Quadruplet Loss For Improving the Robustness to Face Morphing Attacks

Iurii Medvedev
[1] *Institute of Systems and Robotics, University of Coimbra,* Coimbra, Portugal
iurii.medvedev@isr.uc.pt

Nuno Gonçalves [1,2]
[2] *Portuguese Mint and Official Printing Office (INCM), Lisbon, Portugal*
nunogon@deec.uc.pt

***Abstract*—Recent advancements in deep learning have revolutionized technology and security measures, necessitating robust identification methods. Biometric approaches, leveraging personalized characteristics, offer a promising solution. However, Face Recognition Systems are vulnerable to sophisticated attacks, notably face morphing techniques, enabling the creation of fraudulent documents. In this study, we introduce a novel quadruplet loss function for increasing the robustness of face recognition systems against morphing attacks. Our approach involves specific sampling of face image quadruplets, combined with face morphs, for network training. Experimental results demonstrate the efficiency of our strategy in improving the robustness of face recognition networks against morphing attacks.**

***Index Terms*—quadruplet loss, face morphing, face recognition, computer vision, deep learning**

I. Introduction

In recent years, the evolution of deep learning technologies has profoundly influenced the technology and modern security measures. Security concerns continue to expand prompting the need for more robust and reliable identification methods. To address these challenges, biometric approaches have emerged as a promising solution, personalized characteristics for the recognition process.

Among the various modalities of biometric identification, face recognition has garnered significant attention and adoption due to its simplicity of acquisition and recent advancements in computer vision techniques. The widespread use of Face Recognition Systems (FRSs) underscores the importance of facial traits in modern biometric applications, facilitating identification and verification processes across various domains. However, FRSs remain vulnerable to attacks, particularly from sophisticated image manipulation techniques that aim to deceive the system.

One such technique is face morphing, which involves merging or blending two or more digital face images to create a synthetic image that shares biometric properties with the originals, potentially matching different individuals. Face morphing can facilitate acceptance of manipulated face images for the creation of fraudulent documents, which can then be utilized by unauthorized individuals engaging in fraudulent activities. While such fraudulent documents are occasionally detected during border control procedures, the true extent of the existence of fraudulent documents remains unknown [1]. The emergence of deep learning techniques has significantly advanced the field of face recognition, yet face morphing poses a persistent security risk, challenging traditional human or computer-based recognition methods. Ensuring the robustness of face recognition systems against face morphing attacks is crucial for maintaining the integrity of biometric identification.

In this work, we follow and extend contrastive methods in face recognition and introduce a novel quadruplet loss function aimed at enhancing the robustness of face recognition systems to morphing attacks. Our approach involves the specific sampling of quadruplets of face images, combined with face morphs, to train the network. Results of our experiments demonstrate that our strategy effectively increases the robustness of face recognition networks against morphing attacks, thereby contributing to the advancement of biometric security measures.

II. Related Work

In this section, we will review recent advancements in face morphing and face recognition that are relevant to our research objectives.

A. Face Morphing

Generation of a face morph from original images usually includes several stages: *extracting facial features* $\rightarrow$ *averaging features* $\rightarrow$ *generating a morphed image from features* $\rightarrow$ *image refining/retouching*. Early methods [2], [3] relied on facial landmark features for aligning original images and generating face morphs through image blending. However, many of these techniques are prone to producing images with blending artifacts.

Recent advancements in generative deep learning methods have led to the development of face morphing methods that utilize the deep latent feature domain. These approaches may employ a range of deep learning tools, such as variational autoencoders (VAE) [4] or generative adversarial networks (GANs) [5], [6] or diffusion autoencoders [7].

Significant number of recent research studies has been dedicated to the detection of face morphs within biometric systems. [8]–[11].

At the same time the issue of the robustness and stability of facial template (representation) is also very important [12], and even famous private benchmarks begin to support performance estimation for this [13].

For instance Marriott et al. [14] tested several facial recognition algorithms to the robustness against face-morphing attacks and demonstrated the potential threat of face morphing attacks.

