IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 22, 2022, accepted April 5, 2022, date of publication April 25, 2022, date of current version May 16, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3170037

Open-Source Face Recognition Frameworks: A Review of the Landscape

DAVID WANYONYI[1] **AND TURGAY CELIK**[1,2,3]

[1]School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg 2000, South Africa
[2]Wits Institute of Data Science, University of the Witwatersrand, Johannesburg 2000, South Africa
[3]Faculty of Engineering and Science, University of Agder, 4630 Kristiansand, Norway

Corresponding author: Turgay Celik (celikturgay@gmail.com)

ABSTRACT From holistic low-dimension feature-based segmentation to deep polynomial neural networks, Face Recognition (FR) accuracy has increased dramatically since its early days. The advancement and maturity of open-source FR frameworks have contributed to this trend, influencing many open-source research publications available in the public domain. The availability of modern accelerated computing capabilities through Graphics Process Unit (GPU) technology has played a substantial role in advancing open-source FR capabilities. The evolution and success of the open-source DL algorithms on FR, leveraging GPU technologies, have benefited from open datasets, resulting in many FR open-source implementations. This paper reviews the landscape of open-source FR frameworks, covering components of the FR pipeline across open datasets, face detection, face alignment, face representation, identification and verification, and deployment environments. We also discuss the current challenges and emerging directions in FR research.

INDEX TERMS Face recognition, face detection, face verification, face identification, open-source software, deep learning, review.

I. INTRODUCTION

Face Recognition (FR) constitutes visual identification and/or verification of a person using a face picture. Face verification is "a one-to-one mapping of a given face against a known identity (e.g. is this the person?)" [1]. Face identification is "a one-to-many mapping for a given face against a database of known faces (e.g. who is this person?)." FR technology is available through open-source projects and commercial vendors providing services for biometric identification, access control, video surveillance and contact tracing, for instance, under pandemic outbreaks like COVID-19. In FR, face images can be extracted from still photos or video streams. Behind the state-of-art FR is Deep Learning (DL), which is inspired by how the human brain functions and tries to mimic its behaviour through artificial neural networks. Popular DL architectures include Deep Neural Networks (DNNs), Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs) and Generative Adversarial Networks (GANs). Particularly, CNNs have proved to be very good and popular in solving computer vision problems. As such, they have been widely explored in FR research, starting with initial work done by LeCun *et al.* [2] on handwriting recognition. A typical FR pipeline with Deep Convolutional Neural Networks (DCNNs) is composed of labelling face dataset for training, configuring loss functions, optimizing network parameters during training, and generating a trained model for inferencing.

The availability of Graphics Process Unit (GPU) technology, open datasets, and open-source implementations have played substantial roles in advancing open-source FR capabilities, resulting in a landscape of open-source FR methods. There are several review papers published on FR [3]–[8] comparing methods such as illumination and pose invariance and reviewing various FR methods, including 3D approaches. Except for some of the recent publications [7], [8], most of these predate recent advances in FR and do not cover some of the latest DL methods. Importantly, none of these studies give any specific attention to open-source frameworks on FR, the subject of our research. This study presents a review of the landscape of open-source frameworks in FR with the following contributions:

- We review open FR datasets, cover the different categories of these datasets, and highlight their constraints;
- We provide an analysis of some of the notable open-source face detection and face alignment frameworks

The associate editor coordinating the review of this manuscript and approving it for publication was Seifedine Kadry.

in the last decade that have been instrumental in the evolution of FR;
- We review the evolution of the deep network architectures for FR, emphasizing open-source implementations; and
- We discuss gaps, emerging directions and challenges in open-source FR.

The remainder of this review paper is organized as follows. Section II provides a snapshot of the evolution of FR methods. Section III provides a review of open datasets that have been instrumental in developing open-source FR frameworks and their limitations. Section IV and Section V respectively review the most commonly used face detection and face alignment methods. Section VI covers face representation, identification, and verification methods used in open-source frameworks. Section VII presents research gaps and emerging directions. Finally, Section VIII concludes the paper.

II. FACE RECOGNITION EVOLUTION

FR has gradually evolved from holistic approaches to state-of-the-art DL methods as summarized in Figure 1 [9]. Significant work in FR began in the early 90s into the 2000s when Moghaddam *et al.* [10] published one of the earliest methods using low-dimension feature-based segmentation: Eigenfaces. In their approach, each facial image in the training set is split into multiple small feature sections called Eigenfaces. Facial classification is achieved by linearly projecting the subject image over the feature-space of Eigenfaces and computing the difference in position of each Eigenfaces for the subject image. Other early works using holistic approaches with low-dimension feature-based segmentation are presented in [11], [12]. Fischerfaces by Belhumeur *et al.* [13], one of the outstanding works, improved on Eigenfaces by introducing models that are invariant to lighting direction and facial expression. The major shortcoming of the holistic approach is the inability to handle uncontrolled and unexpected facial changes outside the variations captured in the training dataset. This shortcoming motivated researchers to search for new solutions based on handcrafted local facial feature representations. Many significant studies emerged in the early 2000s, such as a Gabor method for FR using local features [14], a local binary feature-based approach [15], and the high dimension feature-based compression [16]. These advances mainly focused on high dimension feature representation and produced better results in FR than holistic approaches. However, their reliance on handcrafted features still limited their effectiveness in practical, diverse, complex environments of FR.

The limitations of handcrafted features resulted in the emergence of learning-based approaches [17], [18]. These methods, however, underperform when faced with complex variations in facial appearance outside the variations captured in the training data. Faced with these challenges, the research community increased efforts to address the fundamental problem of poor performance of FR under non-linear variations in facial appearances and expression. Most of the research at this point focused on methods such as feature coding with histograms, dictionary atoms distribution, feature transformation and use of local descriptors. Still, these did little to address the problem.

DL, addressing the issues mentioned above, have dominated the FR research in the last decade. DL imitates how the human brain processes data, often implemented as layered networks in CNNs as illustrated in Figure 2. CNNs gained attention on image recognition tasks when a then-impressive accuracy was achieved by AlexNet [19] at the ImageNet competition in 2012. In 2014, a then never-before FR accuracy of 97.35% was achieved by DeepFace [20] on the Labeled Faces in the Wild (LFW) dataset, coming quite close to the human performance of 97.53%. Not long after this, human performance was surpassed on the LFW dataset with DL-based methods such as FaceNet [21] with an FR accuracy of over 99%. Research in CNN in 2015 saw the emergence of new network blocks such as gated skip connections and cross-layer channel connectivity to improve the convergence rate and performance of deep CNN architectures [22], [23]. The year 2016 witnessed that stacking multiple transformations in parallel as well as depth-wise improved the learning representation for complex problems [24]. From 2017 onwards, we see a focus on improving the network representation through special blocks. These blocks enhance network performance at any layer in CNN architecture [25]. Recent works [26]–[28] focus on the use of generic blocks that use feature maps to control information assignment using attention. In 2018, Khan *et al.* [29] introduced the concept of channel boosting through transfer learning. We can see another demonstration of transfer learning in 2019 by Kolesnikov *et al.* [30]. In 2020, we saw the emergence of one of the top-performing models that use polynomial neural networks [31]. Table 1 summarizes the evolution chronologically, citing some of the notable works in FR.

III. OPEN DATASETS

Accessibility of open datasets has played a crucial role in advancing deep learning-based open-source FR frameworks. The datasets have appeared in different forms, ranging from those assembled from diverse sources to single sources, from large scale to small scale. Others range from real-world unconstrained conditions to laboratory-controlled, from image-only content to videos. This section reviews the most commonly used datasets that have been invaluable to FR researchers for training and evaluating the performance of the open-source frameworks and their limitations.

A. OVERVIEW

1) LABELED FACES IN THE WILD (LFW)

Labeled Faces in the Wild (LFW) [59] is a well-known benchmark dataset for studying FR in unconstrained environments, popularized with the advancement of CNNs in FR tasks. The dataset consists of 13,233 face images with

TABLE 1. The evolution of deep learning-based FR methods and their performances.

Name	Loss Function	Architecture	Dataset	Year	LFW Accuracy%
DeepFace[32]	Sofmax	Alexnet	Facebook (4.4M,4K)	2014	97.35
DeepID2[33]	Contrastive Loss	Alexnet	CelebFaces+ (0.2M,10K)	2014	99.15
DeepID3[34]	Contrastive Loss	VGGNet-10	CelebFaces+ (0.2M,10K)	2015	99.53
FaceNet[21]	Triplet Loss	GoogleNet-14	Google (500M,10M)	2015	99.63
Baidu[35]	Triplet Loss	CNN-9	Baidu (1.2M,18K)	2015	99.77
VGGFace[20]	Triplet Loss	VGGNet-16	Google (500M,10M)	2015	98.95
Light-CNN[36]	Softmax	Light CNN	MS-Celeb-1M (8.4M,100K)	2015	98.8
Center-Loss[37]	Center Loss	Lenet	CASIA-WebFace, CACD2000,Celebrity+ (0.7M,17K)	2016	99.28
L-Softmax	L-Softmax	VGGNet-18	CASIA-WebFace (0.49M,10K)	2016	98.71
Range Loss[38]	Range Loss	VGGNet-16	MS-Celeb-1M, CASIA-WebFace(5M,100K)	2016	99.52
L2-Softmax[39]	L2-Softmax	ResNet-27	MS-Celeb-1M (3.7M,58K)	2017	99.78
Normface[40]	Contrastive Loss	GoogleNet-14	Google (500M,10M)	2017	99.19
CoCo Loss[41]	Coco Loss	-	MS-Celeb-1M (3M,80K)	2017	99.86
vMF[42] Loss	vMF Loss	ResNet-27	MS-Celeb-1M (4.6M,60K)	2017	99.58
Marginal Loss[43]	Marginal Loss	ResNet-27	MMS-Celeb-1M (4M,80K)	2017	99.48
SphereFace[44]	A-softmax	ResNet-64	CASIA-WebFace (0.49M,10K)	2017	99.42
CCL[45]	Center Invariant Loss	ResNet-27	CASIA-WebFace (0.49M,10K)	2018	99.12
AMS Loss[46]	AMS Loss	ResNet-20	CASIA-WebFace (0.49M,10K)	2018	99.12
Cosface[47]	Cosface	ResNet-64	CASIA-WebFace (0.49M,10K)	2018	99.33
Arcface[32]	Arcface	ResNet-100	MS-Celeb-1M (3.8M,85K)	2018	99.83
Ring loss[48]	Ring Loss	ResNet-64	MS-Celeb-1M (3.5M,31K)	2018	99.83
Ring loss[48]	Ring Loss	ResNet-64	MS-Celeb-1M (3.5M,31K)	2018	99.83
AdaCos[49]	AdaCos	ResNet50	CASIA-WebFace MS1M	2019	99.71
VarGfacenet[50]	Angular Arcface	ResNet	MS1M	2019	99.683
Curricularface[51]	Curricularface	ResNet50 RestNet100	CASIA-WebFace MS1MV2	2020	99.80
GroupFace[52]	Cosface Arcface	ResNet	MS-Celeb-1M	2020	99.85
BroadFace[53]	Arcface	RestNet100	MSCeleb-1M	2020	99.85
MultiFace[54]	Softmax Arcface Cosface	ResNet MobileNet	MS1MV2 CASIA-Webface	2021	99.74
LarNet[55]	ArcFace	ResNet	MS1MV2 CASIA-Webface	2021	99.74
Face.evoLVe[56]	Softmax, Focal etc	RestNet DenseNet	WebFace260M	2021	99.78
MixFaceNets[57]	ArcFace	MixNet	MS1MV2	2021	99.83
MagFace[58]	MagFace	RestNet	MS1MV2	2021	99.83

5,749 individuals. It provides labeled faces that exhibit appearances of varying conditions in daily life, ranging from pose, lighting, background, occlusions, race, and gender. The images are presented as JPEG files of 250 × 250, mostly in color, with a few exceptions in grayscale. The LWF dataset is available on LFW Website [60].

2) CUHK FACE SKETCH FERET DATABASE

Based on the FERET dataset, CUHK Face Sketch FERET (CUFSF) dataset was introduced for research on face sketch synthesis and face sketch recognition by Zhang *et al.* [61]. The dataset consists of paired images of photos and the corresponding sketches. The dataset features one image per subject for 1,194 images with varying lighting conditions. CUFSF is featured by Fu *et al.* [62] for heterogeneous FR.

3) REPLAY-ATTACK

Replay Attack dataset [63] consists of 1,300 video clips that feature photo and video attack attempts on 50 identities, all under different lighting conditions. The videos are generated by recording an actual person attempting to access a laptop using a webcam or using a video recording of the same

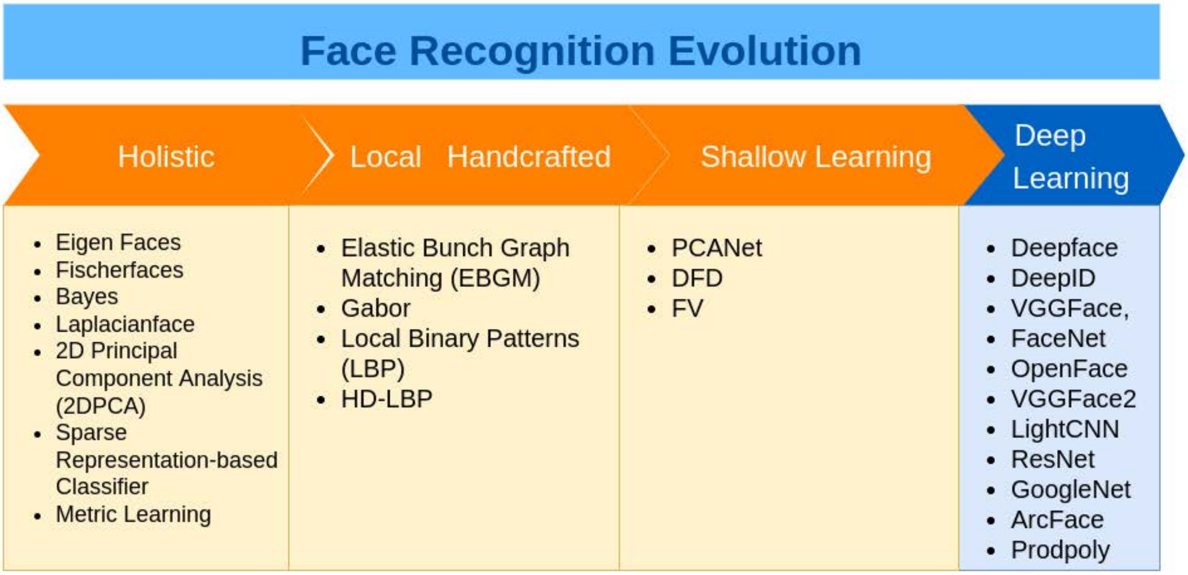

FIGURE 1. In the early days, popular FR methods used low-dimension holistic approaches. This evolved into handcrafted local descriptors and local feature learning. Then deep learning revolutionised FR with emergence of frameworks such as DeepFace [20].

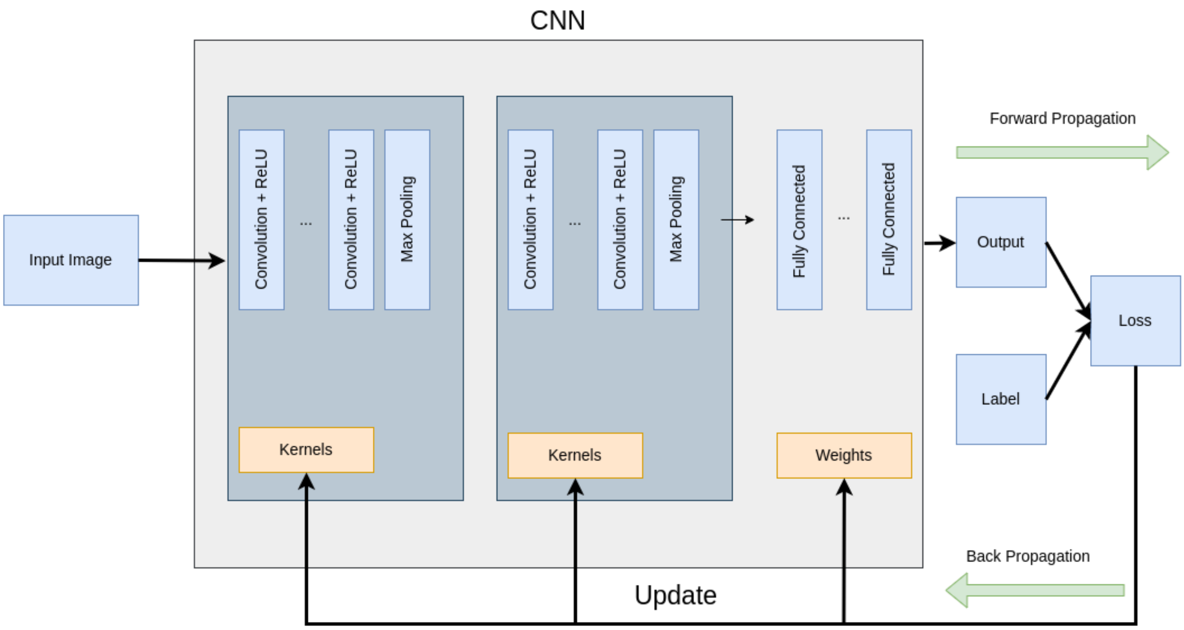

FIGURE 2. Typical layered view of a Convolutional Neural Network during training and inference processes. During the training process, the kernel parameters of CNN and weights of fully connected layers are updated based on backpropagating gradients of loss function at the end of the classifier. During the inference i.e forward propagation, the updated kernels and weights are used for inference.

person for a short time. Yang *et al.* [64] used the dataset for CNN-based face anti-spoofing.

