A Comprehensive Survey of Lightweight Neural Networks for Face Recognition

Yongli Zhang · Jaekyung Yang[†]

Department of Industrial and Information Systems Engineering, Jeonbuk National University

얼굴 인식을 위한 경량 인공 신경망 연구 조사

장영립 · 양재경[†]

전북대학교 산업정보시스템공학과

Lightweight face recognition models, as one of the most popular and long-standing topics in the field of computer vision, has achieved vigorous development and has been widely used in many real-world applications due to fewer number of parameters, lower floating-point operations, and smaller model size. However, few surveys reviewed lightweight models and reimplemented these lightweight models by using the same calculating resource and training dataset. In this survey article, we present a comprehensive review about the recent research advances on the end-to-end efficient lightweight face recognition models and reimplement several of the most popular models. To start with, we introduce the overview of face recognition with lightweight models. Then, based on the construction of models, we categorize the lightweight models into: (1) artificially designing lightweight FR models, (2) pruned models to face recognition, (3) efficient automatic neural network architecture design based on neural architecture searching, (4) Knowledge distillation and (5) low-rank decomposition. As an example, we also introduce the SqueezeFaceNet and EfficientFaceNet by pruning SqueezeNet and EfficientNet. Additionally, we reimplement and present a detailed performance comparison of different lightweight models on the nine different test benchmarks. At last, the challenges and future works are provided. There are three main contributions in our survey: firstly, the categorized lightweight models can be conveniently identified so that we can explore new lightweight models for face recognition; secondly, the comprehensive performance comparisons are carried out so that ones can choose models when a state-of-the-art end-to-end face recognition system is deployed on mobile devices; thirdly, the challenges and future trends are stated to inspire our future works.

Keywords : Face Recognition, Lightweight Networks, Mobile Device, Efficient Face Models

1. Introduction

With the great advance of Deep Convolution Neural Networks (DCNNs) techniques [26, 31, 36, 37, 41, 62, 67] and large-scale datasets [13, 58] in the field of computer vision and image understanding, the DCNNs-based techniques have become an extensive research topic in Face Recognition (FR) tasks and also have been widely used in many real-world applications. Therefore, more and more novel and efficient FR DCNNs [9, 55, 60, 63, 64, 65, 66, 68] and the margin-based loss functions [14, 35, 43, 73] were designed and replaced traditional methods [4, 10, 16, 27, 38, 42, 53, 71, 77, 83], becoming the mainstream methods of FR. However, the large floating-point operations (FLOPs), large number of parameters, and huge models size lead to high computational complex-

ity, making it difficult to deploy the DCNNs models on the Internet of Things (IoT) or mobile devices with limited memory in practical applications [9, 60, 68], such as video surveillance, law enforcement, access control, marketing, smartphones, embedded systems, wearable devices, etc.. And large-scale face datasets with the pose changes, illumination changes, low resolution, and motion blur will also lead to a challenge in recognition accuracy [15, 49]. Therefore, to solve these problems, some scholars are working on efficient and effective lightweight models with fewer parameters, lower FLOPs, and smaller model sizes.

To get lightweight FR models, many efforts have been devoted to keep an optimal trade-off between accuracy and efficiency. So, it is necessary to review recent lightweight FR models to inspire the development and application of lightweight FR. Specifically, there have been certain surveys [1, 2, 17, 34, 49, 50, 72, 87] about FR, which review almost all FR DCNNs models. However, they do not cover recent published lightweight FR models; and some of them focus on some specific tasks; besides, few surveys reimplement reviewed lightweight models. And, to the best of our knowledge, there has been only one article [49] that fulfilled relevant survey for lightweight FR models, which provided a benchmark of lightweight FR. In summary, the end-to-end lightweight FR models need to be systematically reviewed from a variety of different perspectives, however, few of the existing surveys attach importance to this task. So, our survey will focus on reviewing lightweight FR models the most needed in practical applications. Different from [49], this article will classify the existing state-of-the-art end-to-end lightweight models into five categories and reimplement several mainstream models. The main contributions can be summarized as follows:

Firstly, we categorize the lightweight FR models into five aspects to further explore new lightweight FR models.

Secondly, we introduce the SqueezeFaceNet and EfficientFaceNet by pruning SqueezeNet and EfficientNet and reimplement several most popular lightweight FR models. Meanwhile, the comprehensive performance comparisons are also presented.

Thirdly, we present some the challenges and future trends to inspire our future works.

The remainder of the paper is organized as follows. Section 2 briefly reviews what are the FR tasks and lightweight models, and the categories of existing state-of-the-art end-to-end lightweight models are also described. As an example, we introduce

the SqueezeFaceNet and EfficientFaceNet by pruning SqueezeNet and EfficientNet in Section 3. In Section 4, we reimplement and present a detailed performance comparison of different lightweight models on nine different test benchmarks. In Section 5, some challenges and future works are presented. Section 6 concludes this work.

2. Face Recognition and Lightweight Model

2.1 Face Recognition

An automatic face recognition system, which consisted of four main components: face detection, face alignment, feature extraction and face matching, aims at implementing two different tasks, namely, one-to-one (1:1) Face Verification or one-to-many (1:N) Face Identification [17]. Face verification refers to judging whether two face images belong to the same identity or not, which does not need to know the identity of the image. It is essentially a binary classification problem, and is usually used in scenarios such as witness comparison and identity verification. Face identification refers to judging the identity of probe face descriptor according to a registered face gallery. It is essentially a multi-classification problem. Common application scenarios include access control systems, venue sign-in systems and so on.

In view of the broad application prospects of FR, a series of traditional FR methods sprang up from the early 1990s, such as Eigenfaces [71], Fisherfaces [4], Bayesian eigenfaces [53], Laplacianfaces [27], Sparse Representation [16, 77, 83], Gabor feature [42] and learning-based descriptor [10, 38]. However, the above methods either fail to address the uncontrolled facial changes, or lack of distinctiveness and compactness, or shallow representations [72]. To address above problems, based on the LeNet [37] which was applied to handwritten digit recognition in 1990, and AlexNet [37], NiN [41], VGG [62], GoogleNet [67], ResNet [26], DenseNet [31], which won the ImageNet competition champion, many FR DCNNs models have achieved earth-shaking changes, such as Facenet [60], Deepface [68], DeepID seiers [63, 64, 65, 66], VGGFace [55] and VGGFace2 [9]. Especially, various novel margin-based loss functions, ArcFace [14] SphereFace [43] Cosface [73] Adaface [35], also greatly promote the recognition performance.

A standard pipeline of automatic FR system is shown in the <Figure 1>. When the system gets still images or video
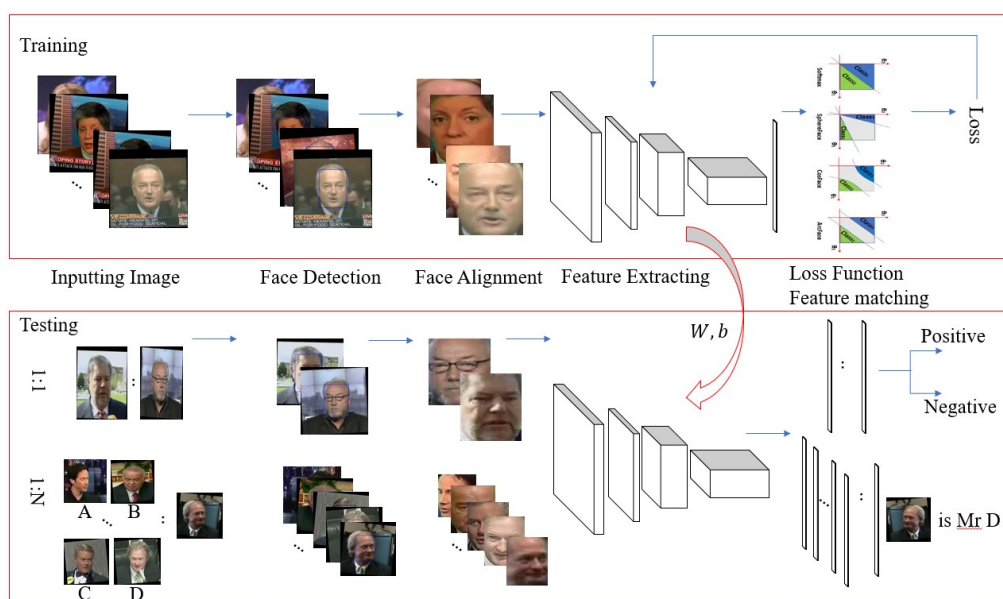
frames as input, face detection locates the face regions of input; then, located face is calibrate and resize normalized pixel in the face alignment stage; additionally, the feature extraction stage extracts the discriminative features using DCNNs; at last, face verification or face identification task is conducted in the face matching stage.

However, it is difficult to deploy these FR DCNNs on the IoT or mobile devices with limited memory due to large parameters, FLOPs, and model size. So, implementing optimal trade-off between accuracy and efficiency is becoming more and more important.

2.2 Lightweight Model

Considering the huge demand for deploying FR models on IoT or mobile devices with limited memory, the light-weight FR models have been researched to keep an optimal trade-off between performance and efficiency in recent years. And these lightweight FR models are also subject to standard pipeline of automatic FR system shown in the <Figure 1>. Nowadays, there have been a variety of lightweight archi-tectures achieving state of the art (SOTA) performance, which can be categorized into: (1) artificially designing lightweight FR models (ADLM), (2) pruned models to face recognition (PM), (3) efficient automatic neural network architecture de-sign based on neural architecture searching (ANND_NAS), (4) Knowledge distillation (KD) and (5) Low-rank decom-position (LRD). The <Table 1> lists the categorization of lightweight FR models. We mainly present the name or author of models, input size, MFLOPs, the number of parameters, model size, accuracy on LFW [32] and published years to facilitate comparison.

<Figure 1> The Standard Pipeline of End-to-End Lightweight FR System

<Table 1> The Categorization of Lightweight FR Models

Category	Models	Input	MFLOPs	Para(M)	ModelSize(MB)	LFW	Years
ADLM	LightCNN-4	128×128	1500	4.095	-	97.97	
	LightCNN-9 [78]	128×128	1000	5.556	-	98.80	2018
	LightCNN-29	128×128	3900	12.637	-	99.33	
	ConvFaceNeXt [29]	112×112	410.59	4.49	4.49	99.30	2022
	MobiFace [19]	112×112	-	-	9.3	99.72	2019
	Airface [39]	112×112	1000	-	-	99.27	2019
	ELANet [84]	112×112	550	1.61	-	99.68	2022
	MaYidong [47]	112×112	1000	-	20	99.80	2019
	LWDistill-DSE-LSE [44]	112×112	-	1.35	5.36	99.67	2021

<Table 1> The Categorization of Lightweight FR Models(Continued)

Category	Models	Input	MFLOPs	Para(M)	ModelSize(MB)	LFW	Years
PM	MobileFaceNetV1 [49]	112×112	1145.89	3.43	13.7	99.4	2021
	MobileFaceNets [11]	112×112	473.15	1.03	4.5	99.55	2021
	ShuffleFaceNet 0.5x	112×112	66.9	0.5	1.9	99.07	2019
	ShuffleFaceNet 1x [48]	112×112	275.8	1.4	5.6	99.45	
	ShuffleFaceNet 1.5x	112×112	557.5	2.6	10.5	99.67	
	ShuffleFaceNet 2x	112×112	1050	4.5	18	99.68	
	SqueezeFaceNet	112×112	1024	0.663	4.18	99.47	2023
	EfficientFaceNet	112×112	3987	0.09	15.87	99.50	2023
	MixFaceNet -s [6]	112×112	161.9	1.04	-	99.60	2021
	VarGFaceNet [80]	112×112	1022	5	-	99.68	2019
	Xiao [79]	112×112	-	-	3.4	98.52	2021
	QuantFace [5]	112×112	-	-	1.1	99.35	2022
	LightQNet [12]	96×96	77	0.45	-	99.86	2021
	Lyu [45]	112×112	-	-	19.8	99.80	2019
	LD-MobileFaceNe [24]	112×112			3	97.02	2019
	Fu model-6 [21]	-	570	0.26	-	-	2022
	SqueezeFacePoseNet [3]	113×113		1.24	4.41	-	2021
NAS	ProxylessFaceNAS [49]	112×112	873.95	3.01	12.4	99.2	2021
	PocketNet [7]	112×112	587.11	0.925	-	99.58	2022
	NAS-C [89]	-	-	-	16	99.89	2020
	TNAS [75]	144×144	550	-	-	99.63	2021
	GP-NAS [40]	-	151	0.83	-	99.17	2020
KD	EC-KD-(2x)128 [74]	144×144	1440	-	16	99.34	2020
	ShrinkTeaNet-MFNR [18]	112×112	-	3.73	-	99.77	2019
	SKD [22]	16×16	2.43	0.79	-	95.03	2018
	Triplet distillation-s [20]	112×112	230	0.59	2.3	99.27	2020
	Bridge Distillation [23]	16×16	1.51	0.11	-	85.46	2020
LRD	SILR [81]	32×32	-	-	-	82.98	2021
	LRRNet [86]	80×60	-	-	-	-	2017

For the first class, the ADLM means the researcher artificially designed an efficient lightweight RF model that keeps an optimal trade-off between performance and efficiency. Wu et al. developed a LightCNN [78] family architectures, which are LightCNN-4, LightCNN-9, LightCNN-29, respectively, to learn a robust face representation on a noisy labeled dataset. There are 4.095M, 5.556M, 12.637M parameters, and about 1500 MFLOPs, 1000 MFLOPs, 3900 MFLOPs in the LightCNN-4, LightCNN-9, LightCNN-29, respectively. ConvFaceNeXt [29] is designed by Hoo et al., which stacks the stem, bottleneck, and embedding partitions to construct a ConvFaceNeXt family. The largest model contains 1.05M parameters and 410.59 MFLOPs.

For the second class, the PM means the researcher tailored FR models by pruning commonly used SOTA lightweight networks, which include the MobileNet series [30, 59], ShuffleNetV2 [46], Mixconv [69], VarGNet [85] and so on, to construct MobileFaceNetV1 [49], MobileFaceNets [11], ShuffleFaceNet [48], MixFaceNet [6], VarGFaceNet [80] and so on. Based on the MobileNetV2 [59], MobileFaceNetV1

[49] and MobileFaceNets [11] are designed. The MobileFaceNets adopts a typical reverse residual block, which needs to train about 1.03M parameters with 473.15M FLOPs. And a family of ShuffleFaceNet [48] are built based on the ShuffleNetV2, in which the smallest is ShuffleFaceNet0.5× with about 0.5M parameters and 66.9M FLOPs, but it does not keep significant accuracy. Therefore, ShuffleFaceNet1× is commonly used, which contains about 1.4M parameters and 275.8M FLOPs.

For the third class, the ANND_NAS means to directly learn neural network architectures for FR based on the neural architecture searching (NAS) [90]. The article [49] introduces a modified version of the ProxylessFaceNAS based on the ProxylessNAS [8], and there are 3.01M parameters, and about 873.95 MFLOPs. Boutros proposed a family of PocketNet based on the NAS and knowledge distillation [28], which contains 0.925M parameters and 587.11 MFLOPs.

For the fourth class, the Knowledge distillation (KD) [57, 82], training a small student network based on the large teacher network, is to train a compact neural network which can reimple-

ment the output of the large networks. The KD-based EC-KD [74] and ShrinkTeaNet [18] can be found in <Table 1>.

For the fifth class, the Low-rank decomposition (LRD) refers to use the low rank matrix to approximate the weight matrix of DCNNs, getting compressed lightweight DCNNs. The SILR [81] and LRRNet [86] can be found in <Table 1>.

These SOTA models aim to improve the compactness and computational efficiency, overcoming the difficulty of deployment due to trainable massive parameters. Meanwhile, as shown in <Table 1>, we also simply prune the SqueezeNet [33] and EfficientNet [70] to construct SqueezeFaceNet and EfficientFaceNet, and the details are introduced in section 3.

3. SqueezeFaceNet and EfficientFaceNet

As the simple supplement of the PM, we introduce the SqueezeFaceNet and EfficientFaceNet based on the SqueezeNet [33] and EfficientNet [70] in this section.

3.1 SqueezeFaceNet

SqueezeFaceNet is simply pruned from the SqueezeNet [33], which consists of fire modules and a Global Depthwise Convolution (GDC) [11] layer. The fire modules [33] extracts useful features and GDC treats different units with different importance [11]. At last, we adopt novel margin-based ArcFace [14] as loss function. The <Table 2> shows the architecture of SqueezeFaceNet.

<Table 2> SqueezeFaceNet Architectures

Input	Operator	Channel	Kernel	Stride
112×112×3	Conv1_3x3	96	3	1
112×112×96	MaxPooling1	-	2	2
56×56×96	Fire2	128	-	-
56×56×128	Fire3	128	-	-
56×56×128	Fire4	256	-	-
56×56×256	MaxPooling2	-	2	2
28×28×256	Fire5	256	-	-
28×28×256	Fire6	384	-	-
28×28×384	Fire7	384	-	-
28×28×384	Fire8	512	-	-
28×28×512	MaxPooling3	-	2	2
14×14×512	Fire9	512	-	-
14×14×512	AvgPooling	-	2	2
7×7×512	GDConv_7x7	512	7	1

<Table 3> EfficientFaceNet-s Architectures

Input	Operator	Channel	Kernel	Stride
112×112×3	Conv_3x3	32	3	2
56×56×32	MBConv1	16	-	-
56×56×16	MBConv6	24	-	-
28×28×24	MBConv6	40	-	-
14×14×40	MBConv6	80	-	-
7×7×80	MBConv6	112	-	-
7×7×112	MBConv6	192	-	-
3×3×192	MBConv6	512	-	-
3×3×512	GDConv_7x7	512	3	1

3.2 EfficientFaceNet

EfficientFaceNet is simply pruned from the EfficientNet [70], which consists of MBConv1, MBConv6 modules and a GDC [11] layer. The MBConv [70] extracts useful features and GDC treats different units with different importance [11]. At last, we adopt novel margin-based ArcFace [14] as loss function. The <Table 3> shows the architecture of EfficientFaceNet.

4. Performance Comparison

This section comprehensively presents the performance of the most common models which are listed <Table 1>. We reimplement the results of EfficientFaceNet, SqueezeFaceNet, MixFaceNet, ShuffleFaceNet0.5x, MobileFaceNet, LightCNN-9 to make fair comparison by using the same computing resources, train dataset and setting the same hyper-parameters. All codes refer to the original papers and ArcFace[2] [14].

4.1 Experimental Settings

Train Datasets. We use the MS1MV3 [15] dataset containing approximately 93K identities and 5.2M images as the train dataset in our experiments, which is semi-automatically cleaned from MS1MV0 (about 10M images of 100K identities) [14] and which is also an enhanced version of the MS1MV2 (about 5.8M images of 85K identities) [25].

Test Datasets. In order to systematically compare the listed lightweight methods, there are nine challenging test datasets. The details are listed in <Table 4>. In the test stage, 1:N Face Identification are conducted on IJB-B and IJB-C, and 1:1 Face Verification are done on the other datasets, IJB-B and IJB-C.

2) https://github.com/deepinsight/insightface

<Table 4> Benchmarks Used for Testing on The Different Face Recognition Scenarios

Datasets	Images/Videos	Identities	Task	Metric	Pairs(P/N)	Type
IJB-B [76]	21.8K/7011	1845	1:1/1:N	TAR@FAR/ Rank-1/5	10K/8M	Large-scale with Unconstrained image
IJB-C [51]	31.3K/11779	3531	1:1/1:N	TAR@FAR/ Rank-1/5	19.5K/15.6M	Large-scale with Unconstrained image
LFW [32]	13233	5749	1:1	Acc	3K/3K	Unconstrained image
CA-LFW [87]	13233	5749	1:1	Acc	3K/3K	Cross-Age
CP-LFW [87]	13233	5749	1:1	Acc	3K/3K	Cross-Pose
AgeDB-30 [54]	16488	568	1:1	Acc	3K/3K	Cross-Age
CFP-FP [61]	7000	500	1:1	Acc	3.5K/3.5K	Cross-Pose
CFP-FF [61]	7000	500	1:1	Acc	3.5K/3.5K	Front-Front
VGG2-FP [9]	9000	1800	1:1	Acc	2.5K/2.5K	Cross-Pose

Implementation Details. In our experiments, all training details are following the ArcFace [14]. The batch-size is 128, and the optimizer is SGD with a learning rate of 0.1, a momentum of 0.9, and a weight decay of 1e-4. The train epochs are 40, and the scale parameter s and margin value of ArcFace loss are set to 64 and 0.5, respectively. We use a Linux machine (Ubuntu 18.04.1 LTS) with Intel(R) Core(TM) i9-9900KS CPU @ 4, 32 G RAM, and 3 Nvidia GeForce RTX 2060 (6Gb) GPUs to finish all experiments. Meanwhile, the experiments are implemented by PyTorch v1.11.0 [56] and mixed-precision [52] is employed to save GPUs memory and accelerate training.

4.2 Evaluation and Comparison

The 1:N Face Identification and 1:1 Face Verification are conducted in this section. All the experiments not only compare the number of parameters, model size and FLOPs of different models, but also evaluate the recognition accuracy. The parameters, model size and MFLOPs are the less the better, on the contrary, the accuracies are the higher the better. As shown in the <Table 5>, <Table 6> and <Table 7>, which all are divided into two parts by a dotted line, the upper data of dotted line refer to related papers, and the bottom data of dotted line come from our reimplemented experiments.

<Table 5> 1:1 Face Verification Performance (%) on 7 Different Datasets

Model	Params.(M)	MFLOPs	ModelSize(M)	LFW	CA-LFW	CP-LFW	CFP-FF	CFP-FP	AgeDB-30	VGG2-FP
EfficientNet [70, 84]	6.58	1140	-	99.53	95.78	90.92	99.50	96.32	-	94.32
VarGFaceNet [49, 80]	5.0	1022	-	99.70	95.15	88.55	99.50	96.90	-	-
MobileFaceNetV1 [29, 49]	3.43	1145.89	13.7	99.02	93.50	85.38	98.96	87.64	92.95	88.14
ProxylessFaceNAS [49, 84]	3.2	900	12.4	99.20	92.55	84.17	98.80	94.70	-	-
ShuffleFaceNet 1.5x [29, 48]	2.6	577.5	10.5	99.03	93.23	85.28	99.03	87.97	92.18	88.20
MobileNet-v2 [59, 84]	2.26	430	-	99.55	95.34	89.43	99.48	93.17	-	91.58
PocketNetM-256 [7]	1.75	1099.15	-	99.58	95.63	90.03	-	95.66	97.17	-
ELANet [84]	1.61	550	-	99.68	96.07	92.17	99.76	97.16	-	94.66
ShuffleFaceNet 1x [29, 48]	1.46	278.47	5.6	98.87	93.13	84.47	98.84	86.44	92.33	87.08
LW+DSE+LSE [44]	1.35	-	5.36	99.40	95.50	89.00	99.50	93.33	96.03	-
QuantFace(w8a8) [5]	1.1	-	-	99.35	94.78	87.73	-	-	94.37	-
ConvFaceNeXt_TE [29]	1.05	410.59	4.49	99.05	93.42	85.58	99.07	88.21	93.15	89.36
ShuffleMixFaceNet-XS [6]	1.04	161.9	-	99.53	-	-	-	-	95.62	-
MobileFaceNets [11, 84]	1.03	473.15	4.5	99.53	95.42	90.34	99.55	95.26	-	92.49
EfficientFaceNet_0*	3.987	90	15.87	99.50	94.81	88.17	99.41	94.19	95.15	89.70
SqueezeFaceNet*	1.024	663	4.18	99.47	94.45	87.62	99.47	93.63	95.20	90.84
MixFaceNet*	1.040	84	4.29	99.45	95.48	89.43	99.61	95.56	96.17	91.62
ShuffleFaceNet0.5x*	0.526	36	2.22	98.81	92.97	86.37	98.33	89.76	90.53	86.56
MobileFaceNet*	1.201	230	4.89	99.67	95.60	90.20	99.70	96.46	96.83	92.84
LightCNN-9*	5.556	1000	17.59	99.30	94.80	87.55	99.54	93.59	94.82	89.56

<Table 6> 1:1 Face Verification Performance (%) on IJB-B and IJB-C

Model	Params.(M)	MFLOPs	ModelSize(M)	IJB-C TAR@ FAR=			IJB-B TAR@ FAR=		
				1e-06	1e-05	1e-04	1e-06	1e-05	1e-04
VarGFaceNet [49, 80]	5.0	1022	-	-	91.6	94.7	-	87.7	92.9
MixFaceNet-M [6]	3.95	626.1	-			93.42			91.55
MobileFaceNetV1 [49]	3.43	1145.89	13.7	-	92.2	94.7	-	87.9	92.8
ProxylessFaceNAS [49]	3.2	900	12.4	-	83.0	89.7	-	75.8	87.1
ShuffleFaceNet 1.5x [29, 48]	2.6	577.5	10.5	-	70.06	81.36	-	63.54	78.17
PocketNetM-256 [7]	1.75	1099.15	-			91.33			89.31
ELANet [84]	1.61	550	-	-	-	94.4	-	-	92.7
ShuffleFaceNet 1x [29, 48]	1.46	278.47	5.6	-	71.49	81.62	-	64.41	77.96
ConvFaceNeXt_TE [29]	1.05	410.59	4.49	-	75.11	84.21	-	67.45	80.61
ShuffleMixFaceNet-XS [6]	1.04	161.9	-	-	-	90.43	-	-	87.86
MobileFaceNets [11, 49]	1.03	473.15	4.5	-	90.4	93.9	-	85.7	92.0
EfficientFaceNet_0*	3.987	90	15.87	80.32	87.03	92.03	33.03	81.18	89.81
SqueezeFaceNet*	1.024	663	4.18	69.10	81.78	89.57	32.47	72.93	86.79
MixFaceNet*	1.040	84	4.29	84.53	89.43	93.22	30.56	84.51	91.16
ShuffleFaceNet0.5x*	0.526	36	2.22	71.78	79.49	86.88	29.95	84.16	90.24
MobileFaceNet*	1.201	230	4.89	86.25	90.69	93.84	32.67	85.48	91.79
LightCNN-9*	5.556	1000	17.59	76.63	84.63	90.17	32.55	78.80	87.68

<Table 7> 1:N Face Identification performance (%) on IJB-B and IJB-C

Model	Params.(M)	MFLOPs	ModelSize(M)	IJB-C		IJB-B	
				Rank-1	Rank -5	Rank -1	Rank -5
VarGFaceNet [49, 80]	5.0	1022	-	95.3	96.9	94.0	96.5
MobileFaceNetV1 [49]	3.43	1145.89	13.7	94.7	96.4	93.2	95.8
ProxylessFaceNAS [49]	3.2	900	12.4	91.8	94.8	90.7	94.4
ShuffleFaceNet [49]	2.6	577.5	10.5	95.0	96.7	93.6	96.2
MobileFaceNets [49]	1.03	473.15	4.5	95.3	96.9	93.9	96.5
SqueezeFaceNet*	1.024	663	4.18	92.25	95.01	91.15	94.49
MixFaceNet*	1.040	84	4.29	94.67	96.53	93.58	95.80
MobileFaceNet*	1.201	230	4.89	95.17	96.80	94.01	96.20
LightCNN-9*	4.503	812	17.59	92.36	95.18	91.08	94.48

1:1 Face Verification Evaluation Results.

A series of 1:1 face verification experiments are conducted on the datasets of LFW, CA-LFW, AgeDB-30, CP-LFW, CFP-FP, CFP-FF and VGG2-FP, and the performances are reported as the accuracy of 10-fold cross-validation. The referred data and reimplemented data are all listed into the <Table 5>.
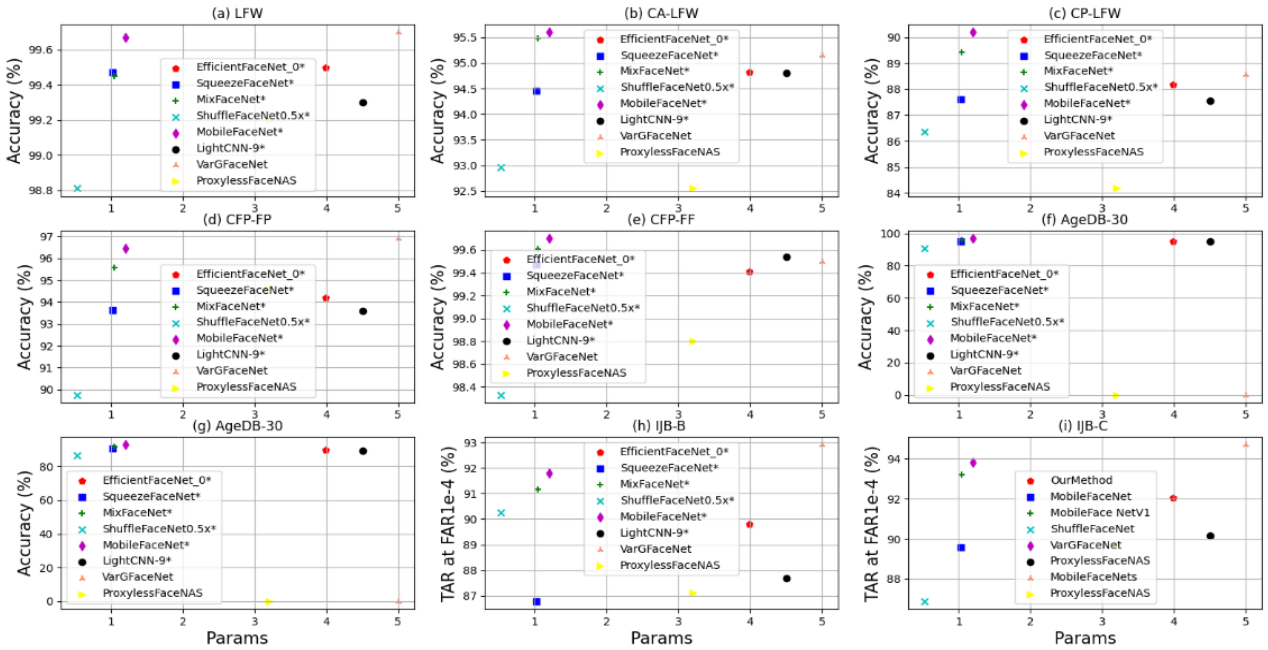
For better investigating performance of lightweight models on two IJB datasets which combine high- or low-quality images and video frames, a series of one-to-one (1:1) face verification experiments are conducted on both IJBs. TAR@FAR are reported as metrics of face verification, which is the higher the better. The referred data and reimplemented data are all listed into <Table 6>.

1:N Face Identification Evaluation Results.

A series of 1:N face identification experiments are conducted on both IJBs. Rank-1 and Rank-5 accuracy are reported as metrics of face identification, which are the higher the better. The referred data and reimplemented data are all listed into <Table 7>.

Seen from <Table 5>, <Table 6> and <Table 7>, generally speaking, more trainable parameters mean higher accuracy, but also higher computational complexity and larger model size. Meanwhile, we also can find there are some small and refined models, such as MobileFaceNet, ELANet, MixFaceNet and so on.

〈Figure 2〉 The Number of Parameters (M) of The Models vs. Performance

4.3 Performance vs. the Number of Parameters

This section plots the number of parameters vs. the performance of lightweight models to present the trade-off between the performance and the number of parameters. As shown in the 〈Figure 2〉, the x-axis of each subplots from (a) to (i) represents the LFW (accuracy), CALFW(accuracy), CPLFW (accuracy), CFP-FF (accuracy), CFP-FP (accuracy), AgeDB-30 (accuracy), VGG2-FP (accuracy), IJB-C (TAR at FAR1e-4), and IJB-B (TAR at FAR1e-4), and the y-axis is the number of parameters. Different methods are highlighted with different markers and different colors. The models locating the conner of the left upper in every subplot indicate high performance and less number of parameters. We can find the MobileFaceNet and MixFaceNet report much better trade-off.

5. Challenge and Future Work

Although the lightweight FR has achieved remarkable performance, there are still many challenges. In my opinion, one challenge is how to design small and refined lightweight model which keeps optimal trade-off between performance and efficiency. In addition, how to response diverse train and test datasets which include large facial pose, extreme expression, occlusion, facial scale, motion blur, low illumination, large-scale. Finally, the interpretability of the lightweight model is also anther worthy challenge.

So, our future work mainly focuses on addressing the above challenges. For designing lightweight model, we will consider designing lightweight model, or using automatic machine learning methods like NAS to search for networks. For diverse train and test datasets, we will consider the cross age, pose, race, and so on. For interpretability of the lightweight model, we mainly consider explanation in spatial and scale dimensions.

6. Discussion

In this survey, we comprehensively reviewed the recent lightweight models for face recognition. Firstly, a standard pipeline of automatic lightweight FR system was presented. Secondly, we categorized the listed lightweight models into ADLM, PM, ANND_NAS, KD and LRD according to the different designing mode of models. Thirdly, we introduceed the SqueezeFaceNet and EfficientFaceNet by pruning SqueezeNet and EfficientNet. Fourthly, we reimplemented the results of EfficientFaceNet, SqueezeFaceNet, MixFaceNet, ShuffleFaceNet0.5x, MobileFaceNet, LightCNN-9 to eliminate the influence of different computing resources, train data-

set and setting hyper-parameters. At last, these results can be seen as benchmark for comparison and direct reference. This survey inspires the future work and indicates that the future models should be small and refined, can response diverse dataset, and have interpretability.

References

[1] Abate, A.F., Nappi, M., Riccio, D., and Sabatino, G., 2D and 3D face recognition: A survey, *Pattern Recognition Letters*, 2007,Vol.28,No.14,pp.1885-1906.

[2] Adjabi, I., Ouahabi, A., Benzaoui, A., and Taleb-Ahmed, A., Past, present, and future of face recognition: A review, *Electronics*, 2020, Vol. 9, No. 8, pp. 1188.

[3] Alonso-Fernandez, F., Barrachina, J., Hernandez-Diaz, K., and Bigun, J., SqueezeFacePoseNet: Lightweight Face Verification Across Different Poses for Mobile Platforms, *International Conference on Pattern Recognition*, 2021, pp. 139-153.

[4] Belhumeur, P.N., Hespanha, J.P., and Kriegman, D.J., Eigenfaces vs. fisherfaces: Recognition using class specific linear projection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, Vol, 19, No. 7, pp. 711-720.

[5] Boutros F., Damer, N., and Kuijper, A., QuantFace: Towards lightweight face recognition by synthetic data low-bit quantization, arXiv preprint arXiv:2206.10526, 2022, pp. 855-862.

[6] Boutros, F., Damer, N., Fang, M., Kirchbuchner, F., and Kuijper, A., Mixfacenets: Extremely efficient face recognition networks, *2021 IEEE International Joint Conference on Biometrics (IJCB)*, 2021, pp. 1-8.

[7] Boutros, F., Siebke, P., Klemt, M., Damer, N., Kirchbuchner, F., and Kuijper, A., PocketNet: extreme lightweight face recognition network using neural architecture search and multistep knowledge distillation, *IEEEAccess*, 2022, pp. 46823-46833.

[8] Cai, H., Zhu, L., and Han, S., Proxylessnas: Direct neural architecture search on target task and hardware, arXiv preprint arXiv:1812.00332, 2018.

[9] Cao, Q., Shen, L., Xie, W., Parkhi, O. M., and Zisserman, A., Vggface2: A dataset for recognising faces across pose and age, *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, 2018.

[10] Cao, Z., Yin, Q., Tang, X., and Sun, J., Face recognition with learning-based descriptor, *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010, pp. 2707-2714.

[11] Chen S., Liu, Y., Gao, X., and Han, Z., Mobilefacenets: Efficient cnns for accurate real-time face verification on mobile devices, *Chinese Conference on Biometric Recognition*, 2018, pp. 428-438.

[12] Chen, K., Taihe, Y., and Qi, L., LightQNet: Lightweight Deep Face Quality Assessment for Risk-Controlled Face Recognition, *IEEE Signal Processing Letters*, 2021, Vol. 28, pp. 1878-1882.

[13] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., and Fei-Fei, L., Imagenet: A large-scale hierarchical image database, *2009 IEEE Conference on Computer Vision and Pattern Recognition,* IEEE, 2009, pp. 248-255.

[14] Deng, J., Guo, J., Xue, N., and Zafeiriou, S., Arcface: Additive angular margin loss for deep face recognition, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,* 2019, 4690-4699.

[15] Deng, J., Guo, J., Zhang, D., Deng, Y., Lu, X., and Shi, S., Lightweight face recognition challenge, *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops.* 2019.

[16] Deng, W., Hu, J., and Guo, J., Extended SRC: Undersampled face recognition via intraclass variant dictionary, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, Vol. 34, No. 9, pp. 1864-1870.

[17] Du, H., Shi, H., Zeng, D., Zhang, X.P., and Mei, T., The elements of end-to-end deep face recognition: A survey of recent advances, *ACM Computing Surveys (CSUR)*, 2022, Vol.54, No.10s, pp. 1-42.

[18] Duong, C.N., Luu, K., Quach, K.G., and Le, N., Shrinkteanet: Million-scale lightweight face recognition via shrinking teacher-student networks, arXiv preprint arXiv:1905.10620, 2019.

[19] Duong, C.N., Quach, K.G., Jalata, I., Le, N., and Luu, K., Mobiface: A lightweight deep learning face recognition on mobile devices, *2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, 2019, pp. 1-6.

[20] Feng, Y., Wang, H., Hu, H.R., Yu, L., Wang, W., and Wang, S., Triplet distillation for deep face recognition, *2020 IEEE International Conference on Image Processing (ICIP)*, 2020, pp. 808-812.

[21] Fu, C., Zhou, X., He, W., and He, R., Towards

Lightweight Pixel-Wise Hallucination for Heterogeneous Face Recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.

[22] Ge, S., Zhao, S., Li, C., and Li, J., Low-resolution face recognition in the wild via selective knowledge distillation, *IEEE Transactions on Image Processing*, 2018, Vol.28, No. 4, pp. 2051-2062.

[23] Ge, S., Zhao, S., Li, C., Zhang, Y., and Li, J., Efficient low-resolution face recognition via bridge distillation, *IEEE Transactions on Image Processing*, 2020, Vol.29, pp. 6898-6908.

[24] Guo, L., Bai, H., and Zhao, Y., A lightweight and robust face recognition network on noisy condition, *2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 2019, pp. 1964-1969.

[25] Guo, Y., Zhang, L., Hu, Y., He, X., and Gao, J., Ms-celeb-1m: A dataset and benchmark for large-scale face recognition, *European Conference on Computer Vision*, 2016, pp. 87-102.

[26] He, K., Zhang, X., Ren, S., and Sun, J., Deep residual learning for image recognition, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770-778.

[27] He, X., Yan, S., Hu, Y., Niyogi, P., and Zhang, H.J., Face recognition using laplacianfaces, *IEEE transactions on Pattern Analysis and Machine Intelligence*, 2005, Vol. 27, No. 3, pp. 328-340.

[28] Hinton, G., Vinyals, O., and Dean, J., Distilling the knowledge in a neural network, arXiv preprint arXiv:1503. 02531 2.7, 2015.

[29] Hoo, S.C., Ibrahim, H., and Suandi, S.A., ConvFaceNeXt: Lightweight Networks for Face Recognition, *Mathematics*, 2022, Vol. 10, No. 19.

[30] Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H., Mobilenets: Efficient convolutional neural networks for mobile vision applications, arXiv preprint arXiv:1704. 04861, 2017.

[31] Huang, G., Liu, Z., VanDer Maaten, L., and Weinberger, K.Q., Densely connected convolutional networks, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 4700-4708.

[32] Huang, G.B., Mattar, M., Berg, T., and Learned-Miller, E., Labeled faces in the wild: A database forstudying face recognition in unconstrained environments.

Workshop on faces in'Real-Life'Images: detection, alignment, and recognition, 2008.

[33] Iandola, F.N., Han, S., Moskewicz, M.W., Ashraf, K., Dally, W.J., and Keutzer, K., SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size, arXiv preprint arXiv:1602.07360, 2016

[34] Jafri, R. and Arabnia, H.R., A survey of face recognition techniques, J*ournal of Information Processing Systems*, 2009, Vol.5, No.2, pp. 41-68.

[35] Kim, M., Jain, A.K., and Liu, X., AdaFace: Quality Adaptive Margin for Face Recognition, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 18750-18759.

[36] Krizhevsky, A., Sutskever, I., and Hinton, G. E., Imagenet classification with deep convolutional neural networks, *Communications of the ACM*, 2017, Vol. 60, No. 6, pp. 84-90.

[37] LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W., and Jackel, L., Handwritten digit recognition with a back-propagation network, *Advances in Neural Information Processing Systems*, 1989, Vol. 2,pp. 396-404.

[38] Lei, Z., Pietikäinen, M., and Li, S.Z., Learning discriminant face descriptor, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, Vol. 36, No. 2, pp. 289-302.

[39] Li, X., Wang, F., Hu, Q., and Leng, C., Airface: Lightweight and efficient model for face recognition, *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, 2019.

[40] Li, Z., Xi, T., Deng, J., Zhang, G., Wen, S., and He, R., Gp-nas: Gaussian process based neural architecture search, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 11933-11942.

[41] Lin, M., Chen, Q., and Yan, S., Network in network, arXiv preprint arXiv:1312.4400, 2013.

[42] Liu, C. and Wechsler, H., Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition, *IEEE Transactions on Image processing*, 2002, Vol. 11, No. 4, pp. 467-476.

[43] Liu, W., Wen, Y., Yu, Z., Li, M., Raj, B., and Song, L., Sphereface: Deep hypersphere embedding for face recognition, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 212-220.

[44] Liu, W., Zhou, L., and Chen, J., Face recognition based on lightweight convolutional neural networks, *Information* 2021, Vol. 12, No. 5, pp. 191.

[45] Lyu, Y., Jiang, J., and Zhang, K., Hua, Y., Cheng, M., Factorizing and reconstituting large-kernel MBConv for lightweight face recognition, *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, 2019.

[46] Ma, N., Zhang, X., Zheng, H.T., and Sun, J., Shufflenet v2: Practical guidelines for efficient cnn architecture design, *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 116-131.

[47] Ma, Y., Effective methods for lightweight image-based and video-based face recognition, *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, 2019.

[48] Martindez-Diaz, Y., Luevano, L.S., Mendez-Vazquez, H., Nicolas-Diaz, M., Chang, L., and Gonzalez-Mendoza, M., Shufflefacenet: A lightweight face architecture for efficient and highly-accurate face recognition, *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, 2019.

[49] Martinez-Diaz, Y., Nicolas-Diaz, M., Mendez-Vazquez, H., Luevano, L.S., Chang, L., Gonzalez-Mendoza, M., and Sucar, L.E., Benchmarking lightweight face architectures on specific face recognition scenarios, *Artificial Intelligence Review*, 2021, pp. 1-44.

[50] Masi, I., Wu, Y., Hassner, T., and Natarajan, P., Deep face recognition: A survey, *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images*, 2018, pp. 471-478.

[51] Maze, B., Adams, J., Duncan, J.A., Kalka, N., Miller, T., Otto, C., Jain, A.K., Niggel, W.T., Anderson, J., Cheney, J., and Grother, P., Iarpa janus benchmark-c: Face dataset and protocol, *2018 International Conference on Biometrics*, 2018, pp 158-165.

[52] Micikevicius P., Narang, S., Alben, J., Diamos, G., Elsen, E., Garcia, D., Ginsburg, B., Houston, M., Kuchaiev, O., Venkatesh, G., and Wu, H., Mixed precision training, arXiv preprint arXiv:1710.03740, 2017.

[53] Moghaddam, B., Jebara, T., and Pentland, A., Bayesian face recognition, *Pattern recognition*, 2000, Vol. 33, No. 11, pp. 1771-1782.

[54] Moschoglou, S., Papaioannou, A., Sagonas, C., Deng, J., Kotsia, I., and Zafeiriou, S., Agedb: the first manually collected, in-thewild age database, *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops*, Honolulu, HI, USA, 2017, pp. 51-59.

[55] Parkhi, O. M., Vedaldi, A., and Zisserman, A., Deep face recognition, 2015.

[56] Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., and Lerer, A., Automatic differentiation in pytorch, NeurIPS workshop 2017.

[57] Romero, A., Ballas, N., Kahou, S.E., Chassang, A., Gatta, C., and Bengio, Y., Fitnets: Hints for thin deep nets, arXiv preprint arXiv:1412.6550, 2014.

[58] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., KArpathy, A., Khosla, A., Bernstein, M., and Fei-Fei, L., Imagenet large scale visual recognition challenge, *International Journal of Computer Vision*, 2015, Vol. 115, pp. 211-252.

[59] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L.C., Mobilenetv2:Inverted residuals and linear bottlenecks, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510-4520.

[60] Schroff, F., Kalenichenko, D., and Philbin, J., Facenet: A unified embedding for face recognition and clustering, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015.

[61] Sengupta, S., Chen, J.C., Castillo, C., Patel, V.M., Chellappa, R., and Jacobs, D.W., Frontal to profile face verification in the wild, *2016 IEEE Winter Conference on Applications of Computer Vision(WACV)*, Lake Placid, NY, USA, 2016, pp.1-9.

[62] Simonyan, K. and Zisserman, A., Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556, 2014.

[63] Sun, , Y., Wang, X., and Tang, X., Deeply learned face representations are sparse, selective, and robust, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 2892-2900.

[64] Sun, Y., Chen, Y., Wang, X., and Tang, X., Deep learning face representation by joint identification-verification, *Advances in Neural Information Processing Systems*, Vol. 27, 2014.

[65] Sun, Y., Liang, D., Wang, X., and Tang, X., Deepid3: Face recognition with very deep neural networks, arXiv preprint arXiv:1502.00873, 2015.

[66] Sun, Y., Wang, X., and Tang, X., Deep learning face representation from predicting10,000 classes, *Proceed-*

ings of the IEEE Conference on Computer Vision and
Pattern Recognition, 2014, pp 1891-1898.

[67] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S.,
Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich,
A., Going deeper with convolutions, Proceedings of the
IEEE Conference on Computer Vision and Pattern
Recognition, 2015, pp.1-9.

[68] Taigman, Y., Yang, M., Ranzato, M. A., and Wolf, L.,
Deepface: Closing the gap to human-level performance
in face verification, Proceedings of the IEEE Conference
on Computer Vision and Pattern Recognition, 2014, pp.
1701-1708.

[69] Tan, M. and Le, V., Mixconv: Mixed depthwise convolu-
tional kernels, arXiv preprint arXiv:1907.09595, 2019.

[70] Tan, M. and Quoc, L., Efficientnet: Rethinking model
scaling for convolutional neural networks, International
Conference on Machine Learning, 2019, pp. 6105-6114.

[71] Turk, M. and Pentland, A., Eigenfaces for recognition,
Journal of cognitive neuroscience, 1991, Vol. 3, No.
1, pp. 71-86.

[72] Wang, M. and Deng, W., Deep face recognition: A survey,
Neurocomputing, 2021, Vol. 429, pp. 215-244.

[73] Wang, H., Wang, Y., Zhou, Z., Ji, X., Gong, D., Zhou,
J., Li, Z., and Liu, W., Cosface: Large margin cosine
loss for deep face recognition, Proceedings of the IEEE
conference on Computer Vision and Pattern Recognition,
2018, pp. 5265-5274.

[74] Wang, X., Fu, T., Liao, S., Wang, S., Lei, Z., and Mei,
T., Exclusivity-consistency regularized knowledge dis-
tillation for face recognition, European Conference on
Computer Vision, 2020, pp. 325-342.

[75] Wang, X., Teacher guided neural architecture search
for face recognition, Proceedings of the AAAI Conference
on Artificial Intelligence, 2021, Vol. 35. No. 4, pp.
2817-2825.

[76] Whitelam, C., Taborsky, E., Blanton, A., Maze, B.,
Adams, J., Miller, T., Kalka, N., Jain, A.K., Duncan,
J.A., Allen, K., and Cheney, J., Iarpa janus benchmark-b
face dataset, Proceedings of the IEEE Conference on
Computer Vision and Pattern Recognition Workshops,
2017, pp 90-98

[77] Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., and
Ma, Y., Robust face recognition via sparse representation,
IEEE Transactions on Pattern Analysis and Machine
Intelligence, 2008, Vol. 31, No. 2, pp. 210-227.

[78] Wu, X., He, R., Sun, Z., and Tan, T., A light CNN

for deep face representation with noisy labels, IEEE
Transactions on Information Forensics and Security,
2018, Vol. 13, No. 11, pp. 2884-2896.

[79] Xiao, J., Jiang, G., and Liu, H., A Lightweight Face
Recognition Model based on MobileFaceNet for Limited
Computation Environment, EAI Endorsed Transactions
on Internet of Things, 2021, Vol. 7, No. 27, pp. 1-9.

[80] Yan, M., Zhao, M., Xu, Z., Zhang, Q., Wang, G., and
Su, Z., Vargfacenet: An efficient variable group convolu-
tional neural network for lightweight face recognition,
Proceedings of the IEEE/CVF International Conference
on Computer Vision Workshops, 2019.

[81] Yang, S., Wen, Y., He, L., Zhou, M., and Abusorrah,
A., Sparse Individual Low-Rank Component Repre-
sentation for Face Recognition in the IoT-Based System,
IEEE Internet of Things Journal, 2021, Vol. 8, No. 24,
17320-17332.

[82] Zagoruyko, S. and Nikos, K., Paying more attention
to attention: Improving the performance of convolutional
neural networks via attention transfer, arXiv preprint
arXiv:1612.03928, 2016.

[83] Zhang, L., Yang, M., and Feng, X., Sparse representation
or collaborative representation: Which helps face recog-
nition?, 2011 International Conference on Computer
Vision, 2011, pp. 471-478.

[84] Zhang, P., Zhao, F., Liu, P., and Li, M., Efficient
LightweightAttention Network for Face Recognition,
IEEEAccess, 2022, Vol. 10, pp. 31740-31750.

[85] Zhang, Q., Li, J., Yao, M., Song, L., Zhou, H., Li, Z.,
Meng, W., Zhang, X., and Wang, G., Vargnet: Variable
group convolutional neural network for efficient em-
bedded computing, arXiv preprint arXiv:1907.05653,
2019.

[86] Zhao, J., Lv, Y., Zhou, Z., and Cao, F., A novel deep
learning algorithm for incomplete face recognition:
Low-rank-recovery network, Neural Networks, 2017,
Vol. 94, pp. 115-124.

[87] Zhao, W., Chellappa, R., Phillips, P.J., and Rosenfeld,
A., Face recognition: A literature survey, ACM
Computing Surveys (CSUR), 2003, Vol. 35, No. 4, pp.
399-458.

[88] Zheng, T., Deng, W., and Hu, J., Cross-age LFW: A
database for studying cross-age face recognition in uncon-
strained environments. arXiv preprint arXiv:1708.08197,
2017.

[89] Zhu, N., Yu, Z., and Kou, C., A new deep neural archi-

tecture search pipeline for face recognition, *IEEE Access*, 2020, Vol. 8, pp. 91303-91310.

[90] Zoph, B. and Le, Q.V., Neural architecture search with reinforcement learning, arXiv preprint arXiv:1611.01578, 2016.

ORCID

Yongli Zhang | http://orcid.org/0000-0002-9786-2912
Jaekyung Yang | http://orcid.org/0000-0002-4904-1351