B. Face Recognition

Deep learning tools, known for their effectiveness in pattern recognition tasks, have found widespread application in various biometric fields, including face recognition [15]. Among these tools, convolutional neural networks (CNNs) have gained prominence for their capacity to extract discriminative features from unconstrained images, making them particularly efficient in this context.

Recent approaches in face recognition primarily focused on extracting a compact facial biometric template from deep features generated by a backbone network. The main objective is to enhance the discriminative capability of this template under specific conditions. The strategies for training deep networks in face recognition can be categorized into classification and contrastive approaches.

Many recent methods or training face recognition networks approach it via a multi-class closed-set classification problem using existing face image datasets. The discriminative information learned by the network is embedded in its hidden feature layer, which can be utilized for open-set identity discrimination tasks. These methods typically employ softmax loss and its variations for classification [16]. Softmax classification-based methods were broadly modified by techniques to enhance intra-class compactness and inter-class separation. [17]–[22] and with sample-specific strategies for additional control of the feature domain with samples hardness, classification uncertainty or quality [23]–[27].

Contrastive methods (or metric learning methods) utilize the target similarity metric (for instance euclidean or cosine distance) to straightforwardly optimize the distance between deep features by matching face image pairs during the learning process [28]. These methods are usually characterized by the high demands of dataset diversity and sophisticated sampling strategies for reliable convergence of the training process [29]. They are shown to be effective on large datasets and useful in the tasks of transfer learning and fine-tuning.

Initial works straightforwardly optimised pairwise distance contrast utilising small neural networks for feature extraction [30], [31].

The major improvement of contrastive methods in face recognition was introduced by the FaceNet approach [32]. FaceNet proposed the use of triplet loss, which directly optimizes the embedding of face images in a high-dimensional space by the distances within triplets of samples. Triplets are composed of an anchor image, a positive image of the same identity as the anchor, and a negative image of a different identity. This approach required sophisticated sampling of those triplets to learns discriminative feature representation that ensures close proximity between embeddings of images of a same individual while maximizing the separation between embeddings of images of different individuals. Many of the following works on contrastive learning methods in face recognition were inspired by the triplet loss strategy.

For instance DocFaceID [33] introduces the MPS loss function for fine-tuning a pair of sibling networks for the ID document photo matching problem. MPS loss aims to enhance representation learning by maximizing the margin between match pair similarities and non-match pair similarities. This approach simulates a scenario where ID photos serve as templates, while selfies from different subjects act as probes seeking verification, or vice versa.

CoReFace [34] revisited the contrastive learning approach for face recognition by incorporating an adaptive margin for better training regularisation. Additionally, it proposed a new pair-coupling protocol to address the similarity issue stemming from pair symmetry, thereby enhancing the effectiveness of the approach.

C. Quadruplet loss

As an evolution of traditional triplet loss, the Quadruplet loss has emerged as a compelling method in the realm of deep metric learning for face recognition. By leveraging four samples instead of three, the Quadruplet loss can offer enhanced discriminative power and robustness. However such approach can increase the complexity of efficient sampling strategy.

The Quadruplet loss found its application in the tasks of Re-Identification. For instance, it was applied to the Re-Identification of baggage [35] or persons [36], [37].

The Quadruplet loss was also used in applications for Heterogeneous Face Recognition where it addresses challenges posed by cross-domain discrepancies and limited training data [38]. It integrates domain-level and class-level alignment within a unified network to learn domain-invariant discriminative features (DIDF). Domain-level alignment reduces distribution discrepancy, while a specialized quadruplet loss further diminishes intra-class variations and enhances inter-class separability.

In applications of visual tracking the Quadruplet loss was used to develop a discriminative model to one-shot learning (recognize objects of the same class by a single exemplar given) and helped to improve tracking accuracy and real-time speed [39].

III. Methodology

In our work we intend to leverage the potential of quadruplet loss by attaching additional branch which inputs the face morph sample to the common triplet.

