years. The discriminator in a typical GAN model determines whether the inputs are real or fake. Simultaneously, the generator learns to generate sharper and more realistic samples that are impossible to distinguish from actual samples. For example, the Pix2Pix model employs conditional GAN to perform supervised image style translation, and CycleGAN preserves key attributes between the input and translated images using cycle consistency loss to reduce the difficulty of obtaining image pairs.

However, when applied to the face photo and sketch synthesis task, all of these frameworks can only learn the relationships between two domains, with the discriminator focusing solely on the differences between the photos and sketches and disregarding any optimization for recognition. In order to solve these problems, each face image must contain its identity-specific information. Therefore, Zhiwu et al. [45] conducted a study on state-of-the-art image-to-image translation networks (Cycle-Consistent Adversarial Networks) of the image to video/video to image translation context using an image-video translation model and an identity preservation model. These concepts also can be applied for the photo to sketch/sketch to photo translation by supervising the relationship between the real target and the fake image with two recognition networks. Other than that, GAN also could generate photos directly from the verbal description [46]. This could help reduce the time consuming of the artist's sketching style in producing the sketches and reduce the transformation loss, which could lead to better recognition accuracy.

VI. CONCLUSION

This study is conducting a systematic review of implementation of face sketch recognition research between 2011 and 2021. The core objective of the SLR is to answer three research questions addressing the subject from three main aspects of face sketch recognition studies which are identifying the types of datasets used in the face sketch recognition, the techniques used in face sketch recognition studies from 2011 to 2021, and the contribution of GANs in face sketch recognition. The research is limited to articles published in journals and conferences paper that proposes a study of face sketch recognition system. In order to select the research papers that are able to provide answers to the research questions, a methodical and careful filtration strategy has been applied.The face sketch recognition accuracy is affected by various factors, including the artists' skills, the matching algorithm, and the type of sketches. Based on the systematic review, deep learning-based techniques have emerged as a new trend in face sketch recognition. The transformation process of deep learning-based approaches is fast and has a better performance compared to the traditional learning-based approaches. However, it required a large dataset for training progress to avoid overfitting, and the sketch images also usually lack facial details, which could lead to serious noise effects. In order to address the challenges mentioned earlier, more future work needs to be presented for the face sketch recognition system approach based on a generative adversarial learning framework, this is because GAN had become more popular and commonly used to solve a wide range of general facial recognition problems such as face synthesis, cross-age recognition, poseinvariant face recognition, video-based and makeup-invariant face recognition, and more.

ACKNOWLEDGMENT

The authors would like to thank Universiti Teknologi MARA, Cawangan Pulau Pinang for supporting the research work, especially the Centre for Electrical Engineering Studies.

REFERENCES

[1] M. S. Thangakrishnan and K. Ramar, “A survey on forensic sketch matching,” *IIOAB J.*, vol. 6, no. 4, pp. 50–54, 2015.

[2] B. Klare, Z. Li, and A. K. Jain, “Matching forensic sketches to mug shot photos,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 3, pp. 639–646, 2011, doi: 10.1109/TPAMI.2010.180.

[3] H. Cheraghi and H. J. Lee, "SP-Net: A novel framework to identify composite sketch," *IEEE Access*, vol. 7, pp. 131749–131757, 2019, doi: 10.1109/ACCESS.2019.2921382.

[4] R. Kokila, M. S. Sannidhan, and A. Bhandary, "A study and analysis of various techniques to match sketches to Mugshot photos," *Proc. Int. Conf. Inven. Commun. Comput. Technol. ICICCT 2017*, no. Icicct, pp. 41–44, 2017, doi: 10.1109/ICICCT.2017.7975243.

[5] A. Akram, N. Wang, J. Li, and X. Gao, "A comparative study on face sketch synthesis," *IEEE Access*, vol. 6, pp. 37084–37093, 2018, doi: 10.1109/ACCESS.2018.2852709.

