IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 24 July 2022, accepted 8 August 2022, date of publication 16 August 2022, date of current version 22 August 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3199014

RESEARCH ARTICLE

Beyond Masks: On the Generalization of Masked Face Recognition Models to Occluded Face Recognition

PEDRO C. NETO[1,2], JOÃO RIBEIRO PINTO[1,2], (Graduate Student Member, IEEE), FADI BOUTROS[3,4], NASER DAMER[3,4], (Member, IEEE), ANA F. SEQUEIRA[1], AND JAIME S. CARDOSO[1,2], (Senior Member, IEEE)

[1]Centre for Telecommunications and Multimedia, INESC TEC, 4200-465 Porto, Portugal
[2]Faculdade de Engenharia, Universidade do Porto, 4200-465 Porto, Portugal
[3]Fraunhofer Institute for Computer Graphics Research IGD, 64283 Darmstadt, Germany
[4]Department of Computer Science, TU Darmstadt, 64289 Darmstadt, Germany

Corresponding author: Pedro C. Neto (pedro.d.carneiro@inesctec.pt)

This work was financed by National Funds through the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia within project LA/P/0063/2020 and the PhD grants "2021.06872.BD" and "SFRH/BD/137720/2018". This research work has been also funded by the German Federal Ministry of Education and Research and the Hessen State Ministry for Higher Education, Research and the Arts within their joint support of the National Research Center for Applied Cybersecurity ATHENE.

ABSTRACT Over the years, the evolution of face recognition (FR) algorithms has been steep and accelerated by a myriad of factors. Motivated by the unexpected elements found in real-world scenarios, researchers have investigated and developed a number of methods for occluded face recognition (OFR). However, due to the SarS-Cov2 pandemic, masked face recognition (MFR) research branched from OFR and became a hot and urgent research challenge. Due to time and data constraints, these models followed different and novel approaches to handle lower face occlusions, i.e., face masks. Hence, this study aims to evaluate the different approaches followed for both MFR and OFR, find linked details about the two conceptually similar research directions and understand future directions for both topics. For this analysis, several occluded and face recognition algorithms from the literature are studied. First, they are evaluated in the task that they were trained on, but also on the other. These methods were picked accordingly to the novelty of their approach, proven state-of-the-art results, and publicly available source code. We present quantitative results on 4 occluded and 5 masked FR datasets, and a qualitative analysis of several MFR and OFR models on the Occ-LFW dataset. The analysis presented, sustain the interoperable deployability of MFR methods on OFR datasets, when the occlusions are of a reasonable size. Thus, solutions proposed for MFR can be effectively deployed for general OFR.

INDEX TERMS Deep learning, biometrics, occluded face recognition, masked face recognition, face biometrics, computer vision.

I. INTRODUCTION

Face recognition methods have evolved over time and are now performing at a human level [1]. Despite the fact that the evaluation of these methods relies on datasets captured in the wild, this evaluation does not address challenges such as occlusions because these datasets are mostly composed of clear images [2]. In real world applications, in fact, there are no guarantees on capturing pictures that show the face free of obstructions. Hence, the training and evaluation of these models are somewhat lacking to include some of the challenges. The recent SarS-Cov2 pandemic led to the mandatory usage of facial masks, which became a frequent obstruction to face images. For this particular occlusion, Damer *et al.* [3] have shown that the performance degradation affects both humans and face recognition systems.

The associate editor coordinating the review of this manuscript and approving it for publication was Mohammad Shorif Uddin.

Parallel to the search for better face recognition models, researchers have also been researching occluded face recognition architectures. Frequently the latter are less discriminative on clear images. Thus, these models also present a trade-off between performance under occlusions and without occlusions. It can already be mitigated by detecting the presence of an occlusion, which, on the other hand, increases the computational requirements. The variety of potential occlusions is also growing as part of technological advances. Examples of this evolution are the popularization of smartphones and the usage of headphones. However, the vast majority of the occlusions are either for aesthetics or random elements, such as high brightness and hands in front of the face [4].

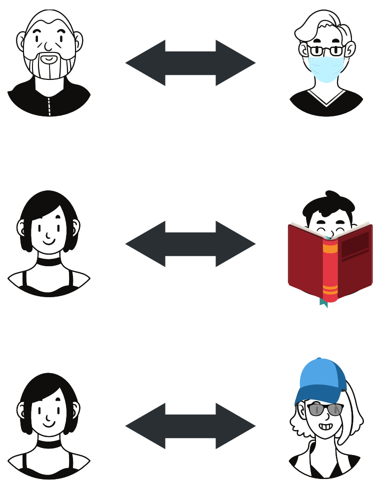

FIGURE 1. Illustration of some occlusions that can highly increase the difficulty of the face verification task.

Due to the pandemic, researchers have directed their focus towards models capable of verifying the identity of someone that is wearing a mask (i.e., masked face recognition). This type of occlusion obscures the majority of facial features below the periocular area. These efforts led to the publication of two competitions [5], [6] and several distinct models based on diverse approaches [7], [8], [9], [10], [11]. Compared to standard face recognition models, the majority of the published models for masked face recognition (MFR) advanced the state of the art. However, as seen in Figure 1, masks are not representative of all the different occlusions. Other types of occlusion might cover other parts of the face. Hence, it remains to be seen if MFR models can also increase the performance on general occluded face recognition. The importance of models that are robust to both problems is proportional to the growth of face recognition adoption across different technologies and scenarios. For instance, it border control it is necessary to prevent the increase in false positives due to specific occlusions, which could lead to security issues. On the other hand, false negatives should remain low in smartphone's face recognition authentication so that the users' experience does not get negatively impacted.

In recent literature, researchers aimed to improve the current robustness of general face recognition for masked face recognition. Thus, knowing that masked faces are a uniform type of face occlusion, the main question posed in this paper is "can the solutions developed for masked face recognition enhance the performance of face recognition for more general occlusion?". Towards that goal, this paper presents several contributions to both masked and occluded face recognition fields. It mainly focuses on understanding the capabilities of models trained for a specific occlusion (face masks) to generalize to others. Validating the performance of a masked face recognition approach in occluded datasets poses two main contributions. First, since these algorithms are fairly lighter than occluded face recognition methods, there is a clear advantage regarding computational costs. Moreover, since masked face recognition approaches leverage knowledge existent in traditional face recognition methods, the growth of traditional face recognition datasets might fuel a performance increase in both masked and occluded face recognition. The main outcomes of the paper show that there is an interoperable deployability of MFR methods on OFR datasets. In fact, MFR methods can be effectively deployed for general OFR and achieve reasonable performance.

II. RELATED WORK

A. FACE RECOGNITION

Advances in face recognition models have been driven by a combination of three elements. One of the elements is the steep increase in computational power. Graphical processing units (GPUs) potentiated the training of larger and more complex models in a reasonable time [12]. This led to deeper models capable of approximating the real function with extraordinary precision. The other is the introduction of novel architectures, such as ResNets [13], that pushed the limits of the State-of-the-Art. Finally, an unprecedented proliferation of the internet led to remarkable growth in the data available and how it can be collected [14], with recent works aiming at replacing such data with privacy-friendly synthetic data [15], [16], [17].

Recently, the design of training losses, aiming at enhancing discriminative ability of the extracted face embeddings, has assumed a preponderant role. Face recognition models highly improved their accuracy and generalization capabilities through the usage of contrastive losses. For instance, triplet loss does not train the model to classify an identity; instead, it optimizes the model to have positive pairs with smaller distances than negative pairs. However, these types of losses have been outperformed by the recent angular margin-based softmax losses [18]. For instance, Center loss [19] uses the learned centres of the features of each identity to decrease intra-class variance. SphereFace [20] (Angular Softmax) adds angular constraints to each identity and normalize its weights, thus creating a hypersphere space. CosFace [21] and ArcFace [18] propose variations of these previous losses. The former introduces an additive margin for the cosine space, whereas the second presents an additive margin for the angular space. MagFace [22] rethinks the margin of loss to be given as a function of the quality of the samples. Finally, ElasticFace [23] proposes a relaxation of the margin constraint by randomly picking its value from a Gaussian distribution.