4) PARTIAL-REID

Partial Person Re-Identification dataset [65] consists of 600 images of 60 subjects, each subject having 5 partial images and 5 full-body images, featuring different backgrounds, viewpoints, and varying high levels of occlusions. The dataset aims to address cases where a partial view of a person is available and can be used to match with other images from different viewpoints for identification. The dataset is used alongside others in experimental biometric recognition on partially captured targets [66], [67]. The dataset is available on Google Drive [68].

5) CASIA-MSDF

CASIA-MSDF [69] is a dataset for face anti-spoofing consisting of 50 subjects with 12 videos. The videos were collected under different lighting conditions and varying resolutions. Zhang *et al.* [69] used 600 video recordings from the dataset consisting of 240 videos from 20 identities for training and 360 videos from 30 identities for testing. Three different spoof attacks were used for experimentation: reply, print, warp, and cut. CASIA-MSDF is one of the datasets used in the Cross Central Difference Convolutions paper published by Yu *et al.* [70].

6) MORPH

The MORPH dataset [71] contains 55,134 facial images from 13,617 subjects, first introduced for academic and development use. The commercial version of the dataset consists of more than 400,000 images from approximately 70,000 subjects. The MORPH dataset is annotated across age, weight, gender, height, and eye coordinates. The dataset is available upon license application to the University of North Carolina, Wilmington. The dataset has been used in studies, including face quality assessment and estimation work [72] and a cross-age face synthesis framework [73].

7) CASIA-WebFace

CASIA-WebFace [74] is a large open facial dataset consisting of annotated images of 10,575 unique people with over 494,414 images in total. CASIA-WebFace was collected by crawling IMDb [75] website and extracting the identities of celebrities born between 1940 to 2014. A multi-view face detector processed the extracted images. FaceNet has an inferencing model trained on the CASIA-WebFace dataset and achieves an accuracy of 99.05% on the LFW benchmark. According ArcFace [76] model, there exists evidence of racial bias in the CASIA-WebFace. It was found that Caucasian distribution has a margin that stands out from the other races implying a higher probability of recognition errors in non-Caucasian subjects.

8) EXTENDED YALE B

Extended Yale B dataset [77] consists of 16,128 images of 28 identities under 9 poses and 64 illumination conditions. Georghiades *et al.* [77] used the dataset to closely crop images for generative appearance under varying lighting and viewpoint. Yale Face Database B is freely available for research purposes and is available on Yale Database [78] and [79].

9) ADIENCE

Adience dataset [80] was introduced in 2014 and is aimed at studying age and gender recognition, consisting of 26,580 face photos of 2,284 identifies labeled in binary gender and with a label from 8 different age groups of five splits. The dataset is available on Yale Database [79] through a simple request via an online form.

10) FaceScrub

FaceScrub [81] was released in 2014, consisting of over 141,000 faces of 695 public figures collected by scrubbing the web. FaceScrub is one of the large datasets used in OpenFace's model, producing the best performing model against LFW at a score of 92.92% $\pm$ 0.0134.

11) WIDER FACE

WIDER FACE [82] is an open face detection dataset first published in 2015. It has 32,000 images of 393,000 labeled faces with many variations. Face-API [83] face detection model named 'SSD Mobilenet V1' is trained on the WIDER FACE dataset.

12) UMDFaces

UMDFaces was introduced in [84] in 2016, consisting of two batches of still images featuring 367,88 annotated faces of 8,277 identities and an annotated set of video frames of over 3.7 million of over 22,000 videos of 3100 identities. The annotations cover estimated poses across yaw, pitch, roll, 21-keypoints, and gender information. The dataset is available on UMDFaces site [85].

13) MS-CELEB-1M

MS-Celeb-1M [86] dataset, appearing in 2016, contains 10M images, becoming the largest open dataset at the time. The dataset was used to recognize 1M celebrities from their face images. Insightface [32], [87] uses MS-Celeb-1M as one of the train datasets, available on Github [88].

14) VGGFace2

VGGFace2 [89] is a large public dataset published in 2017 containing 3.3M images of 9131 identities, averaging about 362.6 images per identity. The dataset has large diversity across illumination and other attributes like career specialization, ethnicity, and age and contains images downloaded using the Google Image Search. FaceNet's top-known model is trained using the VGGFace dataset, achieving an accuracy of 99.65% on the LFW benchmark.

15) RACIAL FACE IN-THE-WILD

Racial Face In-The-Wild (RFW) is a dataset aimed at research on racial bias in FR, introduced by Wang *et al.* [90]. Featuring 40,607 images from a total of 3000 subjects, it is available in 4 sets: African group consisting of 10,415 images from 2,995 subjects, Asian group consisting of 9,688 images from 2,492 subjects, Indian group consisting of 10,308 images of 2,984 subjects, and Caucasian group consisting of 10,196 images of 2,959 subjects. RFW is one of the datasets in FaceX-Zoo open-source project published in [91] and in [92], a study on mitigation of FR bias through the use of a false positive rate penalty loss.

16) IMDb-FACE

Consisting of approximately 1.7 million face pictures from 59,000 identities, IMDb-Face is one of the large-scale controlled datasets introduced in [93]. The source of the images is [75]. IMDb has been used in many face-recognition projects, including works published in [94]–[96].

17) IJB-A

IARPA Janu Benchmark A (IJB-A) [97] is a combination of 5,712 images and 2,085 videos collected from 500 identities that aims to provide additional challenging tasks to FR across the expression, pose, illumination, occlusion, and resolution. IJB-A is used in various research studies such as face clustering [98], 3D face-rotation framework for unsupervised photo-realistic synthesizing [99] and discriminative 3D morphing [100].

18) ARPA JANU BENCHMARK B (IJB-B)

IJB-B [101] is a template-based dataset of unconstrained images and videos collected from the Internet across pose, image quality, and illumination. The dataset has three verification and identification protocols: 1-to-1 template face verification protocol, 1-to-N open-set face identification template-based protocol, and 1-to-N video face identification protocol. The dataset consists of 1845 subjects with 17,754 images.

19) IARPA JANUS BENCHMARK-C

IJB-C [102] is a video-based FR dataset that extends the IJB-A dataset and consists of approximately 11,000 face videos, 138,000 face images, and 10,000 non-face images. IJB-C is used as a benchmark in many research works, including work published on optimization of softmax-based on training with large datasets and in [103] for large-scale noise management in training datasets.

20) MeGlass

MeGlass [104] contains face images designed for eyeglass FR. The dataset consists of 47,817 images collected from 1,710 different subjects, with each having a minimum of two faces with an eyeglass and two without. The dataset can be accessed via links at [105]. MeGlass dataset is used in a paper published by Guo *et al.* [106] in a study for generalized FR on unseen domains and a study by [107] on gaze preservation with eyeglasses.

21) QMUL-SurvFace

QMUL-SurvFace [108] contains 463,507 images from 15,573 subjects from unconstrained real-world surveillance scenes across time-wide location. The dataset can be found at [109] from which test codes and evaluation protocols are published. The dataset is licensed for research purposes.

22) iQIYI-VID

iQIYI-VID is a large dataset introduced by Liu *et al.* [110] consisting of 600,000 video clips from 5,000 celebrities extracted from diverse online videos that include movies, TV series, news broadcasts, and various shows. The dataset is used for multi-modal person identification which involves jointly utilizing face, head, body, and voice features to identify a person. The dataset is also used in a paper published by [111] for large-scale multi-modal person identification under an unconstrained environment.

23) DIVERSITY IN FACE (DiF)

Released by IBM research in 2019, Diversity in Face (DiF) [112] is one of the latest (as of 2020) large open datasets with annotations of one million publicly available face images. It is released for the research community to help researchers study biases in facial diversity across race, ethnicity, culture, and geography. A ten-facial coding scheme with human-understandable facial features was used to generate the annotations.

24) FairFace

FairFace [113] consists of a racially balanced dataset of 108,501 images from seven race groupings: White, Black, Indian, Latino, South East Asian, and Middle Eastern. The dataset features age, race, and gender labels and was collected from the YFCC-100M Flikr dataset.

25) DISGUISED FACES IN THE WILD

Disguised Faces in the Wild (DWF) [113] contains approximately 1K images from 1K identities in different guises. The images were collected from the Internet, exhibiting an unconstrained real-world environment. The paper is featured in the research published in [114] on active learning on disguised faces with adversarial noise.

26) FACEBOOK DATA

In 2021 Facebook AI released a human-annotated dataset consisting of 45,000 videos of 3011 identities for researchers to use to study fairness across a diverse set of ages, genders, and apparent skin tone [115].

B. CHALLENGES

Open datasets consist of photos of people from the Internet, many of them celebrities. This does not accurately reflect daily life realities. When DL algorithms use these datasets for training, the bias reduces the accuracy of DL algorithms outside lab environments.

Depending on source and cleaning methods used, the datasets can also contain noise, affecting the accuracy of trained models. A noticeable level of noise was found in MS-Celeb-1M [32]. Wang *et al.* [93] show that when a dataset contains noise, it severely affects the performance of the trained models, especially for large datasets. They equally show how the degradation caused by noise exponentially increases when there are many classes, especially when used with Softmax classification. Consequently, datasets such as VGGFace2 and MS-Celeb-1M, containing data from celebrities who are mostly young, facially beautiful, and with makeup, are biased in age and facial appearance. Using pre-trained models from these datasets directly can lead to poor performance in the audience that differs from the training dataset.

Klare *et al.* [116] also report demographic bias in ethnicity for black females. The FR systems find it harder to classify the black female faces accurately. Another challenge with datasets is cross-pose face verification and recognition. There have been many efforts in addressing this issue. For instance, Zhao *et al.* [117] published a method for learning pose-invariant faces by jointly learning frontal images and posed versions to produce high-quality frontal reference images, and Chen *et al.* [118] introduced feature extraction using multi-view subspace for improved accuracy. However, these studies do not entirely solve the challenge of cross-pose FR. As discussed by Sengupta *et al.* [119], algorithms still degrade as much as 10% when faced with profile view verification, showing just how hard the cross-pose is still a challenging problem, and public datasets are not yet diverse enough to provide satisfactory results. Some of the datasets that have focused on cross-pose and cross-age domains include CALFW [120], MORPH [71]. Photo-sketch dataset CUFS [62] has been used to address the gap in matching faces across diverse environments. Other common challenges in open datasets include the distribution of images per identity, as evidenced in MS-Celeb-1M [86] where the number of images per person is small.

While researchers have made an effort to create open datasets that reflect diverse real-life situations, the challenges still persist due to unpredictable, complex nonlinear facial appearance.

IV. FACE DETECTION

In FR, face detection is one of the fundamental steps in the recognition pipeline. There are many existing studies on face detection, ranging from keypoint annotation [87], [122], [123] to data augmentation methods [124]. Inherently, FR is built on the essentials of object detection, sharing the same history as generic object detection. Before DL, face detection was based on handcrafted features using methods like Haar-like features [125]. This rapidly evolved into more complex approaches that focused on overcoming variability problems across pose, expression, illumination, occlusion, etc. WIDER FACE [82], one of the most challenging datasets in face detection, has played an important role in accelerating the development of newer methods of face detection, spawning works such as PyramidAchors [124], Dual Shot Face Detector (DSFD) [126], RetinaFace [87] and one of the latest face detectors as of 2021, TinaFace [127].

Face detection methods typically use two approaches [121] as shown in Figure 3: feature-based and image-based approaches. Feature-based approaches extract features from an image and match these against a database of known face features. On the other hand, image-based approaches compare training and testing images for the best match. Please refer to [121] for a detailed literature survey on face detection methods. The rest of this section explores some notable works on face detection that are either open-sourced or have been used in open-source FR frameworks.

A. OVERVIEW

1) TinaFace

TinaFace [127] is an open-source face detection framework using a state-of-the-art average precision test time augmentation. In their paper, Zhu *et al.* [127] show how face detection can be achieved through implementation of methods based on generic face detection. TinaFace uses ResNet [128] for the backbone architecture. It is reported that TinaFace models trained on WIDER FACE [82] achieve a score of 92.4% average precision. The source code for the implementation is available on Github [129]. TinaFace consists of a seemingly simple yet powerful architecture consisting of a feature extractor, an inception block, classification, regression, and Intersection over Union (IoU) [130] module.

2) AInnoFace

AInnoFace [131] is another top-performing face detection framework with state-of-the-art performance on the challenging face detection benchmark WIDER FACE with average precision scores of Easy at 97.0%, Medium at 96.1% and Hard at 91.8% on validation subset. Based on the RetinaNet, it is reported the framework uses IoU loss function [132] for regression and applies a two-step classification and regression for detection. During training, the data-anchor-sampling method for data augmentation [133] is used in training with max-out operation for classification [134]. Although it appears AInnoFace does not directly provide the full source code, it uses open-source code for testing available on Github [135].

3) MTCNN

Multi-Task Cascaded Convolutional Neural Network (MTCNN) [136] is a state-of-the-art face detection framework for detecting faces and facial keypoint locations using a coarse-to-fine approach. A TensorFlow implementation by David Sandberg is available on Github [137] and another excellent implementation by Ivan based on Keras in Python3.4+ is available on Github [138]. David's implementation shows the capabilities of the MTCNN and is provided with pre-trained face detection models.

4) RetinaFaceMask

RetinaFaceMask [76] is an open-source face detector that appeared in 2020 and is hosted on GitHub [139]. The source code is available in Python. It can detect face masks, one of the first dedicated face mask detectors in the wake of the COVID-19 outbreak. RetinaFaceMask architecture is based on a feature pyramid network enhanced with an algorithm for attention detection. Transfer learning is used to train the representations on face masks dataset using baseline pre-trained models from ImageNet and WIDER FACE increasing detection accuracy by 3-4% against the backbone architecture results. The project is licensed under the MIT License terms.

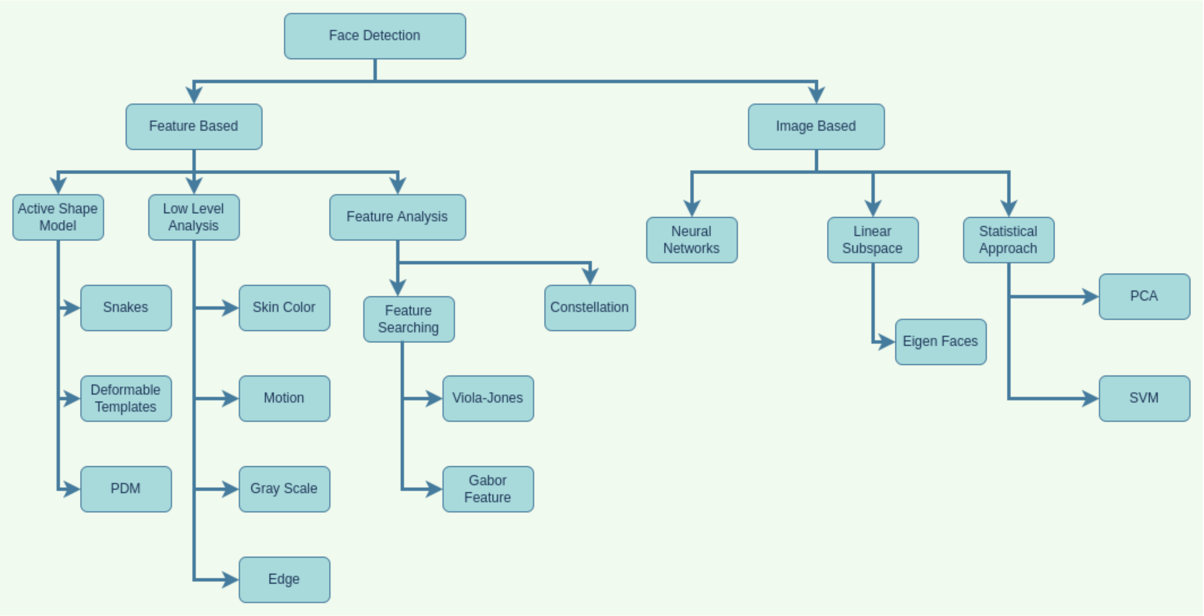

FIGURE 3. Face detection methods broadly fall into one of two categories: feature-based and image-based approaches [121]. Feature-based approaches focus on active shape models, low level analyses, and feature analyses, while image-based approaches focus on neural networks, linear subspace methods, and statistical approaches.