[6] S. Nikkath Bushra and K. Uma Maheswari, "Crime Investigation using DCGAN by Forensic Sketch-to-Face Transformation (STF)- A Review," *Proc. - 5th Int. Conf. Comput. Methodol. Commun. ICCMC 2021*, no. Iccmc, pp. 1343–1348, 2021, doi: 10.1109/ICCMC51019.2021.9418417.

[7] "Module 1: Introduction to conducting systematic reviews | Cochrane Training." https://training.cochrane.org/interactivelearning/module-1-introduction-conducting-systematic-reviews (accessed Apr. 25, 2022).

[8] "Systematic literature reviews - Literature Review - Library Guides at Charles Sturt University." https://libguides.csu.edu.au/review/Systematic (accessed Apr. 25, 2022).

[9] P. Mittal, M. Vatsa, and R. Singh, "Composite sketch recognition via deep network - A transfer learning approach," *Proc. 2015 Int. Conf. Biometrics, ICB 2015*, pp. 251–256, 2015, doi: 10.1109/ICB.2015.7139092.

[10] "Heterogeneous Face Databases — bob.bio.htface 1.0.4 documentation." https://www.idiap.ch/software/bob/docs/bob/bob.bio.htface/v1.0.4/databases.html (accessed Apr. 27, 2022).

[11] H. Kiani Galoogahi and T. Sim, "Face sketch recognition by Local Radon Binary Pattern: LRBP," *Proc. - Int. Conf. Image Process. ICIP*, pp. 1837–1840, 2012, doi: 10.1109/ICIP.2012.6467240.

[12] S. Pramanik and D. Bhattacharjee, "Geometric feature based face-sketch recognition," *Int. Conf. Pattern Recognition, Informatics Med. Eng. PRIME 2012*, pp. 409–415, 2012, doi: 10.1109/ICPRIME.2012.6208381.

[13] H. K. Galoogahi and T. Sim, "Inter-modality face sketch recognition," *Proc. - IEEE Int. Conf. Multimed. Expo*, pp. 224–229, 2012, doi: 10.1109/ICME.2012.128.

[14] A. T. Alex, V. K. Asari, and A. Mathew, "Local difference of gaussian binary pattern: Robust features for face sketch recognition," *Proc. - 2013 IEEE Int. Conf. Syst. Man, Cybern. SMC 2013*, pp. 1211–1216, 2013, doi: 10.1109/SMC.2013.210.

[15] H. G. M. Abel-Aziz, H. M. Ebeid, and M. G. M. Mostafa, "A mean features method for face photo-sketch synthesis and recognition," *ACM Int. Conf. Proceeding Ser.*, vol. 09-11-May-, pp. 107–113, May 2016, doi: 10.1145/2908446.2908492.

[16] H. G. M. Abdel-Aziz, H. M. Ebeid, and M. G. M. Mostafa, "An unsupervised method for face photo-sketch synthesis and recognition," *2016 7th Int. Conf. Inf. Commun. Syst. ICICS 2016*, pp. 221–226, 2016, doi: 10.1109/IACS.2016.7476115.

[17] A. Tharwat, H. Mahdi, A. El Hennawy, and A. E. Hassanien, "Face sketch recognition using local invariant features," *Proc. 2015 7th Int. Conf. Soft Comput. Pattern Recognition, SoCPaR 2015*, pp. 117–122, 2016, doi: 10.1109/SOCPAR.2015.7492793.

[18] W. Wan and H. J. Lee, "FaceNet Based Face Sketch Recognition," *Proc. - 2017 Int. Conf. Comput. Sci. Comput. Intell. CSCI 2017*, pp. 432–436, 2018, doi: 10.1109/CSCI.2017.73.

[19] K. Ounachad, A. Sadiq, and A. Souhar, "Fuzzy Hamming Distance and Perfect Face Ratios Based Face Sketch Recognition," *Colloq. Inf. Sci. Technol. Cist*, vol. 2018-Octob, pp. 317–322, 2018, doi: 10.1109/CIST.2018.8596665.