B. OCCLUDED FACE RECOGNITION

As previously stated, the performance of state-of-the-art face recognition models degrades significantly whenever a face is occluded [4]. Research on how to deal with occlusions led to two main directions: a) predicting the occluded features [24], [25], [26], [27], [28]; and b) removing the occluded features. The former paradigm approached deep learning for the first time in a work proposed by Zhao *et al.* [24]. Their use of a LSTM-Autoencoder to reconstruct the corrupted features led to the loss of identity information, which degraded the results. Ignoring the occlusions is a tough task for deep learning systems due to the difficulty of locating them. To mitigate this difficulty, the MaskNet, introduced by Wan *et al.* [29] presented a middle branch to assign smaller weights to hidden units affected by the corrupted features. However, the training of the middle branch is too relaxed, leading to the inclusion of irrelevant information. The recent state-of-the-art includes two related approaches. The first was proposed by Song *et al.* [30] and it uses a binary mask to clean the corrupted features. It learns a mapping of each block of the face image (divided into $K \times K$ blocks) to the feature mask. The detection of occluded blocks and the processing of the image is conducted in two separate steps. FROM [31], proposed by Qiu *et al.*, removes the need for two different inference stages for training and testing. Their Mask Decoder predicts, in a dynamic fashion, the correct feature masks. Hence, this method is much more efficient at training and testing. Wang and Guo [32] proposed an approach to deal with occlusions and pose variance. Their model, DSA-Face, is based on sparse attention mechanisms. Their results indicate a performance that is slightly above the one delivered by traditional face recognition models. However, it is worth noting that it presents a higher complexity than most known models for face recognition. With a more direct application, Biswas *et al.* [33] proposed a method to tackle Child Sexual Exploitation Material. They state that abusers have their eyes covered in the majority of the content and that traditional face recognition systems fail at recognizing their identity. Hence, they propose a model based on Perceptual hashing to help solve this task and mitigate the humanitarian problem of Child Sexual Exploitation.

C. MASKED FACE RECOGNITION

The effects of the COVID-19 pandemic were felt worldwide. Concurrently, the obligation to use a face mask was present in the majority of the developed countries [34]. These masks occlude a vast part of the face. Hence, due to the lack of those features and the change in the representation nature, face recognition systems are unable to maintain the performance [3], [35], [36]. This drove the research attention towards developing face recognition solutions that aim at maintaining face recognition performance despite the presence of masks. Such solutions were even part of performance competitions organized to consolidate the reporting of advances in this field [5], [6].

Anwar and Raychowdhury [37] showed the benefits of adding synthetic mask augmentations to the training data to increase the algorithm performance on real masks. Boutros *et al.* [6], [7] proposed an efficient on-top solution that can be integrated into any face recognition model. In their solution, when the person wears a mask, a neural network attempts to unmask the embedding produced by the face recognition system. This neural network is trained with a self-restrained triplet loss that put more focus on the more affected genuine pairs, in comparison to impostor training pairs. Neto *et al.* [38] designed a model trained with a modified triplet loss. However, the model performed poorly and was later modified to a multitasking approach based on the ArcFace loss [9]. This model feeds two versions of the same image to the face recognition model. The first contains a mask, whereas the second does not, and the network is constrained to minimize the distance between the two embeddings produced. The multitask optimization for recognition and mask detection was also present in other works [11], while others include multitask learning for face analysis [39].

Geng *et al.* [40] present a generative model to augment the current datasets with masked images. The proposed generative adversarial network (GAN) is specially crafted to retain information regarding the identity of the subject in the image. Mandal *et al.* [41] attempted to approach the problem with a domain adaptation solution. Huber *et al.* [8] intended to approximate embeddings produced by masked and unmasked images. Towards that, he used template level knowledge distillation to a student model, from a teacher model trained on ElasticFace loss. Moreover, to further improve the results, it includes an auxiliary term that relies on the ElasticFace loss to retain identity information within the embedding. Similarly, Li *et al.* [42] proposes a knowledge distillation approach accompanied with image-level face completion. The operation of other subcomponents of face recognition systems has also been affected by mask wearing. Particular attention was paid to presentation attack detection [43], [44] and face image quality assessment [45].

With the raise of masked face recognition as an urgent challenge in face recognition, masked face recognition solutions have evolved as a separate research direction from occluded face recognition. This work tries to link these two conceptually similar research directions by probing the scope of their interoperable deployability.

III. METHODOLOGY

The methodology followed in this document can be divided into three main steps. As previously stated, one of the contributions of this paper is the assessment of the performance of MFR algorithms in OFR problems. Hence, in the first step,

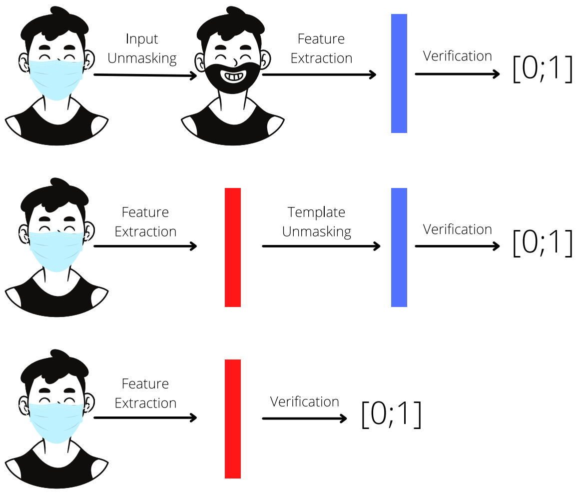

FIGURE 2. **Several approaches can be followed in order to handle masks. These are the three most common approaches: Unmasking the input image, unmasking the template generated by a face recognition model, or using a model that is robust to masks. Blue templates do not contain mask information, whereas red templates contain mask information. Models robust to mask usually have some mask information leaked into their embedding, but in a low proportion when compared to the non-mask information. (Best viewed in colour).**

we evaluated the performance of these algorithms in both masked and occluded face recognition datasets. The algorithms used were carefully selected from recent literature. Afterwards, we repeated this process with a state-of-the-art occluded face recognition algorithm. To further understand the need for both OFR and MFR algorithms, we repeated the evaluation process with algorithms designed for general face recognition.

A. FACE VERIFICATION BY MASKED FACE RECOGNITION MODELS

In the last two years, the number of different models for masked face recognition has grown exponentially. Therefore, it is necessary to select a limited number of them. The selected models must include sufficiently distinct approaches so that the evaluation is diverse. Thus, in the experiments conducted for this research, three masked face recognition models are used. In the context of masked face recognition, models can be evaluated for their performance on masked reference to masked probe (M-M) or unmasked reference to masked probe (U-M) configurations. The former is when both the probe and the reference have the masks, whereas the other has only the probe masked. The different approaches that can be followed for masked face recognition are illustrated in Figure 2.

The first model (FocusFace) [9] presents a multitask neural network that is contrastively optimized to create similar embeddings for masked and unmasked images. FocusFace achieved impressive results in the masked face recognition competition [5] database on M-M face verification and beat the baseline on U-M. The second model (KD) [8] approached this problem through knowledge distillation at the template level, from a teacher network (ElasticFace-Arc) [23], which receives unmasked input, to a student network that receives a mixture of masked and unmsaked faces of the same identities. This enforces the student network to learn how to build embedding from masked images similar to the embedding constructed by the teacher network from unmasked images, while maintaining the discrimination learned by the teacher model on unmasked training data. It is complemented by auxiliary supervision for identity verification through the regularization with the ElasticFace loss. Finally, the last model (EUM) [7] presents a model that operates on top of pre-trained face recognition systems. This model is trained with the self-restrained triplet loss proposed also in [7]. It aims to receive the embedding produced by a face recognition model from a masked input, and map it into an embedding that behaves similarly to the one produced by the same face recognition model from an unmasked input. In other words, it attempts to unmask face embeddings. The self-restrained triplet loss makes sure to automatically adapt the loss so that it puts more attention on the intra-class compactness when required, as this is more affected by the masks as demonstrated in [46]. We trained all the presented models on the MS1MV2 dataset [18], [47] with RGB images aligned (through similarity transfer) and cropped to 112×112, accordingly to the instructions on the original papers. While FocusFace and EUM were also previously tested with a ResNet-50 [13] backbone, in the latter a MobileFaceNet [48] backbone, we decided to train and evaluate these models with a ResNet-100 [13] backbone architecture. Hence, all three models use the same backbone.

B. FACE VERIFICATION BY OCCLUDED FACE RECOGNITION MODELS

Differently from masked face recognition, occluded face recognition requires that its models generalize across occlusions situated at diverse locations on an image. For instance, a scarf can present a challenge similar to a mask. On the other hand, sunglasses create a very different problem. These nuances lead to models and architectures significantly different from those described in the previous section. Research in this field has been short if compared to the efforts dedicated to masked face recognition over the past two years. Hence, the evaluation of occluded face recognition models is focused on a single model. It significantly outperforms others designed for the same task. The selected model (FROM) [31] extends the approach followed by the previous state-of-the-art model PDSN [30]. They present a dynamic approach to correct the feature masks, which they call Mask Decoder. This latter mechanism removes the need for two different train steps and greatly reduces inference time. Differently from previous models, FROM receives aligned images cropped to 112×96. The model's backbone network is initially trained with the CASIA-WebFace dataset for general face recognition. Afterwards, the entire model is fine-tuned end-to-end with the occluded version of the CASIA-WebFace dataset.

C. STATE-OF-THE-ART FACE RECOGNITION

For the evaluation of traditional face recognition algorithms, three different models were chosen from the latest top performing models. They differ from each other by the loss function used for the network optimization. The majority of the other parameters, such as the training set and the backbone network, remain the same. The backbone is the ResNet-100 [13], trained with the MS1MV2 dataset [18], [47]. ArcFace [18] presented a novel angular margin-based softmax loss to increase the discriminative power of the trained model; while it beats the previous State-of-the-Art method, Boutros *et al.* [23] questioned the use of a fixed margin. Hence, they proposed ElasticFace [23], a loss that draws its margin at each iteration from a Gaussian distribution. The ElasticFace has a version where the draw from the Gaussian distribution is also regulated by the samples proximity to their class centers during training, namely the ElasticFace+. Similarly, Meng *et al.* [22] focused their attention on the margin of the loss. Their method, MagFace, presents a novel approach to make the quality of the image affect the margin parameter. All of these models were trained with RGB 112 × 112 cropped and aligned images.

None of these methods was specifically designed to handle occlusions. Hence, their performance can be expected to decrease with the presence of the different occlusions. Their use serves as a baseline for the performance of both the occluded and masked face recognition methods. Masked face recognition algorithms should have, at least, similar performance to these models when confronted with occlusions.

IV. DATABASES

A. MASKED FACE RECOGNITION BENCHMARKS

For masked face recognition, we utilize the masked face benchmarks (proposed in [8]) based on five of the most widely used face recognition benchmarks. Hence, performance comparisons between distinct models are direct and straightforward. The used datasets are: Labeled Faces in the Wild (LFW) [2]; Cross-Age LFW (CALFW) [49]; Cross-pose LFW (CPLFW) [50]; AGEDB-30 [51]; Celebrities in Frontal-Profile (CFP) [52].

1) LABELED FACES IN THE WILD (LFW)

introduced in 2007, this dataset has been widely used to compare the performance of face recognition models on face

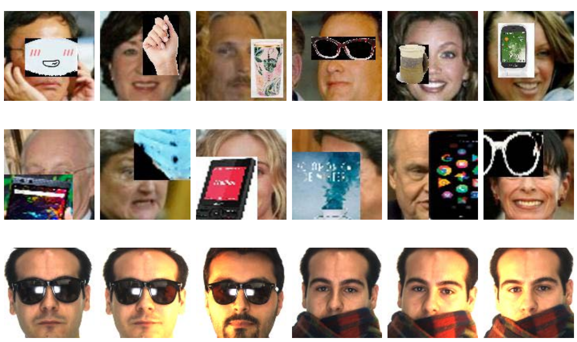

FIGURE 3. Examples of the occluded face verification datasets. The first two rows correspond to Occ-LFW with synthetic occlusions, whereas the third corresponds to the AR database with real occlusions. The first row contains images from the Occ-LFW-1.0, and the second contains images from the Occ-LFW-2.0.

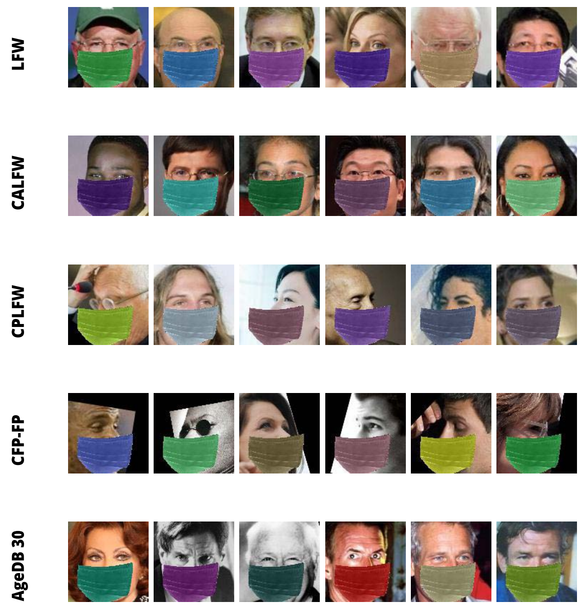

FIGURE 4. Examples of the validation datasets (LFW, CALFW, CPLFW, CFP-FP and AgeDB-30) augmented with face masks for masked face verification. The approach followed for mask augmentation is described by Huber *et al.* [8].

images in unconstrained environments. Moreover, it already includes a protocol, which we strictly follow, for the evaluation of face verification performance on 6,000 image pairs [2].

2) CROSS-AGE LFW (CALFW) AND CROSS-POSE LFW (CPLFW)

are a more challenging version of the original LFW dataset. They present themselves as a renovation of LFW, which is already suggested by their "Cross-Age" and "Cross-Pose" prefix. CALFW introduces images with an age gap, whereas CPLFW presents variations to the pose of the same individual; thus, these two versions increase the intra-class variation to better simulate a real-world face verification scenario.

3) AgeDB-30

is a subset of the larger AgeDB [51]. While the latter has no restrictions on the age difference between images of the same individual, AgeDB-30 imposes a 30-year interval. It is widely used for evaluating the age-invariant face verification performance and contains 16,488 different images of 568 identities. The evaluation protocol followed is also well defined in the literature.

4) CELEBRITIES IN FRONTAL-PROFILE (CFP)

was created to improve the evaluation capabilities on frontal to profile face verification "in the wild". In a sense, the proposed evaluation can be considered similar to occluded face recognition, since, in some extreme poses, many features are completely occluded. It contains 500 subjects, each with 10 frontal and 4 profile images. There are two main evaluation protocols, namely frontal to profile (FP) and frontal to frontal (FF). In this article, only the first protocol is used.

As seen in Figure 4, these datasets were not used as originally proposed. They were modified with face masks as proposed by Huber *et al.* [8] for performance validation on masked face verification. The added masks follow the exact protocol described in the original proposal, and its code is publicly available.[1]

B. OCCLUDED FACE RECOGNITION BENCHMARKS

1) OCC-LFW DATABASE

Song *et al.* [30] presented an approach to synthesizing a dataset of occluded faces with occluders that include sunglasses, scarf, face mask, hand, eye mask, eyeglasses, book, phone, and cup. While it can be argued that these are everyday items that can appear in front of a face, it can also be said that the generated occlusions are not really close to being natural. Nonetheless, it was also the approach followed by Qiu *et al.* [31], hence, also followed in this paper. The images are synthesized by adding a random occluder image on top of a random location of a face. In some scenarios, the full occluder might be absent from the image, since its centre (location of the image selected to add the occlusion) might have a smaller distance to the margin than the size of the occluder.

For this paper, the dataset selected to be occluded is LFW, since it was also used by Qiu *et al.* [31]. It originated three different versions of the LFW. These variations differ from each other due to the different scales of the occluder and the percentage of the face occluded. According to the convention, we name the datasets as follows: Occ-LFW-1.0, Occ-LFW-2.0, and Occ-LFW-3.0. The first has, on average, >18% of the images occluded. The first and the second have >40% and >50% of occluded area, respectively. The same random seed was used, and the code for generating the occlusions is publicly available.[2] Some of these images can be seen in the first two rows of Figure 3.

2) AR DATABASE

The last dataset used is the AR Face Dataset [59]. It contains more than 4,000 face images, which vary in facial expression, illumination conditions and occlusions. These images are from 126 individuals. It presents two different types of real occlusions: sunglasses or a scarf. A person wearing a scarf is more similar to wearing a face mask than a person wearing sunglasses. Occluded images are part of the probe set and the occluded face verification on this dataset is conducted through two different protocols. Some examples can be visualized in the last row of Figure 3.

V. EXPERIMENTAL SETUP

In this section, the performed experiments are presented in detail, as well as the evaluation metrics used to assess the performance of the different models.

[1]https://github.com/fdbtrs/Masked-Face-Recognition-KD
[2]https://github.com/haibo-qiu/FROM

A. EXPERIMENTS

Each experimental setup targets a specific dataset and task. For instance, experiments on the AR dataset are targeted for occluded face recognition. Each of these datasets has well-defined protocols for evaluation, hence, the proposed experiments already benefit from them. All the experimental setups evaluate the verification performance of a plethora of models. For occluded face recognition, the evaluated models were: four distinct versions of ElasticFace [23], MagFace [22], two versions of FocusFace [9], Knowledge Distillation model(KD) [8], and Embedding Unmasking Model (EUM) [7]. For masked face recognition, besides the previously mentioned models, FROM [31] was also evaluated. FocusFace,[3] ElasticFace,[4] KD,[5] EUM[6] and FROM[7] have their code publicly available.

The occluded face verification performance experiments on the AR dataset were conducted on two different protocols [29], [58]. Protocol 1 uses six images per person to form the reference set, whereas protocol 2 uses only one. This dataset includes several previous results published in the literature and is considered to be saturated, especially after the work by Qiu *et al.* [31]. For each protocol, the performance of each method is evaluated individually for both scarves and sunglasses. Thus, it is possible to infer the effect of each on the method.

Experiments conducted on the Occ-LFW dataset are reported individually for each scale value. The defined protocol uses clear images as reference and occluded ones as probe. Further experiments were also conducted on the original LFW dataset as a baseline for the model's performance. The order of the pairs remains the same for all the dataset versions. The experiments conducted for this database are similar to those conducted for the masked validation datasets. The only difference remains in the fact that only unmasked to masked verification was conducted. Due to the applicability of masked face recognition algorithms, the most interesting protocol and the one to be used in practice is when the reference is unmasked and the probe is masked. Hence, for masked face recognition, the masked version of the LFW, CALFW, CPLFW, CFP-FP and AgeDB-30 were used as defined in [8].

We further extend the experiments by selecting a couple of pairs that were either misclassified by all the algorithms or only by the occluded face recognition algorithm. To this, we also add some pairs that were only correctly classified by the occluded face recognition algorithm with the hope that these can help to understand the difficulties and strengths of each approach.

B. EVALUATION METRICS

The evaluation of the models' verification performance is reported accordingly to two different widely used metrics.

[3]https://github.com/NetoPedro/FocusFace
[4]https://github.com/fdbtrs/ElasticFace
[5]https://github.com/fdbtrs/Masked-Face-Recognition-KD
[6]https://github.com/fdbtrs/Self-restrained-Triplet-Loss
[7]https://github.com/haibo-qiu/from

These metrics report a quantitative performance value. The first metric is the accuracy:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

where TP, TN, FP and FN are true positives, true negatives, false positives and false negatives, respectively. However, since the accuracy is somewhat lacking in scenarios where false positives are worse than false negatives, the second metric adopted is TAR (True Accepted Rate) under FAR (False Accepted Rate), which is defined as:

$$TAR = \frac{TP}{TP + FN} \quad (2)$$

$$FAR = \frac{FP}{FP + TN} \quad (3)$$

We acknowledge the evaluation metrics in the ISO/IEC 19795-1 [60] standard, however, to enable the comparability with previous works, we follow the evaluation metrics defined in the utilized benchmarks.

VI. RESULTS

This section presents the quantitative results of all the conducted experiences, and a discussion/interpretation of the meaning and the reasons behind certain results.

A. AR FACE DATABASE

As one of the most used databases for research, it is possible to report the results of a myriad of methods over the years. As seen in Table 1, the performance has been steadily increasing for occluded face recognition models. However, recent general face recognition models have shown impressive results, especially when considering that they were not trained to handle occlusions. For instance, ElasticFace is only surpassed by FROM, which achieves a perfect score. PDSN is capable of achieving a better result, only on protocol 1, and when the occlusion is a scarf.

While masked face recognition models achieve lower results than FROM, they are highly competitive with PDSN and surpass previous models trained for occluded face recognition. Moreover, these models are lighter than FROM and PDSN, in the sense, that they require less computational resources and have fewer parameters. While these results can be interpreted with optimism, the performance of general face recognition methods is similar to the one displayed by masked face recognition models. Thus, it is still unclear how much of this performance is due to an improvement of the former.

TABLE 1. Face verification accuracy(%) on the AR dataset under protocols 1 and 2. "Sg" and "Sf" stand for "Sunglasses" and "Scarf", respectively. The first group of methods represents general face recognition models; the second group of methods represents occluded face recognition models; the last group of methods represents masked face recognition models.

Type	Method	Protocol 1		Protocol 2	
		Sg	Sf	Sg	Sf
FR	CosFace [21], [31]	99.44	**100.00**	98.06	**100.00**
	SphereFace [20], [31]	88.06	96.25	87.50	95.28
	ArcFace [18], [31]	98.33	**100.00**	95.00	99.72
	ElasticFace - Arc [23]	**100.00**	99.61	99.22	98.82
	ElasticFace - Arc+ [23]	99.87	99.34	99.22	98.69
	ElasticFace - Cos [23]	99.87	99.61	99.22	98.95
	ElasticFace - Cos+ [23]	**100.00**	99.48	99.22	98.69
	MagFace [22]	96.73	96.99	92.41	93.33
OFR	Stringface [53]	-	-	82.00	92.00
	LMA [54]	-	-	96.30	93.70
	SRC [25]	87.00	59.50	-	-
	NMR [55]	96.90	73.50	-	-
	MLERPM [56]	98.00	97.00	-	-
	SCF-PKR [57]	95.65	98.00	-	-
	MaskNet [29]	90.90	96.70	-	-
	RPSM [58]	96.00	97.66	84.84	90.16
	PDSN [30]	99.72	**100.00**	98.19	98.33
	FROM [31]	**100.00**	**100.00**	**100.00**	**100.00**
MFR	FocusFace [9]	96.86	99.22	94.90	98.56
	FocusFace - Pretrained [9]	99.22	99.48	98.16	98.82
	ElasticFace-Arc-Aug [8], [23]	99.87	99.61	99.08	98.95
	KD - HG [8]	99.61	99.61	98.95	98.95
	KD - LG [8]	99.08	99.61	98.95	98.82
	EUM - ResNet100 [7]	**100.00**	99.22	98.69	98.30

B. OCC-LFW DATABASE

The experiments of the various models on the original LFW datasets show that the majority of masked and general face recognition models surpass the performance of FROM, which was trained for occluded face recognition. As the best performing Occluded Face Recognition method, and considering that it was also evaluated in masked face recognition datasets, FROM is the ideal benchmark for the comparison with the MFR methods. These results are displayed in Table 2. Despite the good performance of these methods, it is interesting to note that some masked face recognition, such as FocusFace, suffers from a higher degree of degradation than FROM on general face recognition. FROM performance on the Occ-LFW-1.0 dataset was not reported, hence, it was not included. On the other methods, there is a small degradation in the performance. Nonetheless, two methods are capable of achieving better performance on this dataset than the one attained by FROM on the original LFW.

TABLE 2. Face verification accuracy(%) on the LFW and its occluded versions. All the results, except for LFW, are reported for a verification with clear references and occluded probes. Besides FROM, all the other methods were evaluated for this paper.

Type	Method	LFW	Occ-LFW-1.0	Occ-LFW-2.0
OFR	FROM [31]	99.38	-	**94.70**
FR	ElasticFace - Arc [23]	99.80	**99.49**	90.88
	ElasticFace - Arc+ [23]	**99.82**	99.23	90.98
	ElasticFace - Cos [23]	**99.82**	99.40	91.66
	ElasticFace - Cos+ [23]	99.80	99.22	92.06
MFR	FocusFace [9]	99.08	94.95	76.96
	FocusFace - Pretrained [9]	99.53	98.41	85.86
	ElasticFace-Arc-Aug [8], [23]	99.73	99.32	92.33
	KD - HG [8]	99.68	99.30	91.47
	KD - LG [8]	99.68	99.22	91.77
	EUM - ResNet100 [7]	**99.82**	99.23	91.13

The performance discrepancy is completely removed, and FROM becomes the best performing method when evaluated on the Occ-LFW-2.0 dataset. This is due to a faster degradation in the performance of general and masked face recognition methods when the size of the occlusions increases.

TABLE 3. Face verification TAR @ FAR = $1e-3$(%) on the LFW and its occluded versions. All the results, except for LFW, are reported for a verification with clear references and occluded probes. Besides FROM, all the other methods were evaluated for this paper.

Type	Method	LFW	Occ-LFW-1.0	Occ-LFW-2.0	Occ-LFW-3.0
OFR	FROM [31]	98.63	96.17	76.53	**70.23**
FR	ElasticFace - Arc [23]	99.67	98.43	74.67	60.43
	ElasticFace - Arc+ [23]	**99.70**	**99.10**	76.30	59.50
	ElasticFace - Cos [23]	**99.70**	99.03	**80.30**	65.50
	ElasticFace - Cos+ [23]	99.67	98.67	78.16	62.26
MFR	FocusFace [9]	98,30	87.97	48.80	31.96
	FocusFace - Pretrained [9]	99.26	95.86	62.47	45.26
	ElasticFace-Arc-Aug [8], [23]	99.60	98.46	78.40	59.00
	KD - HG [8]	99.63	98.73	76.37	62.37
	KD - LG [8]	99.60	98.76	78.13	60.10
	EUM - ResNet100 [7]	99.23	98.53	75.20	60.87

Nonetheless, ElasticFace and KD achieve unexpectedly good results on a dataset with more than 40% of the image area occluded.

To further understand how this performance degradation affects the results in a real-world evaluation, we also collected the TAR when the FAR was set to $1e-3$. Considering that on a real application the false acceptance rate must be set to a considerably low value, due to its impact on the practicality of the system, the TAR measures how well it performs when the threshold is selected to optimize the FAR value. Table 3 displays the results of each of the evaluated methods. It is interesting to note that on this specific metric the performance of a general face recognition method, ElasticFace, surpasses both masked and occluded face recognition models for Occ-LFW-1.0 and 2.0. The same remarkable performance is attained by KD and EUM.

These impressive and unexpected results are enlightening. In a sense, they argue against the need for models specifically designed for occluded face recognition, since these have a worse performance when the maximum number of false positives is limited. Moreover, the extra parameters, complexity and requirements for these models diminish their interest. On the other side, the success of masked face recognition models indicates a potential research direction to be followed by occluded face recognition methods.

C. MASKED FACE VERIFICATION

On the other hand, the diversity of different occlusions used to train the occluded face recognition model, do not lead to performance comparable to the performance of masked face recognition models. Not only that, but the performance of the general face recognition models is also superior on all the five datasets tested as displayed in Table 4. One important element to consider is that as seen in the previous sections, some approaches that start from a pretrained general face recognition model and are afterwards tuned for masked face recognition perform well on bother masked and occluded datasets. It is also independent of the tuning approach followed.

Furthermore, the performance difference between occluded masked face recognition models and the others is more noticeable on CALFW and AgeDB-30. Hence, it shows that either the model trained for face recognition under occlusions is unable to handle large age gaps, or that the masks make that task significantly harder for that model.

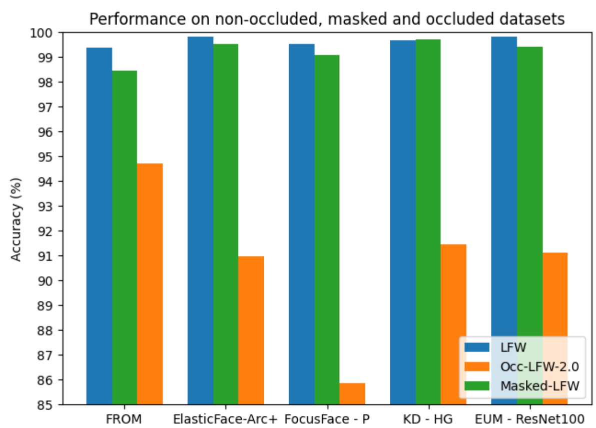

FIGURE 5. Study of the performance of the algorithms when exposed to no occlusions, masks and general occlusions.

Despite the presence of occlusions similar to face masks on the occlusion dataset, the performance of FROM drops significantly on these masked benchmarks. Although it does not vary by more than 1 percent point on the masked LFW when compared to the original LFW, it is a much wider gap than the other models. This performance drop is better seen in Figure 5, which compares the performance of five algorithms in three distinct databases, one for each task. It also shows the impact of Occ-LFW-2.0 in the performance of these algorithms. While the performance drop is significant, we argue in the following sections that this version of the LFW dataset includes unreasonable occlusions. Moreover, if we analyse Figure 6, which studies the performance of the algorithms on progressively more occluded datasets, it is possible to see that the majority of MFR algorithms is comparable to FROM in Occ-LFW-1.0. But as the occlusions increase in size their performance decreases faster.

TABLE 4. Face verification accuracy(%) the masked versions of LFW, CALFW, CPLFW, CFP-FP, AgeDB-30. All the results are reported for a verification with unmasked references and masked probes. All the presented methods were evaluated for this paper. FROM evaluation was conducted on 96 × 112 images, since it was trained on images with those dimensions.

Type	Method	LFW	CFP-FP	AgeDB-30	CALFW	CPLFW
OFR	FROM [31]	98.47	93.01	80.18	89.35	85.70
FR	ArcFace [18]	99.45	95.50	94.68	94.76	89.83
	MagFace [22]	99.33	94.00	94.12	94.15	88.67
	ElasticFace - Arc [23]	99.40	95.29	95.38	94.42	90.40
	ElasticFace - Arc+ [23]	99.52	95.44	95.68	94.62	89.90
	ElasticFace - Cos [23]	99.52	95.43	95.58	94.62	90.18
	ElasticFace - Cos+ [23]	99.37	96.27	95.47	94.78	90.37
MFR	FocusFace [9]	98.82	86.69	93.68	92.50	83.37
	FocusFace - Pretrained [9]	99.10	91.04	94.25	93.80	86.77
	ElasticFace-Arc-Aug [8], [23]	99.58	95.91	96.73	95.58	91.08
	KD - HG [8]	**99.73**	**96.94**	96.69	95.75	91.67
	KD - LG [8]	99.67	96.83	**97.07**	**95.90**	**91.98**
	EUM - ResNet100 [7]	99.43	95.07	95.47	94.90	89.93

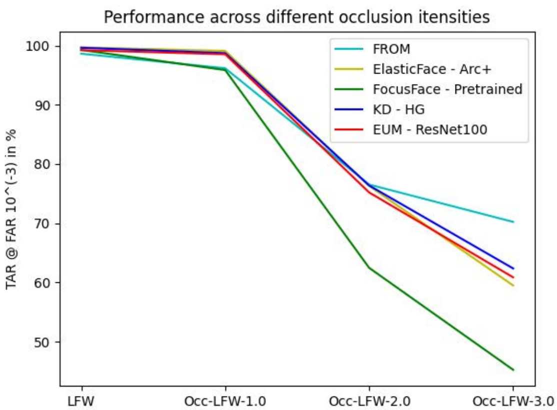

FIGURE 6. Analysis of the performance variation of the algorithms in several versions of the LFW dataset that include different intensity of occlusion.

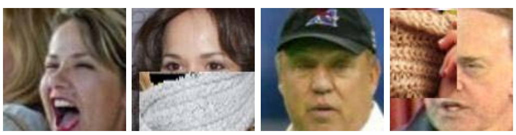

FIGURE 7. Some example pairs from the Occ-LFW-2.0 dataset that were wrongly classified as positive by the following models: FROM, ElasticFace - Arc+, KD - HG and EUM. Each couple of images placed horizontally represents a pair.

D. RESULTS QUALITATIVE ANALYSIS

An analysis was made focusing on some of the misclassified pairs of images, in order to qualitatively study the performance of the algorithms. For this study, the following models were used: FROM, ElasticFace - Arc+, KD - HG and EUM. Only results from Occ-LFW-1.0 and 2.0 were assessed. Figures 7 and 8 show some false positive and false negative samples common to all the models considered. Regarding the false positives, there was no single pair that was common to all models on Occ-LFW-1.0. On 2.0, the two given samples present a considerably difficult task. On the other hand, on the false-negative pairs, on 1.0 the examples displayed indicated a difficulty of all the models on pairs that comprise age gaps.

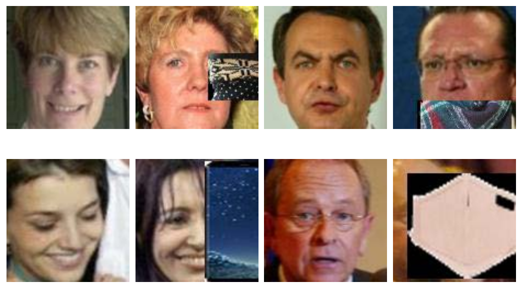

FIGURE 8. Some example pairs from the Occ-LFW dataset that were wrongly classified as negative by the following models: FROM, ElasticFace - Arc+, KD - HG and EUM. First row represents 1.0 dataset and the second the 2.0 dataset. Each couple of images placed horizontally represents a pair.

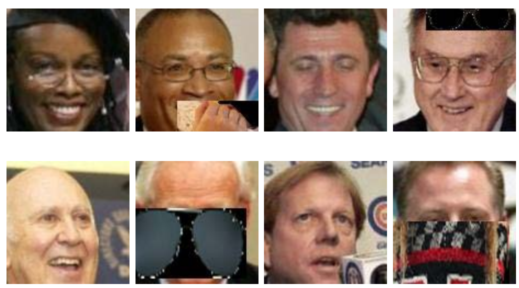

FIGURE 9. Some example pairs from the Occ-LFW dataset that were wrongly classified as positive by the following models: FROM; correctly classified as negative by the following: ElasticFace - Arc+, KD - HG and EUM. First row represents 1.0 dataset and the second the 2.0 dataset. Each couple of images placed horizontally represents a pair.

The second false-negative pair of the 2.0 dataset presents a nonsensical occlusion, which can propel some discussion regarding the validity of the decision, despite being correct or wrong, in these samples.

Figures 9 and 10 show the false-positive and false-negative pairs wrongly classified by FROM and correctly classified by the remaining models. It is interesting to note that the majority of the errors committed exclusively by the occluded face recognition model include occlusions on the lower-face area, similar to face masks. It indicates that the performance

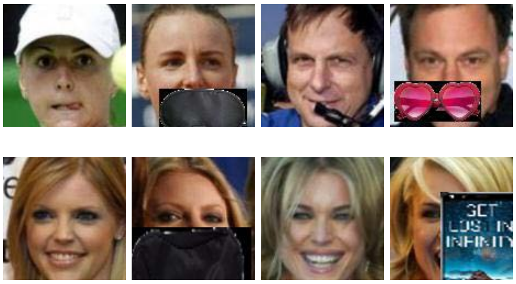

FIGURE 10. **Some example pairs from the Occ-LFW dataset that were wrongly classified as negative by the following models: FROM; correctly classified as positive by the following: ElasticFace - Arc+, KD - HG and EUM. First row represents 1.0 dataset and the second the 2.0 dataset. Each couple of images placed horizontally represents a pair.**

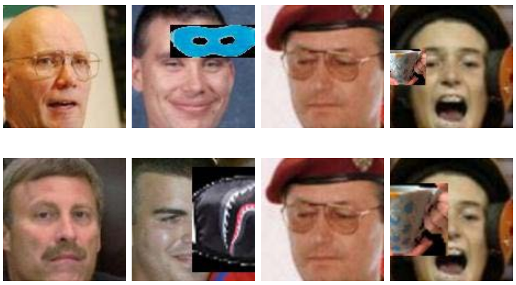

FIGURE 11. **Some example pairs from the Occ-LFW dataset that were wrongly classified as positive by the following models: ElasticFace - Arc+, KD - HG and EUM; correctly classified as negative by the following: FROM. First row represents 1.0 dataset and the second the 2.0 dataset. Each couple of images placed horizontally represents a pair.**

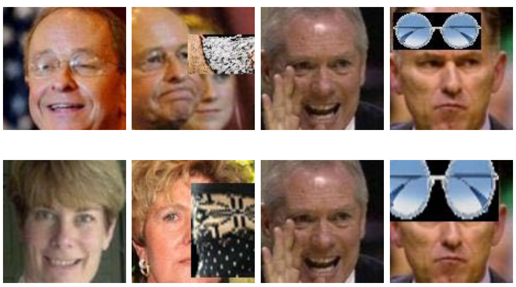

FIGURE 12. **Some example pairs from the Occ-LFW dataset that were wrongly classified as negative by the following models: ElasticFace - Arc+, KD - HG and EUM; correctly classified as positive by the following: FROM. First row represents 1.0 dataset and the second the 2.0 dataset. Each couple of images placed horizontally represents a pair.**

of masked face recognition models generalizes sufficiently well for these synthetic occlusions. Nonetheless, in Figure 9 the occlusions generated are particularly tough and arguably too much unrealistic.

Finally, we analyse Figures 11 and 12, which represent the false positives and false negatives correctly predicted by FROM and wrongly predicted by the other models. Differently from the previous two figures, these indicate that masked face recognition models are particularly ill against occlusions on the ocular area or forehand. Nevertheless, it must be noted that the number of samples wrongly classified by these models and correctly classified by FROM is considerably small. Due to their training nature, this type of error was already expected from the start. Figure 11 shows a false positive between two images from persons of distinct genders.

VII. CONCLUSION

In this paper, we studied the interoperable deployability of two similar but still distinct research directions. We carefully crafted a set of experiments, on masked face recognition and occluded face recognition to analyse the performance of these models in the other task. We aimed to learn the differences between these models and the implications, positive and negatives, of deploying them for a task that they were not trained for. From the set of experiments conducted, it became clear that the performance of the current masked face recognition models is more generalizable, robust and uses fewer parameters. This is an important finding since it shows that building on top of the current face recognition models is a viable option that does not increase the complexity of the solution. These experiments were conducted on datasets frequently used to benchmark each of these two research directions.

The strong capabilities displayed by masked face recognition models foreshadow the future of occluded face recognition. The former, propelled by the urgency of the pandemic, focused on increasing the robustness of current models, which led to simpler, faster to train and more robust solutions when compared to occluded face recognition. However, these findings should not stop researchers from pursuing other research directions within occluded face recognition. We believe that diversity in research and a variety of research directions leads to better breakthroughs. Nonetheless, we still provided sufficient findings to support the investment in an unexplored research direction. Hence, it is possible to conclude that not only masked face recognition models can be used on other occlusions with improved results, but also present a refreshing approach to the task of recognizing under occlusions.

REFERENCES

[1] 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.

[2] B. G. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," Univ. Massachusetts, Amherst, MA, USA, Tech. Rep., 07-49, Oct. 2007.

[3] N. Damer, F. Boutros, M. Süßmilch, M. Fang, F. Kirchbuchner, and A. Kuijper, "Masked face recognition: Human versus machine," *IET Biometrics*, vol. 22, pp. 1–17, May 2022.

[4] D. Zeng, R. Veldhuis, and L. Spreeuwers, "A survey of face recognition techniques under occlusion," *IET Biometrics*, vol. 10, no. 6, pp. 581–606, Nov. 2021.

[5] F. Boutros *et al.*, "MFR 2021: Masked face recognition competition," in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Aug. 2021, pp. 1–10.

[6] J. Deng, J. Guo, X. An, Z. Zhu, and S. Zafeiriou, "Masked face recognition challenge: The InsightFace track report," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2021, pp. 1437–1444.

[7] F. Boutros, N. Damer, F. Kirchbuchner, and A. Kuijper, "Self-restrained triplet loss for accurate masked face recognition," *Pattern Recognit.*, vol. 124, Apr. 2022, Art. no. 108473.

[8] M. Huber, F. Boutros, F. Kirchbuchner, and N. Damer, "Mask-invariant face recognition through template-level knowledge distillation," in *Proc. 16th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, Dec. 2021, pp. 1–8.

[9] P. C. Neto, F. Boutros, J. R. Pinto, N. Damer, A. F. Sequeira, and J. S. Cardoso, "FocusFace: Multi-task contrastive learning for masked face recognition," in *Proc. 16th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, Dec. 2021, pp. 1–8.

[10] Y. Li, K. Guo, Y. Lu, and L. Liu, "Cropping and attention based approach for masked face recognition," *Appl. Intell.*, vol. 51, no. 5, pp. 3012–3025, 2021.

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

[12] D. Steinkraus, I. Buck, and P. Y. Simard, "Using GPUs for machine learning algorithms," in *Proc. 8th Int. Conf. Document Anal. Recognit. (ICDAR)*, Seoul, South Korea, 2005, pp. 1115–1119.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[14] 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.

[15] F. Boutros, M. Huber, P. Siebke, T. Rieber, and N. Damer, "SFace: Privacy-friendly and accurate face recognition using synthetic data," 2022, *arXiv:2206.10520*.

[16] F. Boutros, N. Damer, and A. Kuijper, "QuantFace: Towards lightweight face recognition by synthetic data low-bit quantization," 2022, *arXiv:2206.10526*.

[17] N. Damer, C. A. F. López, M. Fang, N. Spiller, M. V. Pham, and F. Boutros, "Privacy-friendly synthetic data for the development of face morphing attack detectors," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR) Workshops*, Jun. 2022, pp. 1606–1617.

[18] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "ArcFace: Additive angular margin loss for deep face recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 4690–4699.

[19] Y. Xu, W. Yan, G. Yang, J. Luo, T. Li, and J. He, "CenterFace: Joint face detection and alignment using face as point," *Sci. Program.*, vol. 2020, pp. 1–8, Jul. 2020.

[20] 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.

[21] 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/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 5265–5274.

[22] 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.

[23] F. Boutros, N. Damer, F. Kirchbuchner, and A. Kuijper, "ElasticFace: Elastic margin loss for deep face recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2022, pp. 1578–1587.

[24] F. Zhao, J. Feng, J. Zhao, W. Yang, and S. Yan, "Robust LSTM-autoencoders for face de-occlusion in the wild," *IEEE Trans. Image Process.*, vol. 27, no. 2, pp. 778–790, Feb. 2018.

[25] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[26] Z. Zhou, A. Wagner, H. Mobahi, J. Wright, and Y. Ma, "Face recognition with contiguous occlusion using Markov random fields," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep. 2009, pp. 1050–1057.

[27] M. Yang, L. Zhang, J. Yang, and D. Zhang, "Robust sparse coding for face recognition," in *Proc. CVPR*, Jun. 2011, pp. 625–632.

[28] R. He, W.-S. Zheng, B.-G. Hu, and X.-W. Kong, "A regularized correntropy framework for robust pattern recognition," *Neural Comput.*, vol. 23, no. 8, pp. 2074–2100, 2011.

[29] W. Wan and J. Chen, "Occlusion robust face recognition based on mask learning," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 3795–3799.

[30] L. Song, D. Gong, Z. Li, C. Liu, and W. Liu, "Occlusion robust face recognition based on mask learning with pairwise differential Siamese network," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 773–782.

[31] H. Qiu, D. Gong, Z. Li, W. Liu, and D. Tao, "End2End occluded face recognition by masking corrupted features," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Jul. 26, 2021, doi: 10.1109/TPAMI.2021.3098962.

[32] Q. Wang and G. Guo, "DSA-Face: Diverse and sparse attentions for face recognition robust to pose variation and occlusion," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 4534–4543, 2021.

[33] R. Biswas, V. González-Castro, E. Fidalgo, and E. Alegre, "A new perceptual hashing method for verification and identity classification of occluded faces," *Image Vis. Comput.*, vol. 113, Sep. 2021, Art. no. 104245.

[34] M. Gomez-Barrero, P. Drozdowski, C. Rathgeb, J. Patino, M. Todisco, A. Nautsch, N. Damer, J. Priesnitz, W. D. N. Evans, and C. Busch, "Biometrics in the era of COVID-19: Challenges and opportunities," 2021, *arXiv:2102.09258*.

[35] N. Damer, J. H. Grebe, C. Chen, F. Boutros, F. Kirchbuchner, and A. A. Kuijper, "The effect of wearing a mask on face recognition performance: An exploratory study," in *Proc. Int. Conf. Biometrics Special Interest Group (BIOSIG)*, Sep. 2020, pp. 1–6.

[36] G. Jeevan, G. C. Zacharias, M. S. Nair, and J. Rajan, "An empirical study of the impact of masks on face recognition," *Pattern Recognit.*, vol. 122, Feb. 2022, Art. no. 108308.

[37] A. Anwar and A. Raychowdhury, "Masked face recognition for secure authentication," 2020, *arXiv:2008.11104*.

[38] P. C. Neto, F. Boutros, J. R. Pinto, M. Saffari, N. Damer, A. F. Sequeira, and J. S. Cardoso, "My eyes are up here: Promoting focus on uncovered regions in masked face recognition," in *Proc. Int. Conf. Biometrics Special Interest Group (BIOSIG)*, Sep. 2021, pp. 1–5.

[39] V. S. Patel, Z. Nie, T.-N. Le, and T. V. Nguyen, "Masked face analysis via multi-task deep learning," *J. Imag.*, vol. 7, no. 10, p. 204, Oct. 2021.

[40] M. Geng, P. Peng, Y. Huang, and Y. Tian, "Masked face recognition with generative data augmentation and domain constrained ranking," in *Proc. 28th ACM Int. Conf. Multimedia*, Oct. 2020, pp. 2246–2254.

[41] B. Mandal, A. Okeukwu, and Y. Theis, "Masked face recognition using ResNet-50," 2021, *arXiv:2104.08997*.

[42] C. Li, S. Ge, D. Zhang, and J. Li, "Look through masks: Towards masked face recognition with de-occlusion distillation," in *Proc. 28th ACM Int. Conf. Multimedia*, Oct. 2020, pp. 3016–3024.

[43] M. Fang, F. Boutros, A. Kuijper, and N. Damer, "Partial attack supervision and regional weighted inference for masked face presentation attack detection," in *Proc. 16th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, Jodhpur, India, Dec. 2021, pp. 1–8.

[44] M. Fang, N. Damer, F. Kirchbuchner, and A. Kuijper, "Real masks and spoof faces: On the masked face presentation attack detection," *Pattern Recognit.*, vol. 123, Mar. 2022, Art. no. 108398.

[45] B. Fu, F. Kirchbuchner, and N. Damer, "The effect of wearing a face mask on face image quality," in *Proc. 16th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, Jodhpur, India, Dec. 2021, pp. 1–8.

[46] N. Damer, F. Boutros, M. Süßmilch, F. Kirchbuchner, and A. Kuijper, "Extended evaluation of the effect of real and simulated masks on face recognition performance," *IET Biometrics*, vol. 10, no. 5, pp. 548–561, Sep. 2021.

[47] Y. Guo, L. Zhang, Y. Hu, X. He, and J. Gao, "Ms-Celeb-1 M: A dataset and benchmark for large-scale face recognition," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2016, pp. 87–102.

[48] S. Chen, Y. Liu, X. Gao, and Z. Han, "MobileFaceNets: Efficient CNNs for accurate real-time face verification on mobile devices," in *Proc. Chin. Conf. Biometric Recognit.* Cham, Switzerland: Springer, 2018, pp. 428–438.

[49] T. Zheng, W. Deng, and J. Hu, "Cross-age LFW: A database for studying cross-age face recognition in unconstrained environments," 2017, *arXiv:1708.08197*.

[50] T. Zheng and W. Deng, "Cross-pose LFW: A database for studying cross-pose face recognition in unconstrained environments," Beijing Univ. Posts Telecommun., Beijing, China, Tech. Rep., 5:7, 2018.

[51] S. Moschoglou, A. Papaioannou, C. Sagonas, J. Deng, I. Kotsia, and S. Zafeiriou, "AgeDB: The first manually collected, in-the-wild age database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 51–59.

[52] 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.

[53] W. Chen and Y. Gao, "Recognizing partially occluded faces from a single sample per class using string-based matching," in *Proc. Eur. Conf. Comput. Vis.* Berlin, Germany: Springer, 2010, pp. 496–509.

[54] N. McLaughlin, J. Ming, and D. Crookes, "Largest matching areas for illumination and occlusion robust face recognition," *IEEE Trans. Cybern.*, vol. 47, no. 3, pp. 796–808, Mar. 2017.

[55] J. Yang, L. Luo, J. Qian, Y. Tai, F. Zhang, and Y. Xu, "Nuclear norm based matrix regression with applications to face recognition with occlusion and illumination changes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 1, pp. 156–171, Jan. 2017.

[56] R. Weng, J. Lu, J. Hu, G. Yang, and Y.-P. Tan, "Robust feature set matching for partial face recognition," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 601–608.

[57] M. Yang, L. Zhang, S. C.-K. Shiu, and D. Zhang, "Robust kernel representation with statistical local features for face recognition," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 6, pp. 900–912, Jun. 2013.

[58] R. Weng, J. Lu, and Y.-P. Tan, "Robust point set matching for partial face recognition," *IEEE Trans. Image Process.*, vol. 25, no. 3, pp. 1163–1176, Mar. 2016.

[59] A. Martinez and R. Benavente, "The AR face database," CVC Tech. Rep., 24, 1998. [Online]. Available: https://www2.ece.ohio-state.edu/~aleix/ARdatabase.html

[60] A. Mansfield, *Information Technology—Biometric Performance Testing and Reporting—Part 1: Principles and Framework*, ISO/IEC Standard 19795-1:2006, 2006.

PEDRO C. NETO received the B.Sc. degree (Hons.) in software/informatics engineering from the Instituto Superior de Engenharia do Porto (ISEP), and the M.Sc. degree (Hons.) in computer science from Aalto University, Finland. He is currently pursuing the Ph.D. degree in electrical and computer engineering with the Faculdade de Engenharia da Universidade do Porto (FEUP). He is also a Researcher with the Instituto de Engenharia, Sistemas e Computadores, Tecnologia e Ciência (INESC TEC). His Ph.D. work on "Explainable Artificial Intelligence for Fair and Transparent Biometrics," is directed towards the problems exhibited by current biometric systems, such as, race and gender biased in face recognition. He also conducts research for the CADPath.AI project, which aims to build a deep learning method for automated diagnosis of colorectal cancer from whole slide images.

JOÃO RIBEIRO PINTO (Graduate Student Member, IEEE) received the M.Sc. degree in bioengineering (field of biomedical engineering) from the Faculdade de Engenharia da Universidade do Porto (FEUP), in 2017, where he is currently pursuing the Ph.D. degree in electrical and computer engineering. His research has focused on contributing to make the ECG a viable and stronger biometric characteristic in realistic conditions. His Ph.D. studies are focused on using ECG and face, both acquired almost unnoticeably from vehicle drivers, to recognize them and continuously monitor their wellbeing. His M.Sc. thesis are focused on the use of the electrocardiogram (ECG) for biometric recognition of vehicle drivers. His research interests include biometrics, biosignals, pattern recognition, computer vision, and machine learning in general.

FADI BOUTROS received the master's degree in distributed software system from the Technical University of Darmstadt, in 2019, where he is currently pursuing the Ph.D. degree in embedded biometrics. Since March 2019, he has been working as a Research Scientist with the Competence Center of Smart Living & Biometric Technologies, Fraunhofer IGD. He has published several publications in embedded biometrics. His research interests include the fields of machine learning, computer vision, and biometrics. He received the CAST-Förderpreis 2019 Award and the Best Master Thesis Award at the 21st Darmstadte Computer Graphic Night.

NASER DAMER (Member, IEEE) received the Ph.D. degree in computer science from TU Darmstadt, in 2018. He is currently a Senior Researcher with the Fraunhofer IGD, performing research management, applied research, scientific consulting, and system evaluation. He is also a Principal Investigator with the National Research Center for Applied Cybersecurity ATHENE, Germany. He lectures on human and identity-centric machine learning, and on ambient intelligence at TU Darmstadt. His research interests include the fields of biometrics, machine learning, and information fusion. He is a member of the organizing teams of several conferences, workshops, and special sessions, including being the Program Co-Chair of BIOSIG. He is a member of the IEEE Biometrics Council serving on its Technical Activities Committee. He serves as an Associate Editor for *Pattern Recognition* (Elsevier) and the *Visual Computer* (Springer). He represents the German Institute for Standardization (DIN) in the ISO/IEC SC37 International Biometrics Standardization Committee.

ANA F. SEQUEIRA received the Ph.D. degree in electrical and computing engineering, the master's degree in mathematical engineering, and the Ph.D. degree in mathematics from the University of Porto, Portugal. She is currently a Researcher at INESC TEC, applying computer vision and machine learning techniques to biometrics with focus on presentation attack detection. In the past, she worked with the University of Reading, U.K., collaborated in FASTPASS and PROTECT EU projects related to biometrics in border control. She had a short term collaboration with the company Iris Guard U.K., to research on the spoofing vulnerabilities of EyePay®; technology's and develop a proof-of-concept of a countermeasure. As a Ph.D. degree and PDRA, she has authored and coauthored several research publications in major international conferences and journals recognized by peers with citations.

JAIME S. CARDOSO (Senior Member, IEEE) received the Licenciatura degree in electrical and computer engineering, the M.Sc. degree in mathematical engineering, and the Ph.D. degree in computer vision from the University of Porto, in 1999, 2005, and 2006, respectively. He is currently a Full Professor with the Faculty of Engineering, University of Porto (FEUP), where he has been teaching machine learning and computer vision in Ph.D. programs and multiple courses for graduate studies. He is also the Coordinator of the Centre for Telecommunications and Multimedia, INESC TEC, and the Co-Founder and the Co-Leader of the Breast Research Group and the Visual Computing and Machine Intelligence Group. He is also the Co-Founder of ClusterMedia Laboratories. He has coauthored more than 250 papers, and more than 80 of which were in international journals. His research results have been recognized both by his peers with more than 5900 citations to his publications, and the advertisement in the mainstream media several times. His research interests include computer vision, machine learning, and decision support systems. He was a recipient of numerous awards, including the Honorable Mention in the Exame Informática Award 2011, in the software category, for the Project "Semantic PACS" and the First Place in the ICDAR 2013 Music Scores Competition: Staff Removal (task: staff removal with local noise), in August 2013.