Improving Identification Accuracy on Unconstrained Low-Resolution Tiny Faces via Absolute Cosine Similarity

Ravindra Kumar Soni*
Department of Computer Science and Engineering
Malaviya National Institute of Technology Jaipur, India
Email: 2019rcp9159@mnit.ac.in

Neeta Nain
Department of Computer Science and Engineering
Malaviya National Institute of Technology Jaipur, India
Email: nnain.cse@mnit.ac.in

***Abstract*—Low-resolution unconstrained face recognition is one of the most active research areas in computer vision. Deep learning models trained on high-resolution face images produce significant accuracy. However, the performance degrades when applied to unconstrained low-resolution faces because facial features have extremely low visual information due to blurred, occluded, and varying lighting conditions in the images. In the face identification scenario, a probe face is matched with a list of gallery faces. At Rank-1, if both probe and gallery face pertains to the same identity, then it is a correct match. But if the most similar face in the gallery that matches the probe pertains to a different identity, can result in misidentification. These similar-looking faces are referred to as lookalike faces. This paper proposes a deep learning framework that minimizes the misidentification problem caused due to the lookalike faces and enhances the identification accuracy at top ranks. In our framework, we extract the deep features from the unconstrained tiny faces using deep convolution neural networks (DCNN), aggregate the features into a fixed-length feature vector then applies the subspace learning methodology and uses absolute cosine similarity metric to produce an outstanding performance for matching the identities in the embedding space. Exhaustive quantitative experimental analysis on the open-source TinyFace dataset shows better Rank-1, Rank-20 and Rank-50 accuracy than the state-of-the-art methods.**

I. Introduction

Unconstrained low-resolution face recognition is still evolving due to intra/inter variations on the pose, occlusion, and illumination. In the case of face recognition in low-resolution images, which are very tiny in size (usually 32×32) [1], extracted features contain low inter-class variations, making it a highly challenging identification task. Due to very less variations in extracted features; thereby we see the challenge of lookalike disambiguation [2].

Face detection is a vital stage in face recognition. Deep neural networks have demonstrated robust performance in many computer vision tasks such as object detection [3], object segmentation [4], and face recognition [5], [6], [7]. Using the power of deep neural networks, we can extract in-depth features from the faces, which have enough variations to recognize the person. The speed of the face detector is still a bottleneck for face recognition in the images. Single-shot multi-box face detector (SSD) [8], and You only look once (YOLO) [9] are faster enough for face detection but not efficient for capturing the faces at multiple scales.

Deep learning feature-based algorithms have outperformed the handcrafted feature-based algorithms. In deep convolution neural networks (DCNN), initial layers extract lower-level features (e.g., edge, corner, line etc.). In contrast, deep layers extract the face's global features (e.g., eyes, nose etc.). Receptive fields also play a vital role in extracting the features in the deep learning algorithms. Convolution features at deep layers tend to have large receptive fields for extracting the global features, while small receptive fields at initial layers are responsible for extracting the lower level features [10]. Feature aggregation is a crucial part of face recognition algorithms. Many deep learning algorithms are proposed to extract the features, but the efficient one is that it can produce the fixed-length facial feature vector with enough variation. Schroff extitet al. [7] proposed the deep convolution neural networks with triplet loss for generating a 128 byte dimensional vector of fixed-length embedding of the face [7].

Data is the most crucial part of deep learning-based algorithms. The accuracy of deep learning-based approaches depends on the amount of data used to train the parameters of a model. Contrary to the previous approaches, manifolds and subspace learning have also achieved a lot of attention for image-based face recognition [11], [12]. In subspace learning-based methods, a set of image samples is used as manifolds or subspace and, an appropriate subspace similarity metric is used for the identification (1:N) or verification (1:1). The main advantage of the subspace based methods is that instead of mean, subspace representations encode the correlation information well among the samples. The subspace produced by the correlation of the face embedding space will learn a representation to capture the variations efficiently. Despite the remarkable progress in face recognition using subspace projection, majority of the prior works' efficacy degrades under unconstrained tiny face recognition. Therefore, this paper proposes a method for look-alike disambiguation via absolute cosine similarity for unconstrained tiny faces. Experimental analysis on a public database shows the efficacy of the proposed methodology.