5) PyTorch-JAANet

PyTorch-JAANet is a Python project for facial action unit detection and alignment [140], also hosted on GitHub [141]. It appeared in 2020 and used an adaptive attention learning module to enhance local features' extraction for integration into the global feature map for action unit detection. PyTorch-JAANet achieves competitive performance against the state-of-art benchmarks on partial occlusion and non-frontal faces. The code for the project runs under PyTorch 1.1.0 and Python 3.5. BP4D [142] was used for training. There is no licensing information available on the code repository.

6) KPNet

KPNet [143] is a lightweight face detection and alignment framework that detects facial keypoints using a bottom-up approach in which facial keypoints are detected from a low-resolution image without the use of anchor boxes. The face bounding boxes are then inferred from the key points. Two backbone CNN architectures are used: the hourglass network and DRNet for fast inferencing speed. KPNet was implemented using PyTorch, supporting a 68-point semi-frontal and 39-point profile keypoint detection. The source code is available on GitHub [144].

7) YOLOFace

YOLOFace is a face detection library for DL, implementing the YOLOv3 [145] algorithm. The project first appeared on GitHub [146] in 2018. It consists of an OpenCV module that can be used for face detection tasks using pre-trained DL models from popular frameworks such as TensorFlow, Darknet, Caffe, and Torch. The implementation requires a Python 3.6 environment in Ubuntu, with dependencies including TensorFlow. The pre-trained models were trained on the WIDER FACE dataset. The project is provided with an MIT License.

FIGURE 4. Facial appearance changes posing challenges to face detection.

8) RetinaFace

RetinaFace [87] is a face-detection library that uses a single-shot multi-level face localization approach that combines bounding box detection and localization of keypoints. The annotation is achieved by combining manual and semi-automatic annotation on the WIDER FACE dataset. RetinaFace source is available on Github at [147]. The project is available under the MIT License.

9) SeetaFace

SeetaFace is a CPU-based open-source face detection engine for real-time multi-view face detection [148]. It is written in C++. Face detection is implemented as a real-time component using a funnel-structured (FuSt) cascaded schema. The released code currently works only on Windows.

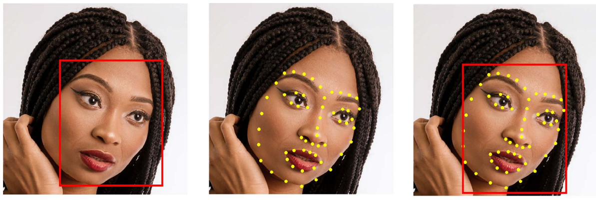

FIGURE 5. Illustration of a typical face alignment process comprising face detection, face keypoints mapping, and finally face alignment.

B. CHALLENGES

Face detection encounters challenges ranging from expressions, poor illumination, skin tone and color, background interference, and even crowded faces in a viewpoint. These challenges reduce detection accuracy and decrease the rate of detection. See Figure 4 for a view of some facial appearances that create challenges for face detection as summarized below:

- **Expression**: Humans are capable of odd expressions that alter their facial appearance making it challenging for accurate face detection.
- **Illumination**: Irregular and poor illumination can change facial appearance reducing the detection accuracy.
- **Viewing distance**: If the subjects' faces are too far from a camera in an uncontrolled environment, the detection rate and accuracy may be reduced.
- **Skin color**: Certain skin colors can be a challenge for face detectors depending on the algorithm used.
- **Occlusion**: When parts of a face are hidden by an object such as a hand, glasses, hats, etc., face detection accuracy may be reduced.
- **Resolution**: An image with poor resolution may challenge face detection, reducing accuracy.
- **Pose and Orientation**: The pose of a face at an angle will hide some sections of a facial image, creating a challenge for face detection systems. Barra *et al.* [149] present a pose estimation method using an adaptation of quad-tree representation of facial keypoints that interestingly does not need the use of neural networks. Starting from face detection and alignment, the authors use quad-tree decomposition resulting in facial representation that is used for pose estimation. Given a sequence of face frames in a video, the authors published a method for selecting best fitting pose that would increase the accuracy of recognition. We acknowledge the positive contribution.
- **Crowded faces in view point**: A face detection system may encounter challenges in accurately detecting crowded faces from a viewpoint.
- **Background interference**: Many visible objects in the background can be a challenge for a face detection system, reducing the rate and accuracy of detection.

V. FACE ALIGNMENT

Face Alignment is another crucial component in the FR pipeline. It localizes facial keypoints on a given facial image and aligns it with a standard face template.

A. OVERVIEW

At the heart of a typical alignment process is facial fiducial points, which refer to predefined facial keypoints on a face typically located around the nose, mouth, and the chin, as illustrated in Figure 5. FR algorithms predominantly use face alignment to reduce the impact of pose variations by affine-warping the input facial image to a standard frontal face template based on facial fiducial points. Face Alignment is also used as a pre-processing method to detect facial accessories such as glasses and model facial deformations. For instance, Kumar *et al.* [150] used 6 facial fiducial points to localize and compute smiles used for face verification in unconstrained conditions.

B. FACE ALIGNMENT METHODS

Face alignment research has progressed over the last years with increasing success. A typical face alignment process aims at progressively aligning a standard face shape template to an input facial image, searching for predefined facial fiducial points on the input. Typically, this commences with an initial coarse shape refined iteratively through several steps and stops when convergence criteria are satisfied. As the search progresses, information on facial appearance and the standard face shape model is jointly utilized to locate facial fiducial points. There are several excellent review papers [151]–[154] covering the progress in face alignment from traditional methods to recent deep learning-based methods. The most popular approaches to face alignment are categorized as either generative or discriminative. Generative methods typically approach a face alignment problem as an optimization task that generates shape and appearance parameters that fit the input facial image. Some of the generative methods include regression-based fitting, gradient descent-based fitting, and part-based generative models [155]–[157]. On the other hand, discriminative methods take each facial fiducial point and use an independent local detector to determine the target location from a facial appearance directly. Some of the discriminative methods include constrained local models (PCA shape model, exemplar shape model), constrained local regression, deformable part models, ensemble regression-voting, cascaded regression (two-level boosted regression, cascaded linear regression), deep neural networks, 3D shape regression and dense 3D model fitting [158]–[161].

One of the mainstream approaches to face localization is the use of heatmap regression [162], [163]. AdaptiveWingLoss is a Pytorch Implementation of a heatmap regression published by Wang *et al.* [164]. It appeared on GitHub [165] in 2019. It is implemented in Python 3.5.7 and Pytorch 1.3. AdaptiveWingLoss is evaluated on face-alignment with

FIGURE 6. Facial appearance changes across expression, illumination, occlusion and pose.

pre-trained models provided. It is trained on datasets 300W [166] and WFLW [167]. The repository does not appear to have any updates since its publication. The code is licensed under Apache License 2.0. Deep Alignment Network (DAN) is a multi-staged face alignment framework that uses keypoint heatmaps based on deep neural network architecture, introduced in [163]. DAN begins with an initial coarse estimate of the face, with subsequent iterations that refine the results. A single-stage deep neural network is used for feature extraction and regression at each iteration through a sequential staged training. During experimentation, DAN was trained on the 300W dataset, achieving an average failure rate of 1.16 with data augmentation.

There is considerable attention on unconstrained face alignment fueled by changes in facial appearance in extreme poses, exaggerated expressions, and heavy occlusions. An implementation geared towards improving face alignment under unconstrained conditions that uses 3D facial keypoint localization is published by Guo *et al.* [168] using U-Net as the backbone CNN architecture. SeetaFace Alignment [159] cascades some Stacked Auto-encoder Networks (SANs) to improve accuracy on the keypoint detection progressively. The keypoint detection framework is trained on 23,000 images and detects 5 facial fiducial keypoints: two eyes, nose tip, and mouth corners. Currently, it only runs on a CPU and is only tested on Windows. The source code is hosted on GitGub [88] under a BSD2 license.

C. CHALLENGES

While the face alignment methods can perform well under controlled facial deformations and environments [150], [169], accuracy is significantly affected negatively by variations in facial expressions and environments. The common challenges include pose, occlusion, expression, and illumination. Pose influences face appearance and local facial features. A face appearance can change depending on whether the pose is frontal, upside, profile, or down, leading to occlusion. Facial expressions also change the appearance of facial features such as eyes and mouth, e.g., deformation of the mouth caused by laughing. Lighting is also known to alter the appearance of a facial image, depending on the intensity, spectra, or source distribution. Figure 6 shows such different facial appearances across expression, illumination, occlusion, and pose.

VI. FACE REPRESENTATION, IDENTIFICATION, AND VERIFICATION

This section reviews open-source frameworks for face representation, identification, and verification. Our review looks at the architecture of each of the libraries/frameworks, the source code location, programming environments, datasets used, and the type of open license available.

A. FaceNet

FaceNet [21] is a TensorFlow face recognizer developed by Google researchers in 2015 and achieved then top accuracy score of 99.6% on the LFW dataset. FaceNet has open-source implementations enabling it to be used by many people and has pre-trained models for extracting high-quality face embeddings. FaceNet uses a deep convolutional network to produce compact 128-D embeddings from face images mapped from euclidean space. The framework implements Zeiler&Fergus [170] (220×220 input image size) and Inception [22] (224×224 input image size) as backbone architectures. FaceNet training uses the Large Margin Nearest Neighbor (LMNN) [171] triplet-based loss function consisting of two matching face images and a third non-matching thumbnail. It groups similar vectors for the same identity to enhance similarity and pushes away vectors for different identities. David Sandberg provides a mature FaceNet implementation in TensorFlow, available on GitHub [172]. The repository provides source code and pretrained models, Inception Resnet (V1), trained using the datasets CASIA-Webface and VGGFace2 and available under MIT license. Pretained models were evaluated under the standard protocol for unrestricted, labeled outside data reporting a mean classification accuracy on the LFW academic test of 99.63% $+-$ 0.09 accuracy. Classification accuracy of 95.12% $+-$0.39 was reported on the Youtube Faces DB. Hiroki Taniai [173] also provides a notable FaceNet implementation with a pre-trained Keras model that can be readily used. It is provided with utilities that can convert the Inception ResNet model from TensorFlow to Keras. It was trained on MS-Celeb-1M dataset for input 160×160 color images.

B. InsightFace

InsightFace [32], [168], [174]–[178] is presented as a framework for analyzing faces both in 2D space and 3D space, available on Github [88], for the most part based on MXNet (version 1.2 – 1.6 at the time of writing this review paper), with a number of third party implementations across TensorFlow and recently Keras. InsightFace [32] uses deep convolutions neural network to implement a loss function with arc-cosine functions for discriminative FR. Normalized 112×112 face images are generated during training based on five facial fiducial points in the prepossessing phase. ResNet50 and ResNet100 are used as backbone CNN architectures, with batch normalization used to generate the final 512-D embedding feature. InsightFace implementations also include backbone architectures such as Dense and MobileNet. The loss functions used are Softmax, SphereFace, CosineFace, ArcFace, and Triplet (Euclidean/Angular) Loss. An original implementation for InsightFace is available on GitHub at [173]. InsightFace's source code is available with an MIT License that has no limitations for use in academic research and commercial projects. It is reported that InsightFace is best trained with GPU servers with Python 3.x and MXNet installed at the time of writing. An accuracy of 99.82% is reported against the LFW benchmark and an accuracy of 98.02 on Youtube Faces. A 94.2% accuracy is reported against IJB-B and 95.6% against IJB-C. InsightFace provides a simple API and sample scripts for testing FR and can easily be used to implement a full-production FR system.

C. IVCLab

IVCLab [180] presents an adaptive threshold-based implementation of face verification. Chou *et al.* [180] employ MTCNN for face detection and L2 normalization for training. The most interesting aspect of this project is the principle of adaptive threshold setting, which can be useful for adaptively selecting optimized thresholds for FR. Evaluation is carried out on LFW (76.46%), Aidance (84.30%), and Color Feret (83.79%). Source code is implemented in Python and uses pre-trained FaceNet models for generating face embeddings. The code is hosted on GitHub [181]. However, the code is only licensed for research purposes and cannot be used for commercial ventures.

D. OpenBR

OpenBR [182] is an open-source framework designed to facilitate rapid algorithm prototyping and includes pre-packaged algorithms for FR, age estimation, and gender estimation. It includes face detection, normalization, representation, feature extraction, and face matching modules. It is built using C/C++ on top of Qt 5.4.1, OpenCV, and Eigen. Integration is achievable via C/C++ API. OpenBR is trained on various data sources. These are FERET [183] and CASIA Web. FR is implemented using the Spectrally Sampled Structural Subspaces Features (4SF) algorithm [184]. Face detection is implemented as a custom wrap on OpenCV 2.4 using the Viola-Jones object detector for frontal face detection with Cascade, with a C++ port of ASEF [185] for eye detection. Normalization uses faces cropped to 128×128 and employs Gaussian methods in the prepossessing steps. LBP [15] and SIFT [186] descriptors are used for face representation with 8×8 pixel local binary patterns with a sliding window of 6-pixel step with PCA decomposition. Feature matching is implemented on a custom algorithm (Lbyte1 distance) based on the L1 distance for vectors. The project's website [182] lists a stable version 1.1.0 released in 2015. However, the source code repository on github [187] shows some sporadic activity, the latest commit appearing in 2021 at the time of writing. OpenBR framework includes clear instructions for building the source code across Linux, OSX, and Windows. OpenBR is licensed under Apache License, Version 2.0.

E. DeepFace–1

DeepFace [188] was created by a Facebook research group for FR. At its release, it achieved an impressive 97.35% accuracy on LFW, comparable to a human accuracy of 97.53%. DeepFace has 8 convolutional neural network layers, a max-pooling layer, a locally connected layer, and a fully connected layer. DeepFace consists of a massive 137,774,071 trainable parameters, only second to VGGFace. The original DeepFace model was trained on the Social Facebook Classification (SFC) dataset comprising 4.4 million photos of 4030 identities. Swarup Gosh [189] also trained models on VGGFace2 dataset. DeepFace model works with face images of size 152×152 resulting in a 4096-dimensional vector. Therefore, an embedding vector generated by DeepFace is of size 4092, and Euclidean distance in L2 can be used to determine the similarity between two vectors. The implementations we found most mature are built-in Python using Keras. One of the notable implementations using DeepFace is deepface [190] discussed in the next section, titled DeepFace-2. A TensorFlow/Keras implementation is available on Github [191]. At the time of writing, this particular implementation uses OpenCV-Python 3.4.4, TensorFlow 1.9.0, and Keras 2.2.0. There are also implementations for Realtime detection [192].