A. Quadruplet Loss With Face Morphing

In order to define the Quadruplet loss in our work we revising the formulation of the triplet loss in the FaceNet approach [32]. The triplet loss is computed for three face images (anchor x_i^a, positive x_i^p and negative x_i^n) where anchor and a positive belong to the same identity and negative belong to another identity. Next the distance between anchor and a positive is minimised and the distance between anchor and a negative is maximised:

$$\|x_i^a - x_i^p\|_2^2 + \alpha < \|x_i^a - x_i^n\|_2^2, \forall(x_i^a, x_i^p, x_i^n) \in \tau \qquad (1)$$

where α is a margin that is enforced between positive and negative pairs. τ is the set of all possible triplets in the training set.

The embedding is represented by $f(x) \in \mathbb{R}^d$ and the loss that is being minimized is then:

$$L = \sum(\|f(x_i^a) - f(x_i^p)\|_2^2 - \|f(x_i^a) - f(x_i^n)\|_2^2 + \alpha) \qquad (2)$$

We modify this formulation by inserting a morph sample into the consideration. This face morph sample may be generated from two face images, which belong to the same identities of positive samples. The source images for generating may be indeed different from other components of the initial triplet. With this modification the cross-sample distance condition have the following formulation:

$$\begin{aligned} &\|x_i^a - x_i^p\|_2^2 + \alpha < \omega_{a/n}\,\|x_i^a - x_i^n\|_2^2 + \\ &+\omega_{a/m}\,\|x_i^a - x_i^m\|_2^2 + \omega_{n/m}\,\|x_i^n - x_i^m\|_2^2 + \\ &+\omega_{p/m}\,\|x_i^p - x_i^m\|_2^2, \forall(x_i^a, x_i^p, x_i^n, x_i^m) \in \mathbb{Q} \end{aligned} \qquad (3)$$

where $x_i^a, x_i^p, x_i^n, x_i^m$ correspond to anchor, positive, negative, and morphed (for positive/negative pairs) in the quadruplet $\mathbb{Q}$ and ω_i signify the weight for each part of the sum, which are selected to balance identity matched and non-matched distances ($\omega_i = 0.25$) on Fig. 1.

Such methods in fact is directed onto achieving the maximum discriminative power of the learned features, while also enlarging variances between original images and their morphed combinations. The approximate visualisation of desired feature distribution (in case of 2D feature layer dimensionality) is depicted on the Fig. 2.

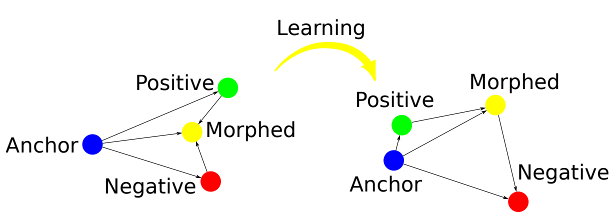

Fig. 1. The proposed Quadruplet Loss minimizes the distance between an anchor and a positive, both of which have the same identity, and maximizes the distance between the anchor and a negative, anchor and morphed, negative and morphed where anchor and negative belongs to different identities and morphed is taken between the positive and negative.

B. Data Harvesting

Our training data is based on the collection, which is developped and used in [10]. It is build on the filtered version of VGGFace2 dataset [40], where quality-based filtering is used to select a subset of images suitable for face morphing

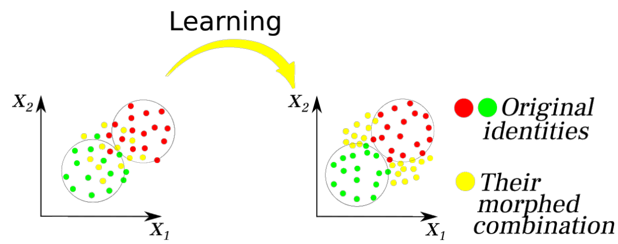

Fig. 2. The distribution of features x (with two dimensional feature layer) with common face recognition approaches (left) and desired feature distribution in proposed approach (right). Each point on 2D surface corresponds to a single image features (data is not based on real experiments).