The remainder of the article is organized as below. Section II presents an overview of the prior works on face recognition and identification. The proposed tiny face identification framework is outlined in Section III. Dataset and experimental protocol are discussed in Section IV. Experiments are discussed in Section V. Finally, conclusions are drawn in Section VI.

II. Related Work

Turk and Pentland [13] proposed the face recognition model using the principal component analysis approach. Parkhi *et al.* [14] proposed a very deep convolution neural networks (CNNs) using the VGGNet for face verification. Peiyun *et al.* [10] proposed the convolution neural network for detecting the TinyFaces in the image. Schroff *et al.* [7] proposed the deep convolution neural networks, which uses the triplet loss and margin parameter for training the network and generates the 128 byte dimension vector for better face recognition in images having pose and illumination. Lamba *et al.* [15] proposed a study on the look-alike faces where the faces of different identities look the same, which is a severe problem in face recognition. Sun *et al.* [16] achieved impressive results on the LFW (Labeled Faces in the Wild) dataset [17] by minimizing the intra variations and maximizing the inter variations among the classes. Ding *et al.* [18] proposed the trunk branch ensemble convolution neural network for recognition of human faces in the video when there is a severe problem of blurriness and pose conditions. Sun *et al.* [19] proposed the high-performance deep convolution network by increasing the dimension of hidden representations and adding supervision to early convolution layers. Zheng *et al.* [20] proposed the automatic system for face recognition in the video using the deep convolution neural network and subspace to subspace learning. Ren *et al.* [3] proposed the faster R-CNN algorithm, which uses the concept of region segmentation for object detection. Praveen and Nain [21] introduced the multi-scale patch GAN for generating the photo-realistic images with preserve identity. Martínez-Díaz, Yoanna and Méndez [22] made a comparative study of very deep to the lightweight deep learning networks. Experiments show an impressive improvement in the results using lightweight networks that are easy to deploy in resource-constraint environments. The Meta Face Recognition (MFR) method was proposed to handle face recognition in unseen domains without updating the model [23]. Chang*et al.* [24] proposed the idea of data uncertainty learning using mean and variance. Zangeneh *et al.* [25] proposed a two-branch deep convolutional neural network architecture for mapping of high-resolution(HR) and low-resolution(LR) images into a common subspace with non-linear transformations.

Zhang *et al.* [26] proposed the Multi-task Cascaded Convolution Networks (MTCNN) for detecting the face at multiple scales in the image. Deng *et al.* [27] proposed the additive angular margin loss in face recognition. When classes are compact enough, the angular margin in the loss, can better classify the classes, which increases the discriminative power of the classification. Chen *et al.* [28] proposed the model based on a deep convolution network which learns the Deep IDentification-verification features and achieved 99.15% face verification accuracy on the challenging LFW dataset [17]. Wen *et al.* [29] invented the idea of training convolution neural networks (CNNs) using a new supervision signal called center loss instead of the Softmax loss function, which is used widely in training. Liu *et al.* [30] proposed the angular softmax (A-Softmax) loss function to learn the angular discriminative features by the convolution neural networks (CNNs).

III. Proposed Unconstrained Tiny Face Recognition Method

There are four major stages in our proposed model, as shown in Fig. 1. The first stage extracts the deep features from the image, passing it through a deep convolution neural network (DCNN). The second stage creates an embedding space from the deep features. The third stage learns the subspace similarity using the principal component analysis (PCA) for a unique representation of identity in the subspace. The final stage uses an appropriate subspace similarity metric for discriminating identities in the subspace. The following sections detail the method.

A. Feature extraction

Deep Convolution Neural Networks (DCNNs): Deep convolution neural network, as shown in Fig. 1 is used to extract the features from the tiny images that produce the face embedding of size $512-$byte dimensional vector. Let I_s be the image and $\phi(I_s) \in R^D$ be the embedding generated after applying the function $\phi \longrightarrow I_s : y$, which embeds the image I_s into a D dimensional embedding space.

$$y = \phi(I_s) \tag{1}$$

Here I_s is the resized (up-sampled) image, passed through the deep convolution network function ϕ that produces an embedding (y) of 512 byte (D) dimensional vector.