[20] W. Wan and H. J. Lee, "Deep feature representation for face sketch recognition," *Adv. Sci. Technol. Eng. Syst.*, vol. 4, no. 2, pp. 107–111, 2019, doi: 10.25046/aj040214.

[21] M. S. Sannidhan, G. Ananth Prabhu, D. E. Robbins, and C. Shasky, "Evaluating the performance of face sketch generation using generative adversarial networks," *Pattern Recognit. Lett.*, vol. 128, pp. 452–458, 2019, doi: 10.1016/j.patrec.2019.10.010.

[22] H. Samma, S. A. Suandi, and J. Mohamad-Saleh, "Face sketch recognition using a hybrid optimization model," *Neural Comput. Appl.*, vol. 31, no. 10, pp. 6493–6508, Oct. 2019, doi: 10.1007/s00521-018-3475-4.

[23] W. Wan and H. J. Lee, "Generative Adversarial Multi-Task Learning for Face Sketch Synthesis and Recognition," *Proc. - Int. Conf. Image Process. ICIP*, vol. 2019-Septe, pp. 4065–4069, 2019, doi: 10.1109/ICIP.2019.8803617.

[24] F. Liu, Y. Ding, F. Xu, and Q. Ye, "Learning Low-Rank Regularized Generic Representation with Block-Sparse Structure for Single Sample Face Recognition," *IEEE Access*, vol. 7, pp. 30573–30587, 2019, doi: 10.1109/ACCESS.2019.2903333.

[25] L. Fan, X. Sun, and P. L. Rosin, "Attention-Modulated Triplet Network for Face Sketch Recognition," *IEEE Access*, vol. 9, pp. 12914–12921, 2021, doi: 10.1109/ACCESS.2021.3049639.

[26] L. Fan, X. Sun, and P. L. Rosin, "Siamese graph convolution network for face sketch recognition," *Proc. - Int. Conf. Pattern Recognit.*, pp. 8008–8014, 2020, doi: 10.1109/ICPR48806.2021.9412917.

[27] V. A. Kumar, K. S. Rajesh, and R. Antony, "Cross Domain Descriptor for Face Sketch-Photo Image Recognition," *ACCESS 2021 - Proc. 2021 2nd Int. Conf. Adv. Comput. Commun. Embed. Secur. Syst.*, no. September, pp. 228–231, 2021, doi: 10.1109/ACCESS51619.2021.9563314.

[28] R. Sindhu, K. Prathyusha, S. Ravi, and M. C. Suman, "Implementation of Digital Forensics Face Sketch Recognition using Fusion Based Deep Learning Convolution Neural Network," *Proc. 5th Int. Conf. Electron. Commun. Aerosp. Technol. ICECA 2021*, pp. 1499–1504, 2021, doi: 10.1109/ICECA52323.2021.9676004.

[29] N. I. Ahmad Sabri and S. Setumin, "One-shot learning for facial sketch recognition using the siamese convolutional neural network," *ISCAIE 2021 - IEEE 11th Symp. Comput. Appl. Ind. Electron.*, pp. 307–312, 2021, doi: 10.1109/ISCAIE51753.2021.9431773.

[30] D. Irgasheva and L. Davronova, "Superpixel Based Face Sketch Recognition Scheme," *Int. Conf. Inf. Sci. Commun. Technol. Appl. Trends Oppor. ICISCT 2021*, pp. 2021–2023, 2021, doi: 10.1109/ICISCT52966.2021.9670266.

[31] C. Galea and R. A. Farrugia, "Fusion of intra-and inter-modality algorithms for face-sketch recognition," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9257, pp. 700–711, 2015, doi: 10.1007/978-3-319-23117-4_60.

[32] G. Cao, M. A. Waris, A. Iosifidis, and M. Gabbouj, "Multi-modal subspace learning with dropout regularization for cross-modal recognition and retrieval," *2016 6th Int. Conf. Image Process. Theory, Tools Appl. IPTA 2016*, pp. 0–5, 2017, doi: 10.1109/IPTA.2016.7821032.

[33] S. Yu, H. Han, S. Shan, A. Dantcheva, and X. Chen, "Improving face sketch recognition via adversarial sketch-photo transformation," *Proc. - 14th IEEE Int. Conf. Autom. Face Gesture Recognition, FG 2019*, 2019, doi: 10.1109/FG.2019.8756563.

[34] D. Liu, C. Peng, N. Wang, J. Li, and X. Gao, "Composite face sketch recognition based on components," *2016 8th Int. Conf. Wirel. Commun. Signal Process. WCSP 2016*, 2016, doi: 10.1109/WCSP.2016.7752734.

[35] L. Fan, H. Liu, and Y. Hou, "An Improved Siamese Network for Face Sketch Recognition," *Proc. - Int. Conf. Mach. Learn. Cybern.*, vol. 2019-July, 2019, doi: 10.1109/ICMLC48188.2019.8949231.

[36] H. T. Chethana and T. C. Nagavi, "A new framework for matching forensic composite sketches with digital images," *Int. J. Digit. Crime Forensics*, vol. 13, no. 5, pp. 1–19, 2021, doi: 10.4018/IJDCF.20210901.oa1.

[37] C. Galea and R. A. Farrugia, "Forensic Face Photo-Sketch Recognition Using a Deep Learning-Based Architecture," *IEEE Signal Process. Lett.*, vol. 24, no. 11, pp. 1586–1590, 2017, doi: 10.1109/LSP.2017.2749266.

[38] S. Patil and S. Dc, "An Integrated Technique for Face Sketch Recognition Using DCNN," *Int. J. Innov. Technol. Explor. Eng.*, no. 10, pp. 2278–3075, 2019, doi: 10.35940/ijitee.J9500.0881019.

[39] H. Mendez-Vazquez, F. Becerra-Riera, A. Morales-Gonzalez, L. Lopez-Avila, and M. Tistarelli, "Local deep features for composite face sketch recognition," *2019 7th Int. Work. Biometrics Forensics, IWBF 2019*, 2019, doi: 10.1109/IWBF.2019.8739212.

[40] J. Xu, X. Xue, J. Li, and X. Mao, "Composite Sketch Recognition Using Multi-Scale HOG Features and Semantic Attributes," *Jisuanji Fuzhu Sheji Yu Tuxingxue Xuebao/Journal Comput. Des. Comput. Graph.*, vol. 32, no. 2, pp. 297–304, 2020, doi: 10.3724/SP.J.1089.2020.17915.

[41] X. Tang and X. Wang, "Face Sketch Recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 1, pp. 50–57, 2004, doi:

10.1109/TCSVT.2003.818353.

[42] R. Kokila, M. S. Sannidhan, and A. Bhandary, "A novel approach for matching composite sketches to mugshot photos using the fusion of SIFT and SURF feature descriptor," *2017 Int. Conf. Adv. Comput. Commun. Informatics, ICACCI 2017*, vol. 2017-Janua, pp. 1458–1464, 2017, doi: 10.1109/ICACCI.2017.8126046.

[43] N. O'Mahony *et al.*, "Deep Learning vs. Traditional Computer Vision," *Adv. Intell. Syst. Comput.*, vol. 943, no. Cv, pp. 128–144, 2020, doi: 10.1007/978-3-030-17795-9_10.

[44] W. Wan, Y. Yang, and H. J. Lee, "Generative adversarial learning for detail-preserving face sketch synthesis," *Neurocomputing*, vol. 438, pp. 107–121, 2021, doi: 10.1016/j.neucom.2021.01.050.

[45] Z. Huang, B. Kratzwald, D. P. Paudel, J. Wu, and L. Van Gool, "Face Translation between Images and Videos using Identity-aware CycleGAN," 2017, [Online]. Available: http://arxiv.org/abs/1712.00971.

[46] R. Sawant, A. Shaikh, S. Sabat, and V. Bhole, "Text to Image Generation using GAN," *SSRN Electron. J.*, no. Icicnis, pp. 1–7, 2021, doi: 10.2139/ssrn.3882570.