(frontal image with acceptable quality). Morphs are generated from those filtered originals with a customized landmark-based morphing approach is utilized with a blending coefficient set to 0.5. Additionally, GAN-based morphs are generated using the StyleGAN method. These morphs are synthesized by projecting original images onto a latent domain, followed by interpolation of their deep representations to produce the resulting morph. We also employ ”selfmorphs” generated by applying face morphing to images of the same identity. These selfmorphs serve as bona fide samples, allowing to focus on the behavior of deep face features while mitigating the influence of perceptual artifacts. In our work we assume that deep discriminative face features remain unaffected by selfmorphing.

C. Data Sampling

Training data generator combines the batches of quadruplets with the above data. In our sampling strategy we go thought the list of available face morphs and for each one we select anchor, positive and negative from the source morph sample identities to combine the quadruplet (see Fig. 3). The resulting quadruplets are combined into batches and fed to the Siamese network input.

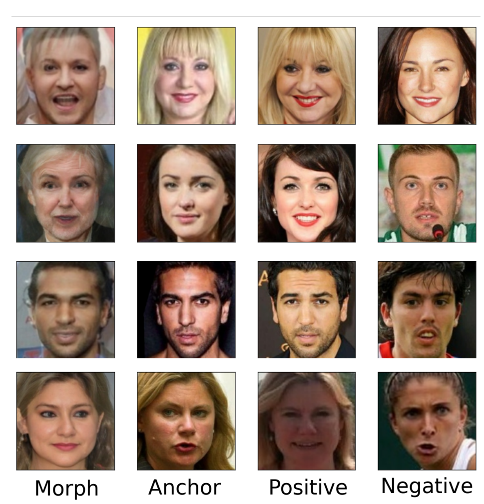

Fig. 3. Quadruplet samples

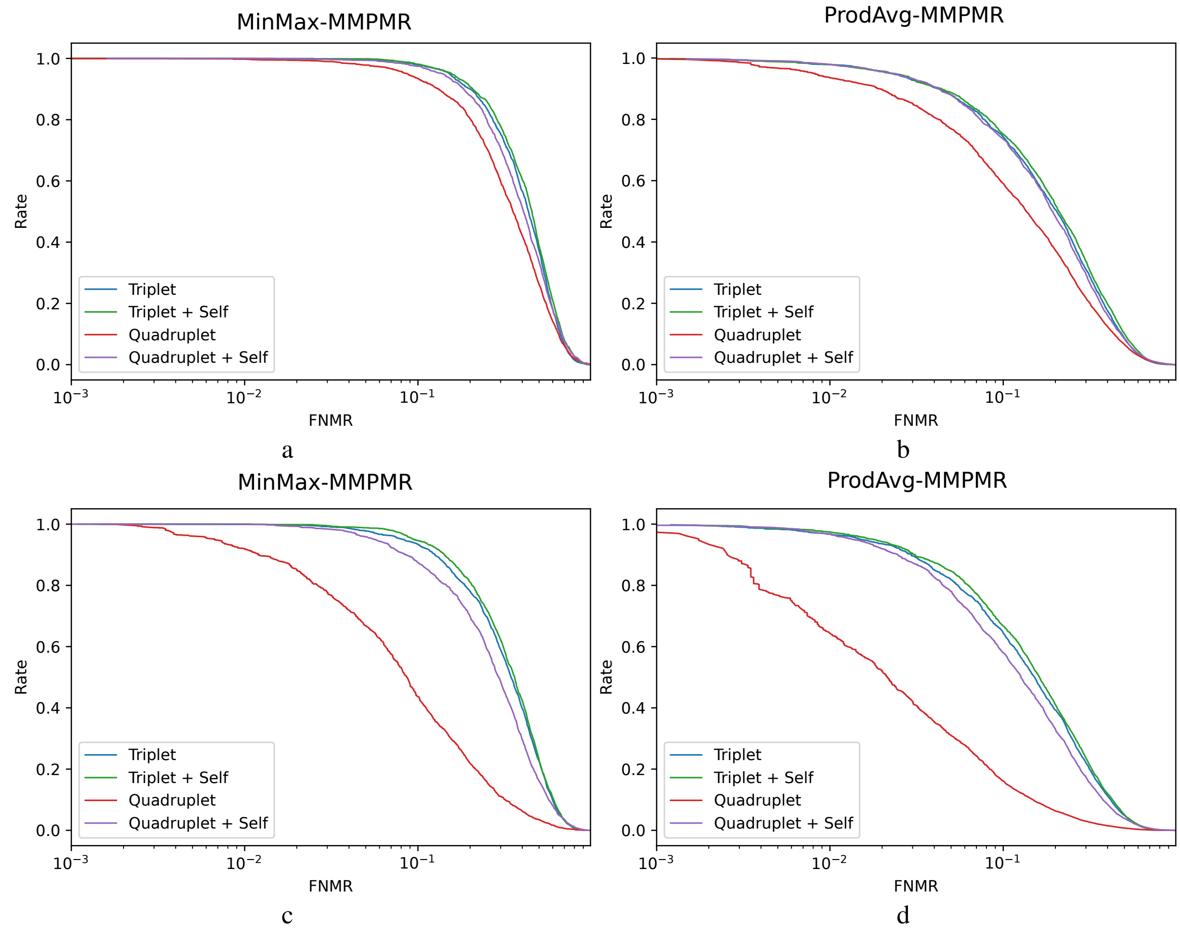

Fig. 4. MinMax-MMPMR and ProdAvg-MMPMR for different models and protocols as a dependency from FNMR. a,b - LDM benchmark; c,d - STG benchmark.

IV. EXPERIMENTS

We perform several experiment with our methodology and selected data. In all our experiments we train the Resnet-50 architecture with image size 300×300 pixels by 10 epochs with Adam optimizer. However first we need to discuss the metrics and benchmarks that will be used for performance estimation.

A. Benchmarking

There are several metrics for estimating the robustness of face recognition system to face morphing. Following the [12] we compute the MinMax-MMPMR (Eq. 4) and ProdAvg-MMPMR (Eq. 5), which are motivated by the international standard ISO/IEC 30107-3 [41].

$$MinMax-MMPMR(\tau) = \frac{1}{M} \cdot \sum_{m=1}^{M} \left\{ \left(\underset{i=1,...,N_m}{min} \left[\underset{i=1,...,I_m^n}{max} S_m^{n,i} \right] \right) > \tau \right\} \quad (4)$$

where I_m^n is the number of samples of subject n within morph m.

$$ProdAvg-MMPMR(\tau) = \frac{1}{M} \cdot \sum_{m=1}^{M} \left\{ \prod_{n=1}^{N_m} \frac{1}{I_m^n} \cdot \sum_{i=1}^{I_m^n} \{ S_m^{n,i} > \tau \} \right\} \quad (5)$$

To estimate those metrics we build a custom benchmark basing on the dataset from [42]. This dataset is collected fur the studies of ICAO-requirements [43] conformity and thus contain both ICAO-compliant and non-compliant images. One of the benefits of the data is that it contain images collected with a professional digital camera and a smartphone giving these two modalities of the data. We use those different modalities to simulate the differences of enrollment and life capture (reference) in our benchmark utilities. Namely we select *enrollment* images from the ones that are collected with professional digital camera and the *reference* from the smartphone images. This also imply strict ICAO-compliance for the enrollment images and relaxed conditions for the reference ones. In total we have 1567 and 8760 images correspondingly.

In order to normalize the result metric rates they are usually computed at the same level of FNMR in some standard face recognition 1-1 performance value. We define such standard 1-1 verification protocol by pairing *enrollment* and *reference* images. This protocols consist of ∼71k of pairs (∼2k match pairs and ∼69k non-match pairs)

We define a standard face pairing protocol for morphing, which standardize the images selection for performing face morphing. This protocol consist of 2142 pairs, thus giving the same amount of face morphs in our benchmarks. face morphs are generated only from the *enrollment* images. Following this protocols we generate face morphs with a landmark based approach (LDM) and a StyleGAN (STG) [44] approach to have in total two different benchmarks for performing comparisons in our work.

B. Results

In our experiments we compare the results of with the Triplet loss baseline (trained on the same date with the same conditions) and also investigating the effect of self-morph presence. We present our results as the MMPMR dependency from the FNMR values (see Fig. 4) and values of MMPMR at the FNMR=0.01 (see Table I).

TABLE I
MINMAX-MMPMR AND PRODAVG-MMPMR FOR DIFFERENT MODELS AND PROTOCOLS AT FNMR=0.01.

Model & Protocol	Rate @ FNMR=0.01	
	MinMax-MMPMR	ProdAvg-MMPMR
Triplet LDM	0.99953	0.97890
Triplet+Self LDM	1.0	0.97849
Triplet STD	0.99906	0.96828
Triplet+Self STG	1.0	0.97456
Quadruplet LDM	0.99766	0.93530
Quadruplet+Self LDM	1.0	0.978870
Quadruplet STG	0.91876	0.642203
Quadruplet+Self STG	0.999533	0.9666627

Withing the same setup Quadruplet Loss with our sampling strategy allowed to achieve better face morphing robustness then the Triplet Loss baseline. At the same time self morphs didn't demonstrate significant impact to the performance.

V. CONCLUSION

In this work we revisit the contrastive methods in face recognition and introduce a novel quadruplet loss function, which is designed to improve the robustness of face recognition systems against morphing attacks. Through the specific sampling of quadruplets of face images, coupled with face morphs, our approach demonstrates promising results in enhancing the resilience of face recognition networks against morphing attacks. By introducing the formulation of the quadruplet loss and conducting extensive experiments across various data settings, we prove the effectiveness of our strategy in resisting face morphing attacks. In our future work we plan to extend this study by enlarging our benchmark utilities with other types of morph samples and also print/scan morphs and perform corresponding experiments. We also plan to perform a extensive study of our benchmark utilities with various face recognition utilities and networks.

REFERENCES

[1] M. Torkar, "Morphing cases in slovenia," NIST IFPS, 2022, ministry of the Interior Police, Slovenia.

[2] M. Ferrara, A. Franco, and D. Maltoni, "The magic passport," *IJCB 2014 - 2014 IEEE/IAPR*, 12 2014.

[3] Biometric System Laboratory, "UBO-Morpher," 2018, http://biolab.csr.unibo.it/Research.asp. (accessed: September 1, 2022).

[4] N. Damer, A. M. Saladié, A. Braun, and A. Kuijper, "MorGAN: Recognition Vulnerability and Attack Detectability of Face Morphing Attacks Created by Generative Adversarial Network," in *2018 IEEE 9th International Conference on BTAS*, 2018, pp. 1–10.

[5] S. K. Venkatesh, H. Zhang, R. Ramachandra, K. B. Raja, N. Damer, and C. Busch, "Can GAN Generated Morphs Threaten Face Recognition Systems Equally as Landmark Based Morphs? - Vulnerability and Detection," *2020 8th IWBF*, pp. 1–6, 2020.

[6] H. Zhang, S. Venkatesh, R. Ramachandra, K. Raja, N. Damer, and C. Busch, "MIPGAN - Generating Robust and High Quality Morph Attacks Using Identity Prior Driven GAN," *ArXiv*, vol. abs/2009.01729, 2020.

[7] N. Damer, M. Fang, P. Siebke, J. N. Kolf, M. Huber, and F. Boutros, "Mordiff: Recognition vulnerability and attack detectability of face morphing attacks created by diffusion autoencoders," 2023. [Online]. Available: https://arxiv.org/abs/2302.01843

[8] M. Ferrara, A. Franco, and D. Maltoni, "Face morphing detection in the presence of printing/scanning and heterogeneous image sources," *ArXiv*, vol. abs/1901.08811, 2019.

[9] S. Lorenz, U. Scherhag, C. Rathgeb, and C. Busch., "Morphing attack detection: A fusion approach," in *IEEE Fusion*, 2021.

[10] I. Medvedev, F. Shadmand, and N. Gonçalves, "Mordeephy: Face morphing detection via fused classification," in *Proceedings of ICPRAM*. SciTePress, 2023, pp. 193–204.

[11] P. C. Neto, T. Gonçalves, M. Huber, N. Damer, A. F. Sequeira, and J. S. Cardoso, "Orthomad: Morphing attack detection through orthogonal identity disentanglement," in *2022 International Conference of the Biometrics Special Interest Group (BIOSIG)*, 2022, pp. 1–5.

[12] U. Scherhag, A. Nautsch, C. Rathgeb, M. Gomez-Barrero, R. Veldhuis, L. Spreeuwers, M. Schils, D. Maltoni, P. Grother, S. Marcel, R. Breithaupt, R. Ramachandra, and C. Busch, "Biometric systems under morphing attacks: Assessment of morphing techniques and vulnerability reporting," in *2017 International Conference of the Biometrics Special Interest Group, BIOSIG 2017*. Germany: Gesellschaft für Informatik, 9 2017.

[13] NIST, "NIST FRVT MORPH," 2022, https://pages.nist.gov/frvt/html/frvt_morph.html.

[14] R. T. Marriott, S. Romdhani, S. Gentric, and L. Chen, "Robustness of facial recognition to gan-based face-morphing attacks," *ArXiv*, vol. abs/2012.10548, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:229339565

[15] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. B., A. C. Berg, and L. Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge," *IJCV*, vol. 115, no. 3, pp. 211–252, 2015.

[16] Y. Sun, X. Wang, and X. Tang, "Deeply learned face representations are sparse, selective, and robust," *2015 IEEE Conference on CVPR*, pp. 2892–2900, 2015.

[17] Y. Wen, K. Zhang, Z. Li, and Y. Qiao, "A Discriminative Feature Learning Approach for Deep Face Recognition," in *Computer Vision – ECCV 2016*, 2016, pp. 499–515.

[18] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj, and L. Song, "SphereFace: Deep Hypersphere Embedding for Face Recognition," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 6738–6746.

[19] 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 *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 5265–5274.

[20] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "ArcFace: Additive Angular Margin Loss for Deep Face Recognition," in *2019 Conference on CVPR*, 2019, pp. 4685–4694.

[21] F. Wang, J. Cheng, W. Liu, and H. Liu, "Additive Margin Softmax for Face Verification," *IEEE Signal Processing Letters*, vol. 25, no. 7, pp. 926–930, 2018.

[22] J. Sun, W. Yang, J. Xue, and Q. Liao, "An Equalized Margin Loss for Face Recognition," *IEEE Transactions on Multimedia*, pp. 1–1, 2020.

[23] D. Zeng, H. Shi, H. Du, J. Wang, Z. Lei, and T. Mei, "NPCFace: A Negative-Positive Cooperation Supervision for Training Large-scale Face Recognition," *CoRR*, vol. abs/2007.10172, 2020. [Online]. Available: https://arxiv.org/abs/2007.10172

[24] Y. Shi and A. Jain, "Probabilistic Face Embeddings," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, pp. 6901–6910.

[25] Q. Meng, S. Zhao, Z. Huang, and F. Zhou, "Magface: A universal representation for face recognition and quality assessment," in *2021 IEEE/CVF Conference on CVPR*, 2021, pp. 14 220–14 229.

[26] J. Tremoço, I. Medvedev, and N. Gonçalves, "QualFace: Adapting Deep Learning Face Recognition for ID and Travel Documents with Quality Assessment," in *2021 International Conference of the BIOSIG*, 2021, pp. 1–6.

[27] I. Medvedev, J. Tremoço, B. Mano, L. E. Santo, and N. Gonçalves, "Towards understanding the character of quality sampling in deep learning face recognition," *IET Biometrics*, vol. 11, no. 5, pp. 498–511, 2022. [Online]. Available: https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/bme2.12095

[28] S. Chopra, R. Hadsell, and Y. LeCun, "Learning a similarity metric discriminatively, with application to face verification," *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1, pp. 539–546 vol. 1, 2005.

[29] A. Mishchuk, D. Mishkin, F. Radenović, and J. Matas, "Working hard to know your neighbor's margins: local descriptor learning loss," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, ser. NIPS'17. Red Hook, NY, USA: Curran Associates Inc., 2017, p. 4829–4840.

[30] H. V. Nguyen and L. Bai, "Cosine similarity metric learning for face verification," in *Asian Conference on Computer Vision*, 2010. [Online]. Available: https://api.semanticscholar.org/CorpusID:18453384

[31] J. Hu, J. Lu, and Y.-P. Tan, "Discriminative deep metric learning for face verification in the wild," *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1875–1882, 2014. [Online]. Available: https://api.semanticscholar.org/CorpusID:3217198

[32] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in *2015 IEEE Conference CVPR*, 2015, pp. 815–823.

[33] Y. Shi and A. K. Jain, "DocFace: Matching ID Document Photos to Selfies*," in *2018 IEEE 9th International Conference on BTAS*, 2018, pp. 1–8.

[34] Y. Song and F. Wang, "Coreface: Sample-guided contrastive regularization for deep face recognition," 2023.

[35] H. Yang, X. Chu, L. Zhang, Y. Sun, D. Li, and S. J. Maybank, "Quadnet: Quadruplet loss for multi-view learning in baggage re-identification," *Pattern Recognition*, vol. 126, p. 108546, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0031320322000279

[36] H. Proença, E. Yaghoubi, and P. Alirezazadeh, "A quadruplet loss for enforcing semantically coherent embeddings in multi-output classification problems," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 800–811, 2021.

[37] W. Chen, X. Chen, J. Zhang, and K. Huang, "Beyond triplet loss: A deep quadruplet network for person re-identification," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1320–1329, 2017. [Online]. Available: https://api.semanticscholar.org/CorpusID:14795862

[38] S. Yang, K. Fu, X. Yang, Y. Lin, J. Zhang, and C. Peng, "Learning domain-invariant discriminative features for heterogeneous face recognition," *IEEE Access*, vol. 8, pp. 209 790–209 801, 2020.

[39] X. Dong, J. Shen, D. Wu, K. Guo, X. Jin, and F. M. Porikli, "Quadruplet network with one-shot learning for fast visual object tracking," *IEEE Transactions on Image Processing*, vol. 28, pp. 3516–3527, 2017. [Online]. Available: https://api.semanticscholar.org/CorpusID:3972862

[40] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman, "Vggface2: A dataset for recognising faces across pose and age," in *International Conference on FG*, 2018.

[41] International Organization for Standardization, "ISO/IEC 30107–3:2017. Information Technology—Biometric Presentation Attack Detection — Part 3: Testing and Reporting," ISO/IEC JTC 1/SC 37 Biometrics, p. 33, 09 2017.

[42] C. Guerra, J. Marcos, and N. Gonçalves, "Automatic validation of ICAO compliance regarding head coverings: An inclusive approach concerning religious circumstances," in *International Conference of the Biometrics Special Interest Group, BIOSIG 2023, Darmstadt, Germany, September 20-22, 2023*, N. Damer, M. Gomez-Barrero, K. B. Raja, C. Rathgeb, A. F. Sequeira, M. Todisco, and A. Uhl, Eds. IEEE, 2023, pp. 1–4. [Online]. Available: https://doi.org/10.1109/BIOSIG58226.2023.10345995

[43] E. V. C. L. Borges, I. L. P. Andrezza, J. R. T. Marques, R. A. T. Mota, and J. J. B. Primo, "Analysis of the Eyes on Face Images for Compliance with ISO/ICAO Requirements," in *2016 29th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, 2016, pp. 173–179.

[44] T. Karras, S. Laine, and T. Aila, "A Style-Based Generator Architecture for Generative Adversarial Networks," in *2019 IEEE/CVF Conference on CVPR*, 2019, pp. 4396–4405.