B. Fusion of features

The next critical task is to fuse the embeddings of an identity into a fixed size or uniform representation. Finding the mean of the embeddings is the most common method to fuse the embeddings. In our methodology, we extract the embeddings from the different tiny faces of a person and produce an embedding matrix. Subspace learning is applied to the embedding matrix using the principal component analysis (PCA). The eigenvector of the highest variance is chosen to represent the identity in the embedding space. In the experimental analysis, we have compared the result of the feature fusion approach using the mean embeddings and subspace learning using the PCA methodology.

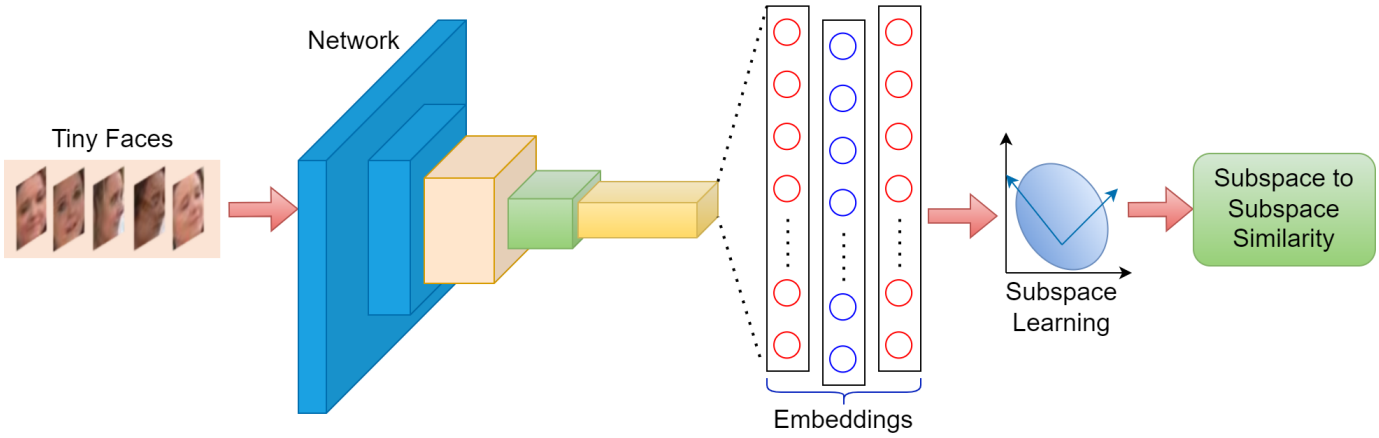

Fig. 1. A deep learning pipeline framework for unconstrained low resolution tiny face recognition.

C. Subspace learning from embedding features:

Let Y $\in R^{D\times n}$ be the embedding matrix, where D denotes the dimensionality of the features and n denotes the number of samples of the images from which deep features are extracted using DCNN. Y_n represents the mean of the embeddings and X: mean centred matrix of the embeddings.

$$Y_n = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{2}$$

$$X = Y - Y_n \tag{3}$$

$$Z_{es} = XX^T = U\lambda U^T \tag{4}$$

The covariance embedding space $XX^T \in R^{D\times D}$ generates the orthonormal bases $P = \{P_1, P_2, P_3, \cdots, P_D\}$ where the dimension of each principal component is D. In TinyFace dataset [31] probe and gallery sets contain the unequal number of images of the identities ($n = 1, 2, 3, ...$). Due to this reason the projected data contains the different number of features for each identity.

Exploiting the above information, we select the principal component of the highest variance as a unique representative of the identity in the embedding space. The embedding space created using the different tiny faces of the same identity has enough feature variations.