F. Deepface–2

Deepface [190], not to be confused with Facebook's DeepFace, is a relatively recent light-weight hybrid FR framework written in Python in Keras and TensorFlow with additional facial attribute analysis capabilities for emotion, gender, and age. It wraps models from VGG-Face, FaceNet, OpenFace, FaceBook's DeepFace, DeepID, and Dlib. The recognition pipeline follows the classic face detection, alignment, representation, and verification stages. Face detection is achieved through a combination of OpenCV's Haar cascade, Single Shot Multibox Detector (SSD), Dlib, and MTCNN. Deepface can detect faces from still pictures as well as real-time videos. The default FR model is VGG-Face, and verification uses cosine similarity. The framework has an API exposed as a

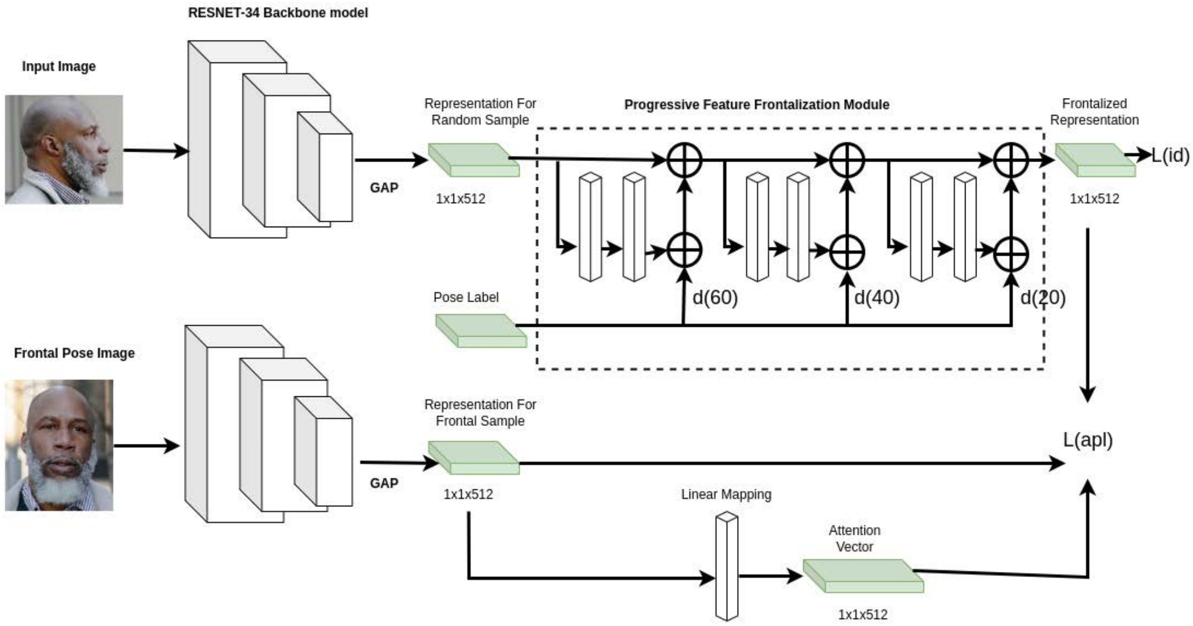

FIGURE 7. An overview of attention-guided Progressive Mapping for Profile Face Recognition framework [179], a light-weight FR framework that uses a step by step pose normalization with attentive pair-wise loss that addresses the common problem of pose variation in FR.

REST service, making it easy to integrate it with a web or mobile app. The source code is available on GitHub at [192] and appears to be under active development as of 2020. It is licensed under MIT License.

G. DeepID

DeepID is a FR model built by researchers from the Chinese University of Hong Kong, published in [193] in 2014. There are two models of DeepID: 1st generation model using 39×31 input 1-channel (grayscale), DeepID1, and 2nd generation model using 55×47 3-channel (RGB) input, DeepID2. It consists of 4 convolutional layers and one fully-connected layer. The softmax layer, used in training, is usually replaced with an early fully connected later represented with 160-dimensional vector for prediction tasks. There are pre-trained models for TensorFlow published by Roy Ruan [194]. A good implementation of DeepID2 is found in the Deepface project. There is also a minimalist implementation by Serengil, published as a notebook [195]. Pre-trained models for the notebook are also available online [195].

H. FACE-API

Face-API is an open-source javascript-based FR API for web browsers implemented on Node.js using TensorFlow.js. Face-API has a face detection model called SSD Mobilenet V1 and is trained on the WIDER FACE dataset. The project also features a lighter version of a Tiny Face Detector detector derived from Tiny YOLO V2 and trained on a custom dataset consisting of approximately 14 thousand images. Face-API implements a lightweight, fast 68-point face keypoint detector. There are two keypoint detectors actually: the default 350kb and the tiny model of 80kb. They all employ depthwise separable convolutions and blocks with dense connections. The models are trained on 35K face images with 68 face keypoint points. A 128-feature vector is used for feature extraction employing an architecture with the same characteristics as the ResNet-34 architecture during FR. Trained models achieve 99.38% on the LFW benchmark. Face-API provides an easy-to-use high-level API, tutorials, live demos, and examples on GitHub [196].

I. face.evoLVe

face.evoLVe [197] is an open-source library for FR tasks such as detection, localization, normalization, and data processing procedures such as augmentation. It uses various backbones such as ResNet and DenseNet, with loss functions such as AmSoftmax, ArcFace, Softmax, Focal, and Triplet. It is built with PyTorch and supports distributed training with multiple GPUs. The source code on the GitHub [198] contains detailed instructions on how to train a model from scratch across a large number of different data sources. There is a reported accuracy of $99.85\% \pm 0.217\%$ on the LFW dataset.

J. SphereFace

SphereFace [44] is a framework that includes face detection, face alignment, and FR modules. The source code is hosted on GitHub and provided under MIT License. SphereFace implementations use various architectures. One such architecture is SphereFace-20, a 20-layer CNN used for training and testing FR on the CAISA-WebFace dataset. The architecture's building blocks are residual units. SphereFace uses Caffe as the implementation platform. Other major libraries required to use the framework are MatLab, MTCNN, and Pdollar toolbox. There are pre-trained models for SphereFace-20 hosted on Google Drive [199] or Baidu [200]. The pre-trained models for SphereFace-4 and on SphereFace-6 are available online at [201]. A top accuracy of 99.30% is reported against the LFW benchmark.

K. SeetaFace

SeetaFace Engine [202] is a CPU-based open-source FR engine developed by the Visual Information Processing and Learning (VIPL) group for face detection, alignment, and recognition written in C++. Multi-view face detection is implemented as a real-time component using a funnel-structured (FuSt) cascaded schema. The face detection code currently works only on Windows. Stacked Auto-encoder Networks [159] are cascaded to achieve facial fiducial points detection. FR is achieved with a modified AlexNet CNN named VIPLFaceNet. This consists of 7 convolutional layers and two full-connected layers having an input size of $256 \times 256 \times 3$, achieving an accuracy of 97.1% on the LFW dataset. It is trained on 1.4 million faces of Mongolians and Caucasians. The code is available on Github [203] under BSD2 license. VisualStudio2013-based examples are provided in the source.

L. AGPM

Attention-guided Progressive Mapping for Profile Face Recognition(AGPM) is a 2021 light-weight FR framework that uses a step-by-step pose normalization with attentive pair-wise loss, introduced in [179] to address the common problem of pose variation in FR. While pose transformation from profile to frontal face can be smooth, it is non-linear along with latent manifolds, posing a large computational cost required to find optimal points for mapping the transformation for FR systems. Huang and Ding [179] approached this problem by decomposing the task into smaller tasks performed at intervals, designed as progressive stacked transformational blocks, as illustrated in Figure 7. An attentive pair-wise loss is used for the supervision of the feature transformation process. The L2 loss is used for channel attention, enabling enhancement of important feature vectors, thus effectively moving profile features closer to frontal ones. Huang and Ding [179] used the MS-Celeb-1M dataset for training the framework, employing MTCNN for face detection, cropping each of the detected faces to 230×230. The framework is implemented in Pytorch, using ResNet for the backbone architecture, with the source code hosted at [204]. Experimental results recorded an accuracy of 89.37% on the CFP dataset. The training data can be downloaded at [205] with pre-trained models available on Baidu [206].

M. PI-NETS

A new group of neural networks that produce outputs of a higher order polynomial of the inputs, referred to as PI-Nets, are introduced in [120]. Chrysos *et al.* [120] formulate the problem as a learning function approximation with a polynomial of input elements. The Pi-Nets, being polynomial networks, can be architectured as generative networks or discriminative networks. The backbone architecture is ResNet for face verification, which is converted into a polynomial network using NCP-Skip. MS1M-RetinaFace is used as training data with images pre-processed to 112×112, with RetinaFace for face detection. Evaluation on the LFW dataset using ResNet50 reports an accuracy of 99.73% with Prodpoly-RestNet reporting 99.83%. Table 2 shows verification performance on other datasets. The authors publish the source of their implementation at [207]. The repository contains implementations in MXNet, PyTorch and Chainer.

TABLE 2. Pi-Net verification performance (%) of ResNet50 and Prodpoly-ResNet50 on LFW, CFP-FF, CFP-FP, CPLFW, AgeDB-30, CALFW and RFW (Caucasian, Indian, Asian and African) [120].

Method	ResNet50	Prodpoly-ResNet50
LFW	99.733 ± 0.309	99.833 ± 0.211
CFP-FF	99.871 ± 0.135	99.886 ± 0.178
CFP-FP	98.800 ± 0.249	98.986 ± 0.274
CFPLFW	92.433 ± 1.245	93.317 ± 1.343
AgeDB-30	98.233 ± 0.655	98.467 ± 0.623
CALFW	95.917 ± 1.209	96.233 ± 1.114
RFW-Caucasian	99.333 ± 0.307	99.700 ± 0.100
RFW-Indian	98.567 ± 0.507	99.300 ± 0.296
RFW-Asian	98.333 ± 0.435	98.950 ± 0.350
RFW-Afican	98.650 ± 0.329	99.417 ± 0.227

N. MagFace

MagFace is a FR framework by Meng *et al.* [58] for learning universal and quality-based face representation. MagFace encourages minimizing intra-class distances and maintaining a cone-like relationship within the classes by learning a universal embedding that pushes ambiguous samples away from class centers. Specifically, MagFace exploits the magnitude property of feature vectors, which is used for quality assessment of faces by adaptively rewarding features with large magnitude and pushing away less learned feature vectors using MagFace loss. MS1M-V2 dataset is used to train and analyze MagFace feature magnitude on 8 (eight) x nVidia 1080i GPUs using stochastic gradient descent. Training data is augmented with random horizontal flips. Results on LFW show a verification score of 99.83% while achieving 98.46% on CFP-FP. The authors released the source code of their work at [208], implemented in Pytorch with an Apache-2.0 license.

VII. RESEARCH GAPS AND EMERGING DIRECTIONS

Major technology companies and research universities drive emerging trends in open-source FR technologies. One of the emerging trends in FR research is 3D FR which builds on the advantages of 2D FR. It is becoming a trend due to its ability to overcome poor lighting conditions or variations in facial expressions and poses. Zhou and Xiao [209] present some of the recent developments in 3D recognition, specifically covering trends in invariant recognition capabilities across the expression, pose, and occlusion. Another trend is continuous improvement in the accuracy of FR. A NIST report [210] shows that there has been an upward trend in FR accuracy since 2013; these gains are so massive that they surpass any improvements made between 2010 and 2013. The FR recognition algorithms developed in 2018 all outperform the then top-performing algorithms of 2013. According to the NIST report, there has been a 20x improvement in FR accuracy between 2013 and 2018.

A. LAW ENFORCEMENT

Recent technologies like 4G/5G networks are increasingly helping users experience richer online services. Availability and relatively affordable complimentary infrastructural technologies, such as Raspberry Pi and other nano smart devices which can host face detection and recognition software, enable law enforcement and security agencies to rapidly deploy services such as public surveillance and detection of suspects, even where there has been challenges related to budgets. Such technology can be valuable to law enforcement agencies by providing FR capabilities for identifying or finding people who are missing. The technology can allow police to capture a photo, detect a face and run it against a face image database of known criminals for matching, an invaluable process for investigators. Because of the portability of nano-devices such as Raspberry PI, wireless cameras can be easily concealed on police clothing and connected via a wireless network to a cloud solution to provide FR capabilities. Early works, such as that published by Chowdhury *et al.* [212] and Shah *et al.* [213] invested a significant amount of research on the capabilities of Raspberry Pi and other resource-constrained devices with a key focus on deploying applications that take advantage of services deployed in the cloud. This foundation has been enhanced, leading to publications such as [214] that show how law enforcement agencies can deploy FR technologies publicly in a cost-efficient manner. Other uses of FR in law enforcement include border checks, where FR technologies are now deployed in some airports, such as Roissy Charles de Gaulle airport in Paris, France. Many states in the US allow law enforcement to use FR for database searches on driver's licenses and ID photos. Deployment of hovering drones in mass events that incorporate FR technologies are used to identify suspects. FR is also increasingly being used to help with identifying and tracking criminals, supporting investigations, and finding missing and exploited children.

B. HEALTH

Face analysis has made significant advances in health. We are increasingly seeing applications that use FR to track patients' use of medication, detect genetic diseases, such as DiGeorge syndrome, and manage pain-related procedures. Remote monitoring of patients is increasingly becoming popular, owing to the maturity of cloud-based health services and increasingly capable handheld devices. FR technologies have been explored in remote patient monitoring to enhance service delivery. The technology can be used to acquire patients' facial images, which could then be sent off to a cloud-based service for analysis in real time. Hossain and Muhammad [215] published a cloud-based face-recognition framework enhanced with speech recognition for remotely monitoring the elderly through the use of portable devices that collect face images which could then be sent over to a cloud-based FR back-end for realtime analysis. Alkawaz *et al.* [216] used FR and cloud computing in conjunction with augmented reality to enable rich access to patient information, especially in emergency cases where such information might be readily available on hand. They present a system that enhances the availability of information to medical practitioners and patients. The technology can also be extended to enable more efficient remote monitoring. FR has also been deployed in mental health monitoring and intervention. FR is well suited for detecting and recognizing human emotion and can be used to predict behavior by analyzing face images. Realtime emotion tracking is equally a useful way of emotion analysis in FR. This can be useful for the mental healthcare industry. Tracking facial keypoints can give useful insight for interpreting patients' inner feelings. Wang [217] published a campus suicide prediction system that can be deployed in schools to identify students with mental problems that are likely to lead to suicide. Data collected from such a system can aid engineers and mental health specialists in designing systems that can monitor and predict suicide behavior and analyze the mental health status of people on the verge of committing suicide. FR is also being deployed for patient check-in and check-out processes at busy hospitals, effectively freeing hospital personnel from paperwork. Such solutions are deployed to help correctly identify patients and eliminate wrong procedures and wrong-patient errors, a mistake that can result in severe temporary or permanent harm or death. Facial recognition software in healthcare is also useful for fraud prevention, especially in cases that could result in patient impersonation. It is often incorporated into a video surveillance system on the hospital premises and can efficiently track criminals such as drug dealers. FR can also be exploited to infringe medical research participants' privacy indirectly. According to Schwarz *et al.* [218], magnetic resonance imaging (MRI) images often contain sufficient data that can be used to reconstruct a participant's face through the use of 3D face reconstruction software. When this happens, it is possible to use FR to establish the identity of the participant in question. Schwarz *et al.* [218] reported that in about 83% of the cases, a FR software could correctly identify a participant from publicly available images using the reconstructed image from the MRI. Figure 9 illustrates how this can be achieved.

C. COVID-19 PANDEMIC

Artificial Intelligence (AI) and FR technologies are at the forefront of tackling the devastating effects of the Covid-19 pandemic. The solutions have been used to track people in

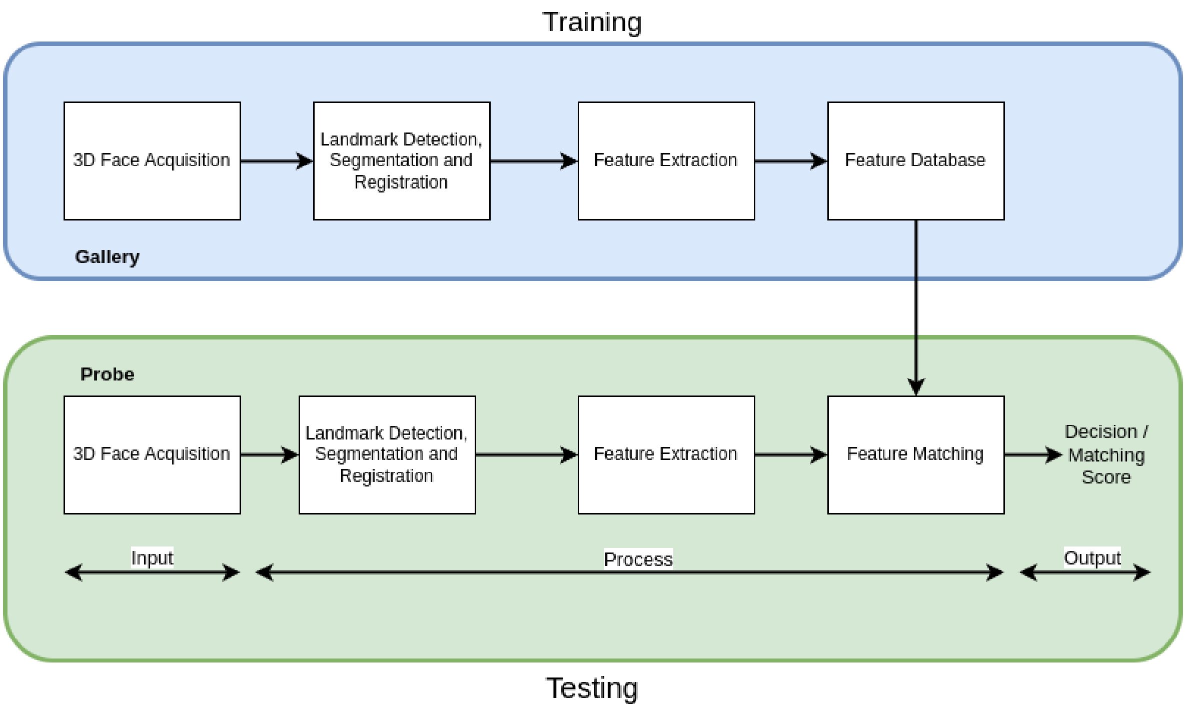

FIGURE 8. A typical 3D face recognition pipeline [211].

quarantine through mobile apps. Operationally, this could include regularly asking users to verify that they are staying inside through self-taken photos or video clips. FR is also being safely used for verification in access-controlled areas in public areas due to its non-contact nature. Other applications include deployment in identifying disease clusters, case monitoring and prediction of future outbreaks, and mortality risk and diagnosis. The advent of COVID-19 led to most people wearing masks, introducing an interesting challenge for FR systems previously trained on non-masked faces. Montero *et al.* [219] highlighted what is likely to be a new feature for FR systems: the ability to detect a face behind a mask with high accuracy. The paper presents Multi-Task Archface (MTArchFace), a new implementation based on Arcface [32], using ResNet-50 as the base architecture, with experimental results showing an increased accuracy averaging 99.78% with mask-use classification

D. RETAIL

Retailers are currently experimenting and deploying FR in retail by placing cameras on shelves and other strategic locations to support analysis of shopper's behavior based on FR and face analysis. Major commercial giants such as Facebook and Amazon are behind some deployments. FR provides enhanced capabilities that allow retail shops to analyze consumer behavior using rich data in real-time. Such analysis could help influence consumer loyalty and consumer satisfaction and increase purchasing probability. Generosi *et al.* [220] described how facial expressions and associated biometric information could be captured and analyzed in retail stores across age, ethnicity, and gender. The authors' system consists of facial expression and emotion recognition modules, gaze detection, and speech recognition modules that utilize CNNs. OpenCV is one of the open-source technologies used in the system. The system can analyze consumers' behaviors and emotions, which can help understand typical and crucial customer actions. The results of such analysis can be used to improve customer experiences in retail spaces. FR has also been applied in the retail banking sector, as reported by Generosi *et al.* [221], citing a number of banks in Russia that have used the technology to prevent violations related to clients' photos. The same technology is used for visual identification when offering services to customers. Generosi *et al.* [222] published technology-enabled personalization research that combines a face-to-face offline personalization with online data-driven personalization in retail spaces. FR is identified as one of the technologies used for enabling consumer-based personalization. The research identifies drivers (such as enhanced control, interaction, integration) and hurdles (such as privacy) to technology-enabled personalization in smart retail. The research also outlines that consumer data is doubling yearly, offering massive opportunities to improve technology-enabled personalization in retail.

E. FACE RECOGNITION COMPUTING

The availability of modern accelerated computing capabilities has played a massive role in the advancement of FR. Specifically, the ever-evolving Graphics Process Unit (GPU) technology makes it possible to quickly carry out massive,

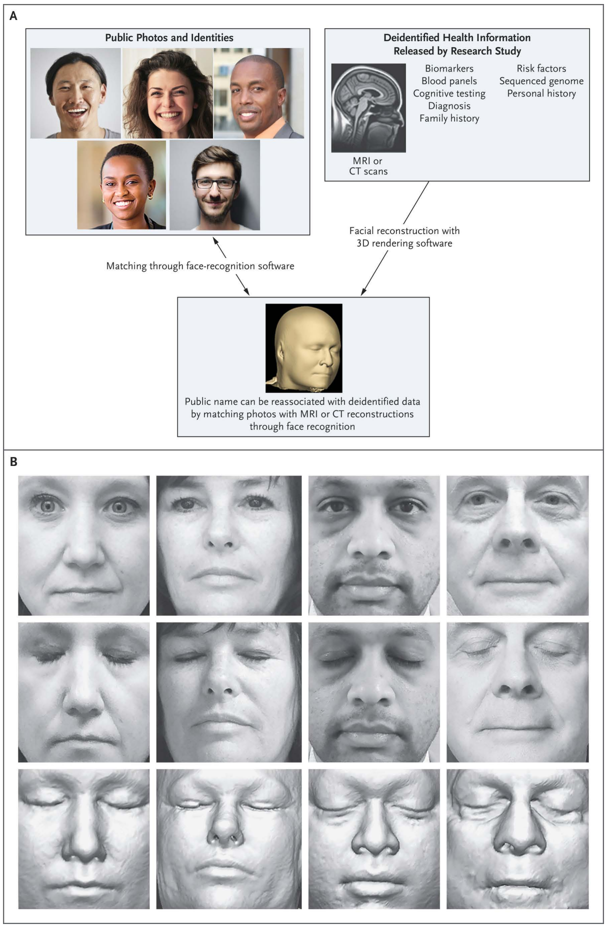

FIGURE 9. Face recognition from reconstructed MRI scans [211].

expensive computations. GPUs are very good at scaling the processing capabilities of FR algorithms and, through parallelism, can churn through a large amount of image data in a reasonable amount of time [223]. GPUs achieve this by highly optimizing computations of matrix operations, a core element of DL algorithms. One of the most popular GPU manufacturers is NVIDIA [224]. NVIDIA provides software accelerated libraries that developers can take advantage of. Computing Unified Device Architecture (CUDA) [225] is one such popular parallel computing & programming framework for GPUs. Available as NVIDIA CUDA Toolkit, the tools expose accelerated computing modules for image and video processing, including mathematical domains of linear algebra. NVIDIA CUDA Deep Neural Network library (cuDNN) is a highly optimized GPU-accelerated implementation of DNN operations such as convolutions, pooling, and normalization, making it suitable for DL training. The implementations are available for frameworks such as Deeplearning4j, TensorFlow, and PyTorch [226]. Other hardware solutions are also dedicated to accelerated computing in the field of FR, such as Google's Tensor Processing Units (TPU) [227], Field Programmable Gate Array (FPGA) [228], IBM's TrueNorth chip [229] and Microsoft's BrainWave [230] among others. A comprehensive list of DL accelerated computing can be found at [231].

F. CLOUD COMPUTING

Cloud computing offers environments that enable convenient, ubiquitous, and on-demand network access to a configurable shared pool of computing resources that are easily provisioned, making it a suitable environment for continuously developing and training FR solutions. Cloud services for artificial intelligence are increasingly being offered as Infrastructure As A Service (IAS), Platform As A Service (PAS), and Software as A Service (SAS). All major computing companies offer FR capabilities packages, often IAS at the highest cost, PAS at a lower fee, and SAS for free. Some of the research work published on the subject include [232], which presents a distributed FR system using support vector machines in a cloud platform. The system uses Hadoop's MapReduce to achieve parallel computing to improve recognition efficiency. The vector machine model reported higher recognition accuracy under cloud computing on Yale B, ORL, and FERET datasets. There are special cases when FR needs intensive computing power that can be dynamically harnessed in a cloud environment. But as reported by Zhang *et al.* [233], cloud computing for FR raises privacy issues.

G. 3D FACE RECOGNITION

The evolution of 3D FR is fueled by the limitations of 2D FR, such as lighting and pose. The ever-improving capabilities of 3D sensors are equally one of the core factors contributing to the advancement of 3D FR. 3D facial data contains more rich geometric information that improves recognition accuracy compared to 2D [234]. As a result, there is an increasing trend of many researchers turning their focus to 3D FR [211]. 3D FR generally consists of training and testing phases with five stages, as illustrated in Figure 8. Training includes data acquisition and pre-processing. 3D face acquisition requires special hardware that can be deployed actively or passively. Active acquisition often uses infrared laser beans to illuminate the target face and use reflection to detect the subject face features. Such systems accomplish their tasks using either triangulation-based or structured light-based [211]. After 3D data is acquired, feature extraction is then used to store face features for later inferencing.

H. LEGISLATION, LEGALITY AND PRIVACY ISSUES

While FR technology can be helpful for efficient bio-metric identification in law enforcement, it can also suffer from bias, leading to potential adverse consequences for the affected individual in the event of mistaken identity. This can cause cases such as mistaken arrests. According to Singer and Metz [235], there has been an increase in concerns about how far government surveillance is permissible by police when using FR technology, prompting the need for legislation. While law enforcement agencies try to address this through stipulated operational processes that are meant to safeguard against bias, these have often proved not adequate, leading to harm and misuse of FR technology [236]. While there are some legislative components covering FR in places such as the United Kingdom (UK) and South Africa, such as the Protection of Personal Information Act 4 of 2013 (POPI) (South Africa), many countries do not yet have clear laws for regulating the use of facial recognition software or other related surveillance technologies. As a result, FR technology can face legal disputes. The *Queen (on application of Edward Bridges) v The Chief Constable of South Wales Police (2020) EWCA Civ. 1058 (R-Bridges)* is one of the first cases in the world where the legality of FR technology was decided in the UK. In South Africa, components of the POPI Act's stipulation point to a regulator's ability to issue a code of conduct applicable to specific conditions, similar to the UK's Surveillance Camera Code of Practice. There was an unreported judgment in South Africa, *Vumacam v Johannesburg Road Agency 2020-08-20 case no. 14867/20*, in which Johannesburg Road Agency (JRA) accused Vumacam [237] of using facial biometric information to spy on peoples' movements, effectively infringing on their right to privacy. The court argued that a legal framework must be in place first; before such technology can be used to collect sensitive bio-metric data, individuals' privacy rights must be respected. In the absence of clear national legislation on regulating FR technology, some countries like the United States of America (USA) have introduced local legislation meant to address and regulate the use of FR technology in law enforcement [236]. But due to competing viewpoints on FR technology, law makers could face a dilemma when determining which regulations to implement. Some cities like San Francisco have enacted local laws banning FR technology by law enforcement and similar agencies. The same enactments are locally in place for Oakland (California) and Somerville (Massachusetts). We also see similar enactments at the state level in the USA (California, New Hampshire, and Oregon), prohibiting the police from using FR footage that is captured on body cameras [236].

In conclusion, FR technology can be used for legitimate investigatory purposes to protect society, but due to lack of uniform regulation, it can be easily abused too, such as bias toward people of color, especially women [238]. Comprehensive regulation needs to be passed in the affected countries where FR is used in law enforcement, with laws mandating public disclosure regarding the deployment of the technology by police.

I. SPOOFING AND ANTI-SPOOFING

Face anti-spoofing involves the prevention of verification of a false photo or video impersonating an authorized person's face. Face spoofing attacks include replay/video attack, which emulates facial movements to achieve a natural look of a face, often achieved by carefully looping over a victim's face. Face spoofing can also involve a straightforward process, such as an attacker presenting a victim's printed or digital photo for verification. A more sophisticated attack consists of using a 3D mask inserted in a face video, introducing fake layers that can deceive depth sensors, emulating natural face movements. Traditionally, anti-spoofing is treated as a binary classification problem involving crafted features detection followed by liveness classification through support vector machines or random forest [239], [240]. Newer methods are CNN-based and can extract more rich features than the traditional hand-crafted approaches [241]–[243]. With FR increasingly moving to mobile and embedded devices, so are face-antispoofing methods. Zhang *et al.* [244] presented a lightweight CNN-based framework for environments that use less computing resources and achieve good accuracy. Central Difference Convolution (CDC) has recently emerged with promising results in face anti-spoofing [245]. Yu *et al.* [70] introduced an improved CDC, Cross Central Difference Convolutions (C-CDC), that utilize differences in local features, employing fewer parameters that result in less computational cost. The methods involve augmenting samples for a wider-ranging domain distribution, delivering state-of-the-art performance on benchmark datasets, Replay-Attack [63] and CASIA-MSDF [69].

J. CHALLENGES

- **Expression**: In FR, face expression is used detect emotions of the subject; for example, in the medical field face expression can be used to analyse how a patient is feeling pain. However, in uncontrolled environments, a person can alter face appearance in an unpredictable way through face expressions, creating a challenge for FR systems not designed and trained for such unpredictable changes. Several studies have been published for building face-expression invariant platforms; Martins *et* . [246] published a 3D-based system that produces 3D-disparity energy models that are face-expression invariant with an experimental recognition rate of 89.33%. A 3D FR system that uses subject-specific curves that are insensitive to intra-face differences is published in [247] with experimental recognition rate of 85% on GavaDB dataset and 88.9% on BU-3DFE datset. Revina and Emmanuel [248] published a method that uses enhanced modified decision based unsymmetric trimmed median filter with additional application of local directional number patterns, dominant gradient local ternary pattern with support vector machines. They report an experimental performance accuracy of 88% on CK and JAFFE datasets.
- **Pose Variation**: In uncontrolled environments where a face can take a wild pose, FR systems still face challenges trying to accurately classify faces in never-before-seen poses. This is especially a challenge if the training data did not have diversified database of images across different poses, degrading classification accuracy [249]. To reduce the negative effects of uncontrolled change in pose, FR systems should be build

to exhibit pose-tolerant features. One way of achieving this is including large-pose variations in the training set, as reported in [250]. Oloyede *et al.* [251] report a number of methods such as local binary feature extraction, divide and rule, 3D reconstruction and others to address the challenge of pose variation using datasets with pose variations such as FERET, MIT and Yale B datasets.

- **Plastic Surgery**: Facial plastic surgery changes appearance of a user depending on the extend of the process. Rhytidectomy is one example where the facial appearance is significantly altered, changing skin texture including eye appearance for youthful appearance. Methods for aiding FR system work over plastic surgery are published in [252]–[254].
- **Illumination**: Illumination is one of the challenging problems still affecting FR. Illumination can cause variations in face appearance creating a challenge for a FR system. Oloyede *et al.* [251] report how early research demonstrated the effects of lighting; changes in illumination can lead to a big difference in image appearance, and these changes can even be bigger compared to inter-face variation. Illumination can also affect the performance of FR system depending on how it affects images used in training and inference. Oloyede *et al.* [251] further report how methods such as image normalization through Gamma intensity corrections, histogram equalisation and 3D face modelling can be used to counter the effects of illumination.
- **Occlusion**: Occlusion can occur in uncontrolled environments when a face is partially blocked for a number of reasons, making accurate FR a challenge. Occlusion can also occur intentionally; a person wearing a hat, sunglasses scarves, etc. Methods developed for dealing with occlusion include a pixel-level occlusion method [255], locally-constrained coding method [256] and a CNN based double-occlusion method in which images with occlusions are present in both training and testing [257].

VIII. CONCLUSION

FR technology is available through numerous open-source projects and commercial vendors providing biometric identification, access control, and other services. This paper looked at FR frameworks landscape in the last ten years. A review of open datasets used by open source projects was carried out. We have identified the following issues: (1) due to privacy, some open datasets consist only of photos of people from the internet (in some cases, only celebrities) and are not a true representative of ordinary life; (2) data noise and demographic bias is also prevalent in some datasets; and (3) cross-pose verification remains a challenge to many datasets. These issues reduce FR accuracy in unconstrained environments.

The open-source face detection and alignment frameworks review revealed that typical face detection methods used are either image-based (often with neural networks) or feature-based. Face detection still encounters challenges across pose, expression, illumination, skin color, orientation, occlusion, etc. However, methods such as 3D facial alignment are poised to overcome some of these issues. The review of open-source frameworks for end-to-end face representation, identification, and verification revealed implementations available via popular environments and tools such as TensorFlow, Pytorch, OpenCV, and Keras, with open-source implementations hosted on platforms such as GitHub. We have observed an upward trend in FR accuracy since 2012, from low-dimension feature-based segmentation implementations to more efficient nets such as deep polynomial neural networks. The majority of contributions to these projects come from top commercial technology companies and research universities. We also found that the increased use of efficient algorithms, 3D technologies, GPUs, and cloud computing has accelerated the upward evolution of these frameworks. These frameworks have implementations that are useful in areas like health, retail, and law enforcement.

In spite of the impressive strides made, FR technologies, including open-source implementations, still face challenges across legislation, legality, privacy, and well-known facial appearance changes due to the environmental factors such as expression, pose, illumination, and occlusion.

REFERENCES

[1] A. K. Jain and S. Z. Li, *Handbook of Face Recognition*, vol. 1. Springer-Verlag, 2005, doi: 10.1007/b138828.
[2] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and D. L. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Comput.*, vol. 1, no. 4, pp. 541–551, 1989.
[3] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM Comput. Surv.*, vol. 35, no. 4, pp. 399–458, 2003.
[4] K. W. Bowyer, K. Chang, and P. Flynn, "A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition," *Comput. Vis. Image Understand.*, vol. 101, no. 1, pp. 1–15, Jan. 2006.
[5] A. Scheenstra, A. Ruifrok, and R. C. Veltkamp, "A survey of 3D face recognition methods," in *Audio- and Video-Based Biometric Person Authentication* (Lecture Notes in Computer Science). Berlin, Germany: Springer, 2005, pp. 891–899, doi: 10.1007/11527923_93.
[6] R. Jafri and H. R. Arabnia, "A survey of face recognition techniques," *J. Inf. Process. Syst.*, vol. 5, no. 2, pp. 41–68, 2009.
[7] I. Masi, Y. Wu, T. Hassner, and P. Natarajan, "Deep face recognition: A survey," in *Proc. 31st SIBGRAPI Conf. Graph., Patterns Images (SIBGRAPI)*, Oct. 2018, pp. 471–478.
[8] Z. An, W. Deng, T. Yuan, and J. Hu, "Deep transfer network with 3D morphable models for face recognition," in *Proc. 13th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, May 2018, pp. 416–422.
[9] M. Wang and W. Deng, "Deep face recognition: A survey," *Neurocomputing*, vol. 429, pp. 215–244, Mar. 2021, doi: 10.1016/j.neucom.2020.10.081.
[10] B. Moghaddam, W. Wahid, and A. Pentland, "Beyond eigenfaces: Probabilistic matching for face recognition," in *Proc. 3rd IEEE Int. Conf. Autom. Face Gesture Recognit.*, 1998, pp. 30–35.
[11] X. He, S. Yan, Y. Hu, P. Niyogi, and H.-J. Zhang, "Face recognition using Laplacianfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 3, pp. 328–340, Mar. 2005.
[12] L. Wiskott, J.-M. Fellous, N. Kuiger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 775–779, Jul. 1997.
[13] P. N. Belhumeur, J. P. Hespanha, and D. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.
[14] C. Liu and H. Wechsler, "A Gabor feature classifier for face recognition," in *Proc. 8th IEEE Int. Conf. Comput. Vis. (ICCV)*, Jul. 2001, pp. 270–275.

[15] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, Dec. 2006.
[16] D. Chen, X. Cao, F. Wen, and J. Sun, "Blessing of dimensionality: High-dimensional feature and its efficient compression for face verification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 3025–3032.
[17] Z. Cao, Q. Yin, X. Tang, and J. Sun, "Face recognition with learning-based descriptor," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 2707–2714.
[18] Z. Lei, M. Pietikäinen, and S. Z. Li, "Learning discriminant face descriptor," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 3, pp. 289–302, Feb. 2013.
[19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, vol. 25, Dec. 2012, pp. 1097–1105.
[20] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in *Proc. Brit. Mach. Vis. Conf. (BMVC)*, G. K. L. T. X. Xie and M. W. Jones, Eds. BMVA Press, Sep. 2015, pp. 41.1–41.12, doi: 10.5244/C.29.41.
[21] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Mar. 2015, pp. 815–823.
[22] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.
[23] R. K. Srivastava, K. Greff, and J. Schmidhuber, "Highway networks," 2015, *arXiv:1505.00387*.
[24] S. Zagoruyko and N. Komodakis, "Wide residual networks," in *Proc. Brit. Mach. Vis. Conf.*, 2016, pp. 1–15.
[25] J. Hu, L. Shen, and G. Sun, "Squeeze- and-excitation networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141.
[26] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 3–19.
[27] F. Wang, M. Jiang, C. Qian, S. Yang, C. Li, H. Zhang, X. Wang, and X. Tang, "Residual attention network for image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 3156–3164.
[28] A. G. Roy, N. Navab, and C. Wachinger, "Concurrent spatial and channel 'squeeze & excitation' in fully convolutional networks," in *Medical Image Computing and Computer Assisted Intervention—MICCAI 2018*. Springer, 2018, pp. 421–429, doi: 10.1007/978-3-030-00928-1_48.
[29] A. Khan, A. Sohail, and A. Ali, "A new channel boosted convolutional neural network using transfer learning," 2018, *arXiv:1804.08528*.
[30] A. Kolesnikov, L. Beyer, X. Zhai, J. Puigcerver, J. Yung, S. Gelly, and N. Houlsby, "Big transfer (BiT): General visual representation learning," 2019, *arXiv:1912.11370*.
[31] G. G. Chrysos, S. Moschoglou, G. Bouritsas, Y. Panagakis, J. Deng, and S. Zafeiriou, "P-Nets: Deep polynomial neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2020, pp. 7325–7335.
[32] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: Closing the gap to human-level performance in face verification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 1701–1708, doi: 10.1109/CVPR.2014.220.
[33] Y. Sun, X. Wang, and X. Tang, "Deeply learned face representations are sparse, selective, and robust," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 2892–2900.
[34] Y. Sun, D. Liang, X. Wang, and X. Tang, "DeepID3: Face recognition with very deep neural networks," 2015, *arXiv:1502.00873*.
[35] J. Liu, Y. Deng, T. Bai, Z. Wei, and C. Huang, "Targeting ultimate accuracy: Face recognition via deep embedding," 2015, *arXiv:1506.07310*.
[36] X. Wu, R. He, Z. Sun, and T. Tan, "A light CNN for deep face representation with noisy labels," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 11, pp. 2884–2896, Nov. 2018.
[37] Y. Wen, K. Zhang, Z. Li, and Y. Qiao, "A discriminative feature learning approach for deep face recognition," in *Computer Vision—ECCV 2016*. Springer, 2016, pp. 499–515, doi: 10.1007/978-3-319-46478-7_31.
[38] X. Zhang, Z. Fang, Y. Wen, Z. Li, and Y. Qiao, "Range loss for deep face recognition with long-tailed training data," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 5409–5418.
[39] R. Ranjan, C. D. Castillo, and R. Chellappa, "L2-constrained softmax loss for discriminative face verification," 2017, *arXiv:1703.09507*.
[40] F. Wang, X. Xiang, J. Cheng, and A. L. Yuille, "Normface: L2 hypersphere embedding for face verification," in *Proc. 25th ACM Int. Conf. Multimedia*, 2017, pp. 1041–1049.
[41] Y. Liu, H. Li, and X. Wang, "Rethinking feature discrimination and polymerization for large-scale recognition," 2017, *arXiv:1710.00870*.
[42] M. A. Hasnat, J. Bohné, J. Milgram, S. Gentric, and L. Chen, "Von mises-Fisher mixture model-based deep learning: Application to face verification," 2017, *arXiv:1706.04264*.
[43] J. Deng, Y. Zhou, and S. Zafeiriou, "Marginal loss for deep face recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Jul. 2017, pp. 60–68.
[44] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj, and L. Song, "SphereFace: Deep hypersphere embedding for face recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 212–220.
[45] X. Qi and L. Zhang, "Face recognition via centralized coordinate learning," 2018, *arXiv:1801.05678*.
[46] F. Wang, J. Cheng, W. Liu, and H. Liu, "Additive margin softmax for face verification," *IEEE Signal Process. Lett.*, vol. 25, no. 7, pp. 926–930, Jul. 2018.
[47] H. Wang, Y. Wang, Z. Zhou, X. Ji, D. Gong, J. Zhou, Z. Li, and W. Liu, "Cosface: Large margin cosine loss for deep face recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 5265–5274.
[48] J. Deng, J. Guo, J. Yang, N. Xue, I. Cotsia, and S. P. Zafeiriou, "ArcFace: Additive angular margin loss for deep face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, p. 1, 2021, doi: 10.1109/tpami.2021.3087709.
[49] X. Zhang, R. Zhao, Y. Qiao, X. Wang, and H. Li, "AdaCos: Adaptively scaling cosine logits for effectively learning deep face representations," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 10823–10832.
[50] M. Yan, M. Zhao, Z. Xu, Q. Zhang, G. Wang, and Z. Su, "VarGFaceNet: An efficient variable group convolutional neural network for lightweight face recognition," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW)*, Oct. 2019, pp. 1–8.
[51] Y. Huang, Y. Wang, Y. Tai, X. Liu, P. Shen, S. Li, J. Li, and F. Huang, "Curricularface: Adaptive curriculum learning loss for deep face recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2020, pp. 5901–5910.
[52] Y. Kim, W. Park, M.-C. Roh, and J. Shin, "GroupFace: Learning latent groups and constructing group-based representations for face recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2020, pp. 5621–5630.
[53] Y. Kim, W. Park, and J. Shin, "BroadFace: Looking at tens of thousands of people at once for face recognition," 2020. [Online]. Available: https://arxiv.org/abs/2008.06674, doi: 10.48550/ARXIV.2008.06674.
[54] J. Xu, T. Guo, Y. Xu, Z. Xu, and K. Bai, "MultiFace: A generic training mechanism for boosting face recognition performance," *Neurocomputing*, vol. 448, pp. 40–47, Aug. 2021.
[55] X. Yang, X. Jia, D. Gong, D.-M. Yan, Z. Li, and W. Liu, "LARNet: Lie algebra residual network for face recognition," in *Proc. Int. Conf. Mach. Learn.*, 2021, pp. 11738–11750.
[56] Q. Wang, P. Zhang, H. Xiong, and J. Zhao, "Face.evoLVe: A high-performance face recognition library," 2021, *arXiv:2107.08621*.
[57] F. Boutros, N. Damer, M. Fang, F. Kirchbuchner, and A. Kuijper, "MixFaceNets: Extremely efficient face recognition networks," in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Aug. 2021, pp. 1–8.
[58] Q. Meng, S. Zhao, Z. Huang, and F. Zhou, "MagFace: A universal representation for face recognition and quality assessment," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 14225–14234.
[59] G. B. Huang, M. Mattar, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," in *Workshop on Faces in 'Real-Life' Images: Detection, Alignment, and Recognition*, E. Learned-Miller, A. Ferencz, and F. Jurie, Eds. Marseille, France, Oct. 2008. [Online]. Available: https://hal.inria.fr/inria-00321923
[60] University of the Massachusetts. *Labeled Faces in the Wild*. Accessed: Mar. 14, 2021. [Online]. Available: http://vis-www.cs.umass.edu/lfw/
[61] W. Zhang, X. Wang, and X. Tang, "Coupled information-theoretic encoding for face photo-sketch recognition," in *Proc. CVPR*, Jun. 2011, pp. 513–520.
[62] C. Fu, X. Wu, Y. Hu, H. Huang, and R. He, "DVG-face: Dual variational generation for heterogeneous face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Jan. 18, 2021, doi: 10.1109/TPAMI.2021.3052549.

[63] I. Chingovska, A. Anjos, and S. Marcel, "On the effectiveness of local binary patterns in face anti-spoofing," in *Proc. Int. Conf. Biometrics Special Interest Group (BIOSIG)*, Sep. 2012, pp. 1–7.
[64] J. Yang, Z. Lei, and S. Z. Li, "Learn convolutional neural network for face anti-spoofing," 2014, *arXiv:1408.5601*.
[65] W.-S. Zheng, X. Li, T. Xiang, S. Liao, J. Lai, and S. Gong, "Partial person re-identification," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 4678–4686.
[66] L. He, Z. Sun, Y. Zhu, and Y. Wang, "Recognizing partial biometric patterns," 2018, *arXiv:1810.07399*.
[67] L. He, X. Liao, W. Liu, X. Liu, P. Cheng, and T. Mei, "FastReID: A pytorch toolbox for general instance re-identification," 2020, *arXiv:2006.02631*.
[68] W.-S. Zheng, X. Li, T. Xiang, S. Liao, J. Lai, and S. Gong. *PartialREID*. Accessed: Mar. 14, 2021. [Online]. Available: https://drive.google.com/file/d/1p7Jvo-RJhU_B6hf9eAhIEFNhvrzM5cdh/view
[69] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, and S. Z. Li, "A face antispoofing database with diverse attacks," in *Proc. 5th IAPR Int. Conf. Biometrics (ICB)*, Mar. 2012, pp. 26–31.
[70] Z. Yu, Y. Qin, H. Zhao, X. Li, and G. Zhao, "Dual-cross central difference network for face anti-spoofing," 2021, *arXiv:2105.01290*.
[71] K. Ricanek and T. Tesafaye, "MORPH: A longitudinal image database of normal adult age-progression," in *Proc. 7th Int. Conf. Autom. Face Gesture Recognit. (FGR)*, Apr. 2006, pp. 341–345.
[72] P. Terhörst, J. N. Kolf, N. Damer, F. Kirchbuchner, and A. Kuijper, "Face quality estimation and its correlation to demographic and non-demographic bias in face recognition," in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Apr. 2020, pp. 1–11.
[73] J. Zhao, Y. Cheng, Y. Cheng, Y. Yang, F. Zhao, J. Li, H. Liu, S. Yan, and J. Feng, "Look across elapse: Disentangled representation learning and photorealistic cross-age face synthesis for age-invariant face recognition," in *Proc. AAAI Conf. Artif. Intell.*, vol. 33, no. 1, 2019, pp. 9251–9258.
[74] D. Yi, Z. Lei, S. Liao, and S. Z. Li, "Learning face representation from scratch," 2014, *arXiv:1411.7923*.
[75] Amazon. *Internet Movies Database*. Accessed: Mar. 14, 2021. [Online]. Available: https://www.imdb.com
[76] X. Fan and M. Jiang, "RetinaFaceMask: A single stage face mask detector for assisting control of the COVID-19 pandemic," 2020, *arXiv:2005.03950*.
[77] A. S. Georghiades, P. N. Belhumeur, and D. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 643–660, Jun. 2001.
[78] D. Yi, Z. Lei, S. Liao, and S. Z. Li. *Yale Face Database B*. Accessed: Apr. 10, 2022. [Online]. Available: http://vision.ucsd.edu/~leekc/ExtYaleDatabase/download.html
[79] Yale-Database. *Adience Dataset*. Accessed: Apr. 10, 2022. [Online]. Available: http://www.cslab.openu.ac.il/download/
[80] E. Eidinger, R. Enbar, and T. Hassner, "Age and gender estimation of unfiltered faces," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 12, pp. 2170–2179, Dec. 2014.
[81] H.-W. Ng and S. Winkler, "A data-driven approach to cleaning large face datasets," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 343–347.
[82] S. Yang, P. Luo, C. C. Loy, and X. Tang, "WIDER FACE: A face detection benchmark," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 5525–5533.
[83] FaceAPIJS. *Face API JS*. Accessed: Mar. 14, 2021. [Online]. Available: https://justadudewhohacks.github.io/face-api.js/docs/index.html
[84] A. Bansal, A. Nanduri, C. D. Castillo, R. Ranjan, and R. Chellappa, "UMDFaces: An annotated face dataset for training deep networks," in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Oct. 2017, pp. 464–473.
[85] A. Bansal, A. Nanduri, C. D. Castillo, R. Ranjan, and R. Chellappa. *UMDFaces Dataset*. Accessed: Mar. 14, 2021. [Online]. Available: https://www.umdfaces.io/
[86] Y. Guo, L. Zhang, Y. Hu, X. He, and J. Gao, "MS-Celeb-1M: A dataset and benchmark for large-scale face recognition," in *Computer Vision—ECCV 2016*. Springer, 2016, pp. 87–102, doi: 10.1007/978-3-319-46487-9_6.
[87] J. Deng, J. Guo, Y. Zhou, J. Yu, I. Kotsia, and S. Zafeiriou, "RetinaFace: Single-stage dense face localisation in the wild," 2019, *arXiv:1905.00641*.
[88] protossw512. *Face Analysis Project on MXNet*. Accessed: Mar. 14, 2021. [Online]. Available: https://github.com/deepinsight/insightface/tree/master/alignment
[89] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman, "VGGFace2: A dataset for recognising faces across pose and age," in *Proc. 13th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, May 2018, pp. 67–74.
[90] M. Wang, W. Deng, J. Hu, X. Tao, and Y. Huang, "Racial faces in the wild: Reducing racial bias by information maximization adaptation network," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, Oct. 2019, pp. 692–702.
[91] J. Wang, Y. Liu, Y. Hu, H. Shi, and T. Mei, "FaceX-zoo: A PyTorch toolbox for face recognition," 2021, *arXiv:2101.04407*.
[92] X. Xu, Y. Huang, P. Shen, S. Li, J. Li, F. Huang, Y. Li, and Z. Cui, "Consistent instance false positive improves fairness in face recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2021, pp. 578–586.
[93] F. Wang, L. Chen, C. Li, S. Huang, Y. Chen, C. Qian, and C. C. Loy, "The devil of face recognition is in the noise," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 765–780.
[94] W. Hu, Y. Huang, F. Zhang, and R. Li, "Noise-tolerant paradigm for training face recognition CNNs," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 11887–11896.
[95] Q. Huang, L. Yang, H. Huang, T. Wu, and D. Lin, "Caption-supervised face recognition: Training a state-of-the-art face model without manual annotation," in *Computer Vision—ECCV 2020*. Springer, 2020, pp. 139–155, doi: 10.1007/978-3-030-58520-4_9.
[96] Y. Shen, Y. Xiong, W. Xia, and S. Soatto, "Towards backward-compatible representation learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 6368–6377.
[97] B. F. Klare, B. Klein, E. Taborsky, A. Blanton, J. Cheney, K. Allen, P. Grother, A. Mah, and A. K. Jain, "Pushing the frontiers of unconstrained face detection and recognition: Iarpa Janus benchmark A," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 1931–1939.
[98] L. Yang, X. Zhan, D. Chen, J. Yan, C. C. Loy, and D. Lin, "Learning to cluster faces on an affinity graph," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 2298–2306.
[99] H. Zhou, J. Liu, Z. Liu, Y. Liu, and X. Wang, "Rotate- and-render: Unsupervised photorealistic face rotation from single-view images," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2020, pp. 5911–5920.
[100] A. T. Tran, T. Hassner, I. Masi, and G. Medioni, "Regressing robust and discriminative 3D morphable models with a very deep neural network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5163–5172.
[101] C. Whitelam, E. Taborsky, A. Blanton, B. Maze, J. Adams, T. Miller, N. Kalka, A. K. Jain, J. A. Duncan, K. Allen, and J. Cheney, "IARPA Janus benchmark-B face dataset," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Jul. 2017, pp. 90–98.
[102] B. Maze, J. Adams, J. A. Duncan, N. Kalka, T. Miller, C. Otto, A. K. Jain, W. T. Niggel, J. Anderson, J. Cheney, and P. Grother, "IARPA Janus benchmark-C: Face dataset and protocol," in *Proc. Int. Conf. Biometrics (ICB)*, Feb. 2018, pp. 158–165.
[103] J. Deng, J. Guo, T. Liu, M. Gong, and S. Zafeiriou, "Sub-center arcface: Boosting face recognition by large-scale noisy web faces," in *Proc. Eur. Conf. Comput. Vis.* Springer, 2020, pp. 741–757.
[104] J. Guo, X. Zhu, Z. Lei, and S. Z. Li, "Face synthesis for eyeglass-robust face recognition," in *Proc. Chin. Conf. Biometric Recognit.* Springer, 2018, pp. 275–284.
[105] MeGlass. *Face Synthesis for Eyeglass-Robust Face Recognition*. Accessed: Apr. 10, 2022. [Online]. Available: https://github.com/cleardusk/MeGlass
[106] J. Guo, X. Zhu, C. Zhao, D. Cao, Z. Lei, and S. Z. Li, "Learning meta face recognition in unseen domains," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2020, pp. 6163–6172.
[107] A. Rangesh, B. Zhang, and M. M. Trivedi, "Gaze preserving CycleGANs for eyeglass removal & persistent gaze estimation," 2020, *arXiv:2002.02077*.
[108] Z. Cheng, X. Zhu, and S. Gong, "Surveillance face recognition challenge," 2018, *arXiv:1804.09691*.
[109] QMUL-SurvFace. *Surveillance Face Recognition Challenge*. Accessed: Apr. 10, 2022. [Online]. Available: https://github.com/cleardusk/MeGlass
[110] Y. Liu, B. Peng, P. Shi, H. Yan, Y. Zhou, B. Han, Y. Zheng, C. Lin, J. Jiang, Y. Fan, T. Gao, G. Wang, J. Liu, X. Lu, and D. Xie, "IQIYI-VID: A large dataset for multi-modal person identification," 2018, *arXiv:1811.07548*.
[111] J. Ye, Y. Guan, J. Liu, X. Huang, and H. Zhang, "Large-scale multi-modal person identification in real unconstrained environments," in *Proc. IEEE Int. Conf. Robot. Biomimetics (ROBIO)*, Dec. 2019, pp. 1–6.

[112] M. Merler, N. Ratha, R. S. Feris, and J. R. Smith, "Diversity in faces," 2019, *arXiv:1901.10436*.
[113] K. Karkkainen and J. Joo, "Fairface: Face attribute dataset for balanced race, gender, and age for bias measurement and mitigation," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, Jan. 2021, pp. 1548–1558.
[114] A. Suri, M. Vatsa, and R. Singh, "A2-LINK: Recognizing disguised faces via active learning and adversarial noise based inter-domain knowledge," *IEEE Trans. Biometrics, Behav., Identity Sci.*, vol. 2, no. 4, pp. 326–336, Oct. 2020.
[115] C. Hazirbas, J. Bitton, B. Dolhansky, J. Pan, A. Gordo, and C. C. Ferrer, "Towards measuring fairness in AI: The casual conversations dataset," 2021, *arXiv:2104.02821*.
[116] B. F. Klare, M. J. Burge, J. C. Klontz, R. W. V. Bruegge, and A. K. Jain, "Face recognition performance: Role of demographic information," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 6, pp. 1789–1801, Dec. 2012.
[117] J. Zhao, Y. Cheng, Y. Xu, L. Xiong, J. Li, F. Zhao, K. Jayashree, S. Pranata, S. Shen, J. Xing, S. Yan, and J. Feng, "Towards pose invariant face recognition in the wild," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 2207–2216.
[118] G. Chen, Y. Shao, C. Tang, Z. Jin, and J. Zhang, "Deep transformation learning for face recognition in the unconstrained scene," *Mach. Vis. Appl.*, vol. 29, no. 3, pp. 513–523, Apr. 2018.
[119] S. Sengupta, J.-C. Chen, C. Castillo, V. M. Patel, R. Chellappa, and D. W. Jacobs, "Frontal to profile face verification in the wild," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2016, pp. 1–9.
[120] G. G. Chrysos, S. Moschoglou, G. Bouritsas, J. Deng, Y. Panagakis, and S. P. Zafeiriou, "Deep polynomial neural networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, p. 1, 2021, doi: 10.1109/tpami.2021.3058891.
[121] M. K. Hasan, M. S. Ahsan, Abdullah-Al-Mamun, S. H. S. Newaz, and G. M. Lee, "Human face detection techniques: A comprehensive review and future research directions," *Electronics*, vol. 10, no. 19, p. 2354, Sep. 2021, doi: 10.3390/electronics10192354.
[122] S. W. F. Earp, P. Noinongyao, J. A. Cairns, and A. Ganguly, "Face detection with feature pyramids and landmarks," 2019, *arXiv:1912.00596*.
[123] D. Yashunin, T. Baydasov, and R. Vlasov, "MaskFace: Multi-task face and landmark detector," 2020, *arXiv:2005.09412*.
[124] Z. Li, X. Tang, J. Han, J. Liu, and R. He, "PyramidBox++: High performance detector for finding tiny face," 2019, *arXiv:1904.00386*.
[125] P. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, 2004.
[126] J. Li, Y. Wang, C. Wang, Y. Tai, J. Qian, J. Yang, C. Wang, J. Li, and F. Huang, "DSFD: Dual shot face detector," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 5060–5069.
[127] Y. Zhu, H. Cai, S. Zhang, C. Wang, and Y. Xiong, "TinaFace: Strong but simple baseline for face detection," 2020, *arXiv:2011.13183*.
[128] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.
[129] Tinaface. *Strong But Simple Baseline for Face Detection*. Accessed: Apr. 10, 2022. [Online]. Available: https://github.com/Media-Smart/vedadet/tree/main/configs/trainval/tinaface
[130] S. Wu, X. Li, and X. Wang, "IoU-aware single-stage object detector for accurate localization," *Image Vis. Comput.*, vol. 97, May 2020, Art. no. 103911.
[131] F. Zhang, X. Fan, G. Ai, J. Song, Y. Qin, and J. Wu, "Accurate face detection for high performance," 2019, *arXiv:1905.01585*.
[132] J. Yu, Y. Jiang, Z. Wang, Z. Cao, and T. Huang, "Unitbox: An advanced object detection network," in *Proc. 24th ACM Int. Conf. Multimedia*, 2016, pp. 516–520.
[133] W. Tian, Z. Wang, H. Shen, W. Deng, Y. Meng, B. Chen, X. Zhang, Y. Zhao, and X. Huang, "Learning better features for face detection with feature fusion and segmentation supervision," 2018, *arXiv:1811.08557*.
[134] S. Zhang, X. Zhu, Z. Lei, H. Shi, X. Wang, and S. Z. Li, "S3FD: Single shot scale-invariant face detector," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 192–201.
[135] WIDER-FACE. *Wider Test*. Accessed: Apr. 10, 2022. [Online]. Available: https://github.com/sfzhang15/SFD/blob/master/sfd_test_code/WIDER_FACE/wider_test.py
[136] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1499–1503, Oct. 2016.
[137] D. Sandberg. *Facenet MTCNN Implementation for TensorFlow*. Accessed: Mar. 14, 2021. [Online]. Available: https://github.com/davidsandberg/facenet/tree/master/src/align
[138] I. de Paz Centeno. *MTCNN Face Detection Implementation for TensorFlow*. Accessed: Mar. 14, 2021. [Online]. Available: https://github.com/ipazc/mtcnn
[139] AIZOO Tech. *Face Mask Detection*. Accessed: Mar. 14, 2021. [Online]. Available: https://github.com/AIZOOTech/FaceMaskDetection
[140] Z. Shao, Z. Liu, J. Cai, and L. Ma, "JAA-Net: Joint facial action unit detection and face alignment via adaptive attention," *Int. J. Comput. Vis.*, vol. 129, no. 2, pp. 321–340, 2021.
[141] Z. Shao. *Pytorch-JAANet*. [Online]. Available: https://github.com/ZhiwenShao/PyTorch-JAANet
[142] X. Li, X. Zhang, H. Yang, W. Duan, W. Dai, and L. Yin, "An EEG-based multi-modal emotion database with both posed and authentic facial actions for emotion analysis," in *Proc. 15th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, Nov. 2020, pp. 328–335.
[143] G. Song, Y. Liu, Y. Zang, X. Wang, B. Leng, and Q. Yuan, "Kpnet: Towards minimal face detector," in *Proc. AAAI Conf. Artif. Intell.*, vol. 34, no. 7, 2020, pp. 12015–12022.
[144] Researcher-MSU. *Pytorch Face Landmark Detection*. Accessed: Apr. 10, 2022. [Online]. Available: https://github.com/cunjian/pytorch_face_landmark
[145] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.
[146] T. Nguyen. *YOLOFace*. Accessed: Mar. 14, 2021. [Online]. Available: https://github.com/sthanhng/yoloface
[147] RetinaFace. *RetinaFace in PyTorch*. Accessed: Mar. 14, 2021. [Online]. Available: https://github.com/biubug6/Pytorch_Retinaface
[148] S. Wu, M. Kan, Z. He, S. Shan, and X. Chen, "Funnel-structured cascade for multi-view face detection with alignment-awareness," *Neurocomputing*, vol. 221, pp. 138–145, Jan. 2017.
[149] P. Barra, C. Bisogni, M. Nappi, and S. Ricciardi, "Fast quadtree-based pose estimation for security applications using face biometrics," in *Proc. Int. Conf. Netw. Syst. Secur.* Springer, 2018, pp. 160–173.
[150] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar, "Attribute and simile classifiers for face verification," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep. 2009, pp. 365–372.
[151] O. Çeliktutan, S. Ulukaya, and B. Sankur, "A comparative study of face landmarking techniques," *EURASIP J. Image Video Process.*, vol. 2013, no. 1, pp. 1–27, Dec. 2013.
[152] H. Yang, X. Jia, C. Change Loy, and P. Robinson, "An empirical study of recent face alignment methods," 2015, *arXiv:1511.05049*.
[153] X. Jin and X. Tan, "Face alignment in-the-wild: A survey," *Comput. Vis. Image Understand.*, vol. 162, pp. 1–22, Sep. 2017.
[154] N. Wang, X. Gao, D. Tao, H. Yang, and X. Li, "Facial feature point detection: A comprehensive survey," *Neurocomputing*, vol. 275, pp. 50–65, Jan. 2018.
[155] J. M. Saragih, S. Lucey, and J. F. Cohn, "Deformable model fitting by regularized landmark mean-shift," *Int. J. Comput. Vis.*, vol. 91, no. 2, pp. 200–215, Jan. 2011.
[156] G. Tzimiropoulos, J. Alabort-i-Medina, S. P. Zafeiriou, and M. Pantic, "Active orientation models for face alignment in-the-wild," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 12, pp. 2024–2034, Dec. 2014.
[157] G. Tzimiropoulos and M. Pantic, "Gauss-Newton deformable part models for face alignment in-the-wild," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1851–1858.
[158] X. Xiong and F. De la Torre, "Supervised descent method and its applications to face alignment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 532–539.
[159] J. Zhang, S. Shan, M. Kan, and X. Chen, "Coarse-to-fine auto-encoder networks (cfan) for real-time face alignment," in *Proc. Eur. Conf. Comput. Vis.* Springer, 2014, pp. 1–16.
[160] G. Trigeorgis, P. Snape, M. A. Nicolaou, E. Antonakos, and S. Zafeiriou, "Mnemonic descent method: A recurrent process applied for End-to-End face alignment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 4177–4187.
[161] A. Jourabloo and X. Liu, "Large-pose face alignment via CNN-based dense 3D model fitting," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 4188–4196.
[162] A. Bulat and G. Tzimiropoulos, "Two-stage convolutional part heatmap regression for the 1st 3D face alignment in the wild (3DFAW) challenge," in *Proc. Eur. Conf. Comput. Vis.* Springer, 2016, pp. 616–624.
[163] M. Kowalski, J. Naruniec, and T. Trzcinski, "Deep alignment network: A convolutional neural network for robust face alignment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Jul. 2017, pp. 88–97.
[164] X. Wang, L. Bo, and L. Fuxin, "Adaptive wing loss for robust face alignment via heatmap regression," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 6971–6981.

[165] protossw512. *Adaptive Wing Loss for Robust Face Alignment via Heatmap Regression—Official Implementation*. Accessed: Mar. 14, 2021. [Online]. Available: https://github.com/protossw512/AdaptiveWingLoss
[166] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, "300 faces in-the-wild challenge: The first facial landmark localization challenge," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops*, Dec. 2013, pp. 397–403.
[167] W. Wu, C. Qian, S. Yang, Q. Wang, Y. Cai, and Q. Zhou, "Look at boundary: A boundary-aware face alignment algorithm," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 2129–2138.
[168] J. Guo, J. Deng, N. Xue, and S. Zafeiriou, "Stacked dense U-nets with dual transformers for robust face alignment," 2018. [Online]. Available: https://arxiv.org/abs/1812.01936, doi: 10.48550/ARXIV.1812.01936.
[169] M. Dantone, J. Gall, G. Fanelli, and L. Van Gool, "Real-time facial feature detection using conditional regression forests," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 2578–2585.
[170] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Proc. Eur. Conf. Comput. Vis.* Springer, 2014, pp. 818–833.
[171] S. Sun and Q. Chen, "Distance metric learning for large margin nearest neighbor classification," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 25, no. 7, pp. 1073–1087, Nov. 2011.
[172] D. SandBerg. *Face Recognition Using Tensorflow*. Accessed: Mar. 15, 2021. [Online]. Available: https://github.com/davidsandberg/facenet
[173] H. Taniai. *Facenet Implementation By Keras2*. Accessed: Mar. 14, 2021. [Online]. Available: https://github.com/noki-mtl/keras-facenet
[174] J. Deng, A. Roussos, G. Chrysos, E. Ververas, I. Kotsia, J. Shen, and S. Zafeiriou, "The menpo benchmark for multi-pose 2D and 3D facial landmark localisation and tracking," *Int. J. Comput. Vis.*, vol. 127, nos. 6–7, pp. 599–624, Jun. 2019.
[175] X. An, X. Zhu, Y. Xiao, L. Wu, M. Zhang, Y. Gao, B. Qin, D. Zhang, and Y. Fu, "Partial FC: Training 10 million identities on a single machine," 2020, *arXiv:2010.05222*.
[176] J. Deng, J. Guo, T. Liu, M. Gong, and S. Zafeiriou, "Sub-center arcface: Boosting face recognition by large-scale noisy web faces," in *Proc. IEEE Conf. Eur. Conf. Comput. Vis.*, 2020, pp. 741–757.
[177] J. Deng, J. Guo, E. Ververas, I. Kotsia, and S. Zafeiriou, "Retinaface: Single-shot multi-level face localisation in the wild," in *Proc. CVPR*, Jun. 2020, pp. 5203–5212.
[178] J. Guo, J. Deng, A. Lattas, and S. Zafeiriou, "Sample and computation redistribution for efficient face detection," 2021, *arXiv:2105.04714*.
[179] J. Huang and C. Ding, "Attention-guided progressive mapping for profile face recognition," in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Aug. 2021, pp. 1–8.
[180] H.-R. Chou, J.-H. Lee, Y.-M. Chan, and C.-S. Chen, "Data-specific adaptive threshold for face recognition and authentication," in *Proc. IEEE Conf. Multimedia Inf. Process. Retr. (MIPR)*, Mar. 2019, pp. 153–156.
[181] H.-R. Chou, J.-H. Lee, Y.-M. Chan, and C.-S. Chen. *Evalution Code of Face Recognition and Authentication With Online Registration*. Accessed: Mar. 15, 2021. [Online]. Available: https://github.com/ivclab/Online-Face-Recognition-and-Authentication
[182] OpenBR Team. *OpenBR: Open Source Biometric Recognition*. Accessed: Mar. 15, 2021. [Online]. Available: http://openbiometrics.org
[183] P. J. Phillips, H. Wechsler, J. Huang, and P. J. Rauss, "The FERET database and evaluation procedure for face-recognition algorithms," *Image Vis. Comput.*, vol. 16, no. 5, pp. 295–306, 1998.
[184] B. Klare, "Spectrally sampled structural subspace features (4SF)," Michigan State Univ., East Lansing, MI, USA, Tech. Rep. MSU-CSE-11-16, 2011.
[185] D. S. Bolme, B. A. Draper, and J. R. Beveridge, "Average of synthetic exact filters," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 2105–2112.
[186] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Feb. 2004.
[187] OpenBR Biometrics. *Open Source Biometrics, Face Recognition*. Accessed: Mar. 15, 2021. [Online]. Available: https://github.com/biometrics/openbr
[188] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: Closing the gap to human-level performance in face verification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1701–1708.
[189] S. Ghosh. *Keras Implementation of Closing the Gap to Human-Level Performance in Face Verification*. Accessed: Mar. 15, 2021. [Online]. Available: https://github.com/swghosh/DeepFace
[190] S. I. Serengil and A. Ozpinar, "LightFace: A hybrid deep face recognition framework," in *Proc. Innov. Intell. Syst. Appl. Conf. (ASYU)*, Oct. 2020, pp. 1–5.
[191] S. I. Serengil. *Facebook-Deepface.ipynb*. Accessed: Mar. 15, 2021. [Online]. Available: https://github.com/serengil/tensorflow-101/blob/master/python/Facebook-Deepface.ipynb
[192] S. Ilkin. *Facebook-Realtime-Deepface.ipynb*. Accessed: Mar. 15, 2021. [Online]. Available: https://github.com/serengil/tensorflow-101/blob/master/python/fb-deepface-real-time.py
[193] Y. Sun, X. Wang, and X. Tang, "Deep learning face representation by joint identification-verification," 2014, *arXiv:1406.4773*.
[194] R. Ran. *DeepID Implementation*. Accessed: Mar. 15, 2021. [Online]. Available: https://github.com/Ruoyiran/DeepID
[195] S. I. Serengil. *TensorFlow 101: Introduction to Deep Learning*. Accessed: Mar. 15, 2021. [Online]. Available: https://github.com/serengil/tensorflow-101
[196] D. E. King. *dlib C++ Library*. Accessed: Mar. 15, 2021. [Online]. Available: https://github.com/davisking/dlib
[197] Z. Wang, J. Zhao, C. Lu, F. Yang, H. Huang, and Y. Guo, "Learning to detect head movement in unconstrained remote gaze estimation in the wild," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, 2020, pp. 3443–3452.
[198] V. Mühler. *JavaScript API for Face Detection and Face Recognition in the Browser and Nodejs With tensorflow.js*. Accessed: Mar. 15, 2021. [Online]. Available: https://github.com/justadudewhohacks/face-api.js
[199] Google. *SphereFace Google Model*. Accessed: Mar. 15, 2021. [Online]. Available: https://drive.google.com/open?id=0B_geeR2lTMegb2F6dmlmOXhWaVk
[200] Baidu. *SphereFace Baidu Model*. Accessed: Mar. 15, 2021. [Online]. Available: http://pan.baidu.com/s/1qY5FTF2
[201] W. Liu. *SphereFace: Deep Hypersphere Embedding for Face Recognition*. Accessed: Mar. 15, 2021. [Online]. Available: https://github.com/wy1iu/sphereface
[202] J. He, D. Li, B. Yang, S. Cao, B. Sun, and L. Yu, "Multi view facial action unit detection based on CNN and BLSTM-RNN," in *Proc. 12th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, May 2017, pp. 848–853.
[203] SeetaFace. *SeetaFace Engine*. Accessed: Mar. 15, 2021. [Online]. Available: https://github.com/seetaface/SeetaFaceEngine
[204] APGM. *A Pytorch Implementation of Attention-Guided Progressive Mapping for Profile Face Recognition*. Accessed: Apr. 10, 2022. [Online]. Available: https://github.com/hjy1312/AGPM
[205] CFP-Data. *Attention-Guided Progressive Mapping for Profile Face Recognition Training Data*. Accessed: Apr. 10, 2022. [Online]. Available: https://pan.baidu.com/s/1_3NHIYsfwHii0APP9-0Llw
[206] CFP-Pretained. *Attention-Guided Progressive Mapping for Profile Face Recognition Pretained Models*. Accessed: Apr. 10, 2022. [Online]. Available: https://pan.baidu.com/s/1ye_eSoIQpSaTWjkQayFNsA
[207] Pi-Nets. *Deep Polynomial Neural Networks*. Accessed: Apr. 10, 2022. [Online]. Available: https://github.com/grigorisg9gr/polynomial_nets
[208] MagFace. *A Universal Representation for Face Recognition and Quality Assessment*. Accessed: Apr. 10, 2022. [Online]. Available: https://github.com/IrvingMeng/MagFace
[209] S. Zhou and S. Xiao, "3D face recognition: A survey," *Hum.-Centric Comput. Inf. Sci.*, vol. 8, no. 1, pp. 1–27, 2018.
[210] P. Grother, M. Ngan, and K. Hanaoka, "Face recognition vendor test (FRVT) part 2," Nat. Inst. Standards Technol., Gaithersburg, MD, USA, Sep. 2019, doi: 10.6028/nist.ir.8271.
[211] S. Soltanpour, B. Boufama, and Q. M. J. Wu, "A survey of local feature methods for 3D face recognition," *Pattern Recognit.*, vol. 72, pp. 391–406, Dec. 2017.
[212] M. N. Chowdhury, M. S. Nooman, and S. Sarker, "Access control of door and home security by raspberry pi through internet," *Int. J. Sci. Eng. Res*, vol. 4, no. 1, pp. 550–558, 2013.
[213] D. Shah and V. Haradi, "IoT based biometrics implementation on Raspberry Pi," *Proc. Comput. Sci.*, vol. 79, pp. 328–336, Jan. 2016.
[214] M. Sajjad, M. Nasir, K. Muhammad, S. Khan, Z. Jan, A. K. Sangaiah, M. Elhoseny, and S. W. Baik, "Raspberry pi assisted face recognition framework for enhanced law-enforcement services in smart cities," *Future Gener. Comput. Syst.*, vol. 108, pp. 995–1007, Jul. 2020.
[215] M. S. Hossain and G. Muhammad, "Cloud-assisted speech and face recognition framework for health monitoring," *Mobile Netw. Appl.*, vol. 20, no. 3, pp. 391–399, Jun. 2015.
[216] M. H. Alkawaz, T. Waili, and S. M. Adnan, "Augmented reality for patient information using face recognition and cloud computing," *Int. J. Perceptive Cognit. Comput.*, vol. 6, no. 1, pp. 24–27, 2020.
[217] Z. Wang, "Campus intelligence mental health searching system based on face recognition technology," *J. Electron. Res. Appl.*, vol. 4, no. 4, Aug. 2020.

[218] C. G. Schwarz, W. K. Kremers, T. M. Therneau, R. R. Sharp, J. L. Gunter, P. Vemuri, A. Arani, A. J. Spychalla, K. Kantarci, D. S. Knopman, R. C. Petersen, and C. R. Jack, "Identification of anonymous MRI research participants with face-recognition software," *New England J. Med.*, vol. 381, no. 17, pp. 1684–1686, Oct. 2019.

[219] D. Montero, M. Nieto, P. Leskovsky, and N. Aginako, "Boosting masked face recognition with multi-task ArcFace," 2021, *arXiv:2104.09874*.

[220] A. Generosi, S. Ceccacci, and M. Mengoni, "A deep learning-based system to track and analyze customer behavior in retail store," in *Proc. IEEE 8th Int. Conf. Consum. Electron. Berlin (ICCE-Berlin)*, Sep. 2018, pp. 1–6.

[221] O. Vaganova, N. Bykanova, A. Grigoryan, and N. Cherepovskaya, "Directions of development of bank technologies applied in the Russian market of retail credit services," *Nat. Acad. Managerial Staff Culture Arts Herald*, no. 3, 2018.

[222] A.-S. Riegger, J. F. Klein, K. Merfeld, and S. Henkel, "Technology-enabled personalization in retail stores: Understanding drivers and barriers," *J. Bus. Res.*, vol. 123, pp. 140–155, Feb. 2021.

[223] A. Cano, "A survey on graphic processing unit computing for large-scale data mining," *Wiley Interdiscipl. Rev., Data Mining Knowl. Discovery*, vol. 8, no. 1, p. e1232, Jan. 2018.

[224] NVIDIA Developer. *Accelerated Computing*. Accessed: Mar. 14, 2021. [Online]. Available: https://developer.nvidia.com/hpc

[225] NVIDIA. *CUDA Zone*. Accessed: Mar. 14, 2021. [Online]. Available: https://developer.nvidia.com/cuda-zone

[226] V. Kovalev, A. Kalinovsky, and S. Kovalev, "Deep learning with theano, torch, caffe, tensorflow, and deeplearning4J: Which one is the best in speed and accuracy?" in *Proc. Int. Conf. Pattern Recognit. Inf. Process.*, 2016.

[227] Google. *Cloud TPU*. Accessed: Mar. 14, 2021. [Online]. Available: https://cloud.google.com/tpu

[228] G. Lacey, G. W. Taylor, and S. Areibi, "Deep learning on FPGAs: Past, present, and future," 2016. [Online]. Available: https://arxiv.org/abs/1602.04283, doi: 10.48550/ARXIV.1602.04283.

[229] IBM. *Introducing a Brain-Inspired Computer*. Accessed: Mar. 14, 2021. [Online]. Available: https://www.research.ibm.com/articles

[230] D. Burger, "Microsoft unveils project brainwave for real-time AI," *Microsoft Res., Microsoft*, vol. 22, 2017.

[231] S. Tang. *A List of ICs and IPs for AI, Machine Learning and Deep Learning*. Accessed: Mar. 14, 2021. [Online]. Available: https://github.com/basicmi/AI-Chip

[232] B. Zhang, "Distributed SVM face recognition based on Hadoop," *Cluster Comput.*, vol. 22, no. S1, pp. 827–834, Jan. 2019.

[233] Y. Zhang, X. Xiao, L.-X. Yang, Y. Xiang, and S. Zhong, "Secure and efficient outsourcing of PCA-based face recognition," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 1683–1695, 2020.

[234] X. Zhu, Z. Lei, J. Yan, D. Yi, and S. Z. Li, "High-fidelity pose and expression normalization for face recognition in the wild," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 787–796.

[235] N. Singer and C. Metz, "Many facial-recognition systems are biased, says us study," The New York Times, New York, NY, USA, Tech. Rep., 2019, vol. 19.

[236] C. Jones, "Law enforcement use of facial recognition: Bias, disparate impacts on people of color, and the need for federal legislation," *NCJL Tech.*, vol. 22, p. 777, Oct. 2020.

[237] VumaCam. *vumacam.co.za*. Accessed: Mar. 14, 2021. [Online]. Available: https://www.vumacam.co.za/

[238] J. Buolamwini and T. Gebru, "Gender shades: Intersectional accuracy disparities in commercial gender classification," in *Proc. Conf. Fairness, Accountability Transparency*, 2018, pp. 77–91.

[239] S. Chakraborty and D. Das, "An overview of face liveness detection," 2014, *arXiv:1405.2227*.

[240] K. Patel, H. Han, and A. Jain, "Secure face unlock: Spoof detection on smartphones," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 10, pp. 2268–2283, Jun. 2016.

[241] L. Feng, L.-M. Po, Y. Li, X. Xu, F. Yuan, T. C.-H. Cheung, and K.-W. Cheung, "Integration of image quality and motion cues for face anti-spoofing: A neural network approach," *J. Vis. Commun. Image Represent.*, vol. 38, no. 1, pp. 451–460, Jul. 2016.

[242] L. Li, X. Feng, Z. Boulkenafet, Z. Xia, M. Li, and A. Hadid, "An original face anti-spoofing approach using partial convolutional neural network," in *Proc. 6th Int. Conf. Image Process. Theory, Tools Appl. (IPTA)*, Dec. 2016, pp. 1–6.

[243] K. Patel, H. Han, and A. K. Jain, "Cross-database face antispoofing with robust feature representation," in *Proc. Chin. Conf. Biometric Recognit.* Springer, 2016, pp. 611–619.

[244] P. Zhang, F. Zou, Z. Wu, N. Dai, S. Mark, M. Fu, J. Zhao, and K. Li, "Feathernets: Convolutional neural networks as light as feather for face anti-spoofing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2019, pp. 1–10.

[245] Z. Yu, C. Zhao, Z. Wang, Y. Qin, Z. Su, X. Li, F. Zhou, and G. Zhao, "Searching central difference convolutional networks for face anti-spoofing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 5295–5305.

[246] J. A. Martins, R. L. Lam, J. M. F. Rodrigues, and J. M. H. du Buf, "Expression-invariant face recognition using a biological disparity energy model," *Neurocomputing*, vol. 297, pp. 82–93, Jul. 2018.

[247] Y. Li, Y. Wang, J. Liu, and W. Hao, "Expression-insensitive 3D face recognition by the fusion of multiple subject-specific curves," *Neurocomputing*, vol. 275, pp. 1295–1307, Jan. 2018.

[248] I. M. Revina and W. R. S. Emmanuel, "Face expression recognition using LDN and dominant gradient local ternary pattern descriptors," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 33, no. 4, pp. 392–398, May 2021.

[249] K. Wang, Z. Chen, Q. M. Jonathan Wu, and C. Liu, "Illumination and pose variable face recognition via adaptively weighted ULBP_MHOG and WSRC," *Signal Process., Image Commun.*, vol. 58, pp. 175–186, Oct. 2017.

[250] C.-H. Ho, P. Morgado, A. Persekian, and N. Vasconcelos, "PIEs: Pose invariant embeddings," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 12377–12386.

[251] M. O. Oloyede, G. P. Hancke, and H. C. Myburgh, "A review on face recognition systems: Recent approaches and challenges," *Multimedia Tools Appl.*, vol. 79, nos. 37–38, pp. 27891–27922, Oct. 2020.

[252] A. H. Sable, S. N. Talbar, and H. A. Dhirbasi, "Recognition of plastic surgery faces and the surgery types: An approach with entropy based scale invariant features," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 31, no. 4, pp. 554–560, Oct. 2019.

[253] C. C. Chude-Olisah, G. Sulong, U. A. K. Chude-Okonkwo, and S. Z. M. Hashim, "Face recognition via edge-based Gabor feature representation for plastic surgery-altered images," *EURASIP J. Adv. Signal Process.*, vol. 2014, no. 1, pp. 1–15, Dec. 2014.

[254] A. S. O. Ali, V. Sagayan, A. Malik, and A. Aziz, "Proposed face recognition system after plastic surgery," *IET Comput. Vision*, vol. 10, no. 5, pp. 342–348, 2016.

[255] S. Zhao, "Pixel-level occlusion detection based on sparse representation for face recognition," *Optik*, vol. 168, pp. 920–930, Sep. 2018.

[256] Y. Fu, X. Wu, Y. Wen, and Y. Xiang, "Efficient locality-constrained occlusion coding for face recognition," *Neurocomputing*, vol. 260, pp. 104–111, Oct. 2017.

[257] Y. Long, F. Zhu, L. Shao, and J. Han, "Face recognition with a small occluded training set using spatial and statistical pooling," *Inf. Sci.*, vols. 430–431, pp. 634–644, Mar. 2018.

DAVID WANYONYI received the M.Sc. degree in computer science from the University of the Western Cape, South Africa. He is currently pursuing the Ph.D. degree in electrical and information engineering in the field of never ending face recognition with deep learning with the University of the Witwatersrand, South Africa. He is with the Department of Information Communication Technology (Wits ICT), University of the Witwatersrand.

TURGAY CELIK received the second Ph.D. degree from the University of Warwick, Coventry, U.K., in 2011. He is currently a Professor of digital transformation and the Director of the Wits Institute of Data Science, University of Witwatersrand, Johannesburg, South Africa. His research interests include computer vision, (explainable) artificial intelligence, (health) data science, data-driven optimal control, and remote sensing. He is an Associate Editor of *BMC Medical Informatics and Decision Making*, *IET ELL*, IEEE Access, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS), and *Signal, Image and Video Processing* (Springer).