D. Similarity metric

Absolute Cosine Similarity: Let $P_1 \in R^{D\times 1}$ and $P_2 \in R^{D\times 1}$ are the two principal components of the two subspaces S_1 and S_2 and θ is the angle between P_1 and P_2 where θ lies in the range of $0 \leq \theta \leq 90^\circ$ then the absolute cosine similarity can be defined as:

$$S_{ABSCos}(P_1, P_2) = \left| \frac{\sum_{i=1}^{D} P_{1i}P_{2i}}{\sqrt{\sum_{i=1}^{D} P_{1i}^2}\sqrt{\sum_{i=1}^{D} P_{2i}^2}} \right| \tag{5}$$

IV. Dataset and Experimental Protocols

A. Synthetic data generation

For generating the synthetic data for low-resolution face recognition, we have used the CASIA-WebFace dataset [32]. In this approach, we first down-sample the high-resolution (HR) original image (112×112) to the size of 7×7 and then up-sample it to the original size, which produces the low-resolution image. To capture the different degradations available in the real world, we down-sample the image to the sizes of 14×14, 28×28 and 56×56, then up-sample them to the original size. This process produces low-resolution images at different scales.

Fig. 2. An example of synthetically generated low resolution (LR) face images using Area interpolation.

From Fig. 2, it is noticeable that Area interpolation degrades the spatial resolution of the image. To simulate the real world degradation processes, different scales are applied to the image.

B. Dataset details

We have trained our model on the CASIA-WebFace synthetically generated images as shown in Fig. 2 and evaluated the identification results on the TinyFace dataset [31]. The Open-source TinyFace dataset implements the open-set face recognition protocol for train and test sets generation, which means there is no identity overlapping in train and test sets. In TinyFace dataset, there are 5139 labelled facial identities defined by 169403 native low resolution (LR) images (average size 20×16) for face recognition. The TinyFace dataset is split into train and test sets, where the train set consists of 2570

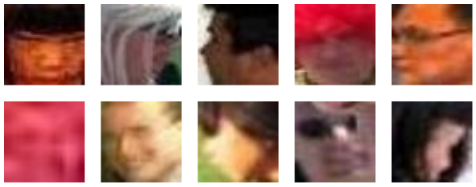

Fig. 3. An example of native low resolution unconstrained images from TinyFace dataset [31].

identities with 7804 images. The test set is further divided into probe, gallery and distractor sets. In the gallery, there are 4443 images having 2569 identities, while in the probe, there are 3728 images having 2569 identities. The rest of the identities images are present in the distractor set. Some of the images being unlabelled in the distractor set, we have not considered them in our experiment. In the TinyFace dataset, images are captured under unconstrained viewing conditions in pose, occlusion, and varying illumination. A sample of the TinyFace dataset is shown in Fig. 3.

C. Implementation details

We have performed our experiments using the ResNet-100 model as a backbone network. The ResNet-100 model contains the property of skip connections that avoids the gradient vanishing problem in deeper networks. During training of the model, An embedding size of 512, Stochastic gradient descent (SGD) optimizer with a weight decay of $5e^{-4}$ and an initial learning rate of 0.1 is used to train the network. We have employed ArcFace loss [27] function to train the model. The learning rate of our model decreases after 3k and 5k iterations, as depicted in Fig. 4. After training our model on CASIA-WebFace synthetically generated images, we have finetuned (FT) it on TinyFace train set.

V. Experiment's Results

A. Comparison with existing face recognition algorithms

TABLE I
GENERIC FACE RECOGNITION (FR) EVALUATION ON TINYFACE [31]

Metric %	Rank-1	Rank-20	Rank-50
DeepID2 [28]	17.4	25.2	28.3
SphereFace [30]	22.3	35.5	40.5
VggFace [14]	30.4	40.4	42.7
CenterFace [29]	32.1	44.5	48.5
ShuffleFacenet [22]	43.1	58.9	64.5
ResNet-100FT(Ours)	**39.1**	**56.9**	**63.0**

As listed in Table I, our model has better Rank-1, Rank-20, and Rank-50 accuracy compared to the state-of-the-art methods like DeepID2 [28], SphereFace [30], VggFace [14] and CenterFace [29] but less than ShuffleFacenet [22].

B. Loss graph

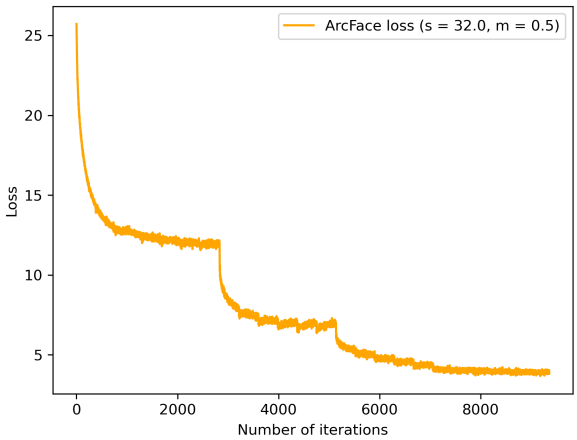

Fig. 4. Training loss of our model decreases as learning rate changes.

In the feature space, even when the classes face the problem of lookalike, absolute cosine similarity separates them well, and principal components can distinctly represent the classes with well-defined distributions in the feature space. At Rank-20 and Rank-50, in the feature space classes are distributed so perfectly that there are minimum number of false matches because the match value lies within the range of the rank value, which increases the Rank-20 and Rank-50 accuracy significantly and the metric function can find the correct match within the specified rank in the embedding space.

C. Effect of cosine similarity

From Fig. 5, it is observed that the principal components of the same identities lies either in the range of $0° - 37°$ (58.72%) or in $151.4° - 175.6°$ (41.28%) using the cosine similarity metric.

D. Effect of absolute cosine similarity

From Fig. 6, it is evident that using the absolute cosine similarity, the angle between all the principal components of the same identities lies in the range of $0° - 37°$ $(58.72\% + 41.28\%)$, emphasizing a strong degree of similarity between the same identities. For $P_{id=3}$, we can verify the angle values in Fig. 5 and Fig. 6 respectively.

As we measure the Rank-1, Rank-20 and Rank-50 accuracy on TinyFaces using absolute cosine similarity and cosine similarity as a metric function, cosine similarity produces false non-matches due to the alignment of the principal components of the same identity in the opposite direction. Due to this, even when the identities are the same, the cosine similarity metric computes the less similarity value for them. Instead, when we apply absolute cosine similarity, principal components of the same identity (in probe and gallery sets) are projected close to each other with a minimum angle

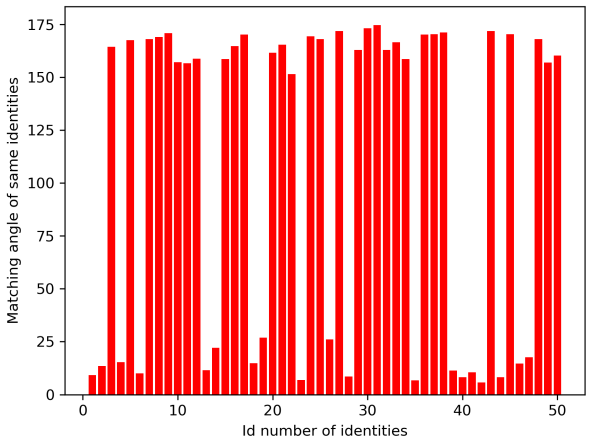

Fig. 5. Matching of same identities in Rank-1 accuracy using cosine similarity (50 identities data is shown).

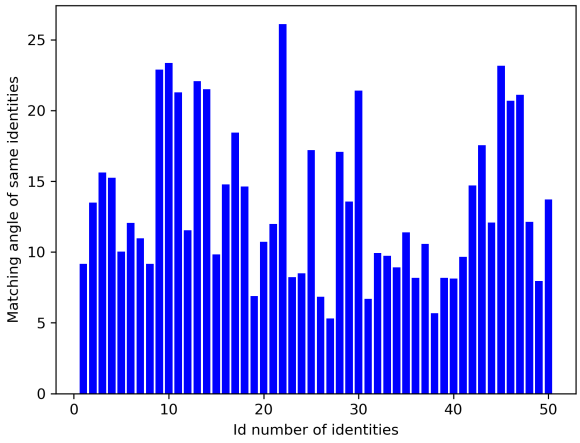

Fig. 6. Matching of same identities in Rank-1 accuracy using absolute cosine similarity.

$(0° - 37°)$ between them means more similarity value, which increases the rank accuracy of our method.

Fig. 7 shows that our model achieves the 39.13% accuracy at Rank-1, 56.9% accuracy at Rank-20 and 63.0% accuracy at Rank-50, and There is a large margin in the accuracy achieved using the absolute cosine similarity (ABSCos) and cosine similarity metric. We conclude that for tiny faces, where the classes suffer the problem of lookalike, absolute cosine similarity separates the classes well where all principal components lie in the range of $0 \leq \theta \leq 90°$.

For example, in Fig. 8, it can be observed that using cosine similarity, the number of false positives for id 33 is ids: 224, 140, 58. In contrast, absolute cosine similarity

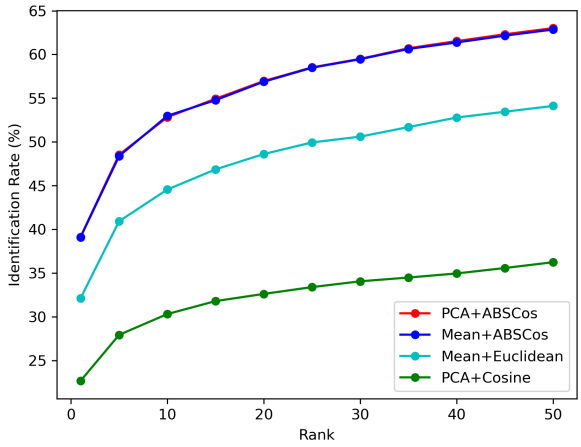

Fig. 7. Cumulative Match Curve(CMC) for the identification accuracy of different metric functions.

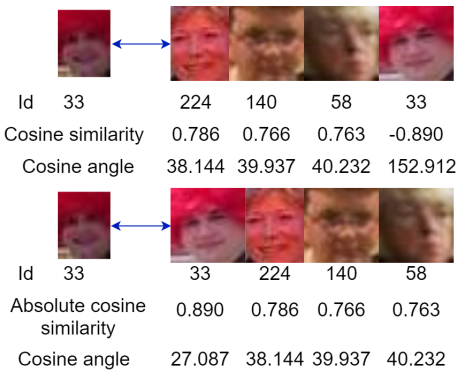

Fig. 8. Absolute cosine similarity for diminishing the problem of look alike.

increases the rank of id 33 from Rank-4 to Rank-1 and pushes down false positives. Ultimately as the number of false positives decreases, the Rank-1 accuracy of our proposed system increases.

E. Ablation study on different metric functions

TABLE II
ABLATION STUDY ON METRIC FUNCTIONS WITH SUBSPACE LEARNING METHODS.

Metric %	**Rank-1**	**Rank-5**	**Rank-10**
PCA+ABSCos	39.13	48.53	52.82
PCA+Cosine	22.66	27.93	30.31
Mean+ABSCos	39.09	48.38	52.98
Mean+Cosine	39.09	48.38	52.98
Mean+Euclidean	32.11	40.92	44.55

Table II depicts that the principal component analysis (PCA) with absolute cosine similarity (ABSCos) produces better

Rank-1 and Rank-5 identification accuracy than mean embedding with absolute cosine similarity. Rank-10 accuracy using the mean embedding with cosine or absolute cosine similarity produces better identification results than PCA with absolute cosine similarity. Rank-1 identification accuracy is prominent; hence PCA with absolute cosine similarity produces the best results in our experimental analysis.

VI. Conclusion

Our proposed methodology shows that the model using the absolute cosine metric produces better results than cosine similarity. Instead of using the image-to-image matching based on class labels, we combine all the embeddings of the identity and generate a unique representation of that identity in the embedding space using the subspace learning method. Our proposed methodology is capable enough to resolve the issue of lookalike and finds the correct match to an extent even when the faces are tiny (20×16), blurred and posed. The subspace to subspace similarity is computed using the absolute cosine similarity metric, where subspaces are learnt using the principal component analysis. In future work, we aim to develop ensemble-based systems, where models trained on synthetically generated datasets will be combined to extract the features from tiny faces having abundant variations to recognize the person in the unconstrained environment. Variation of loss functions and synthetic data generation approaches will also be used to enhance the identification accuracy of the recognition system.

References

[1] A. Torralba, R. Fergus, and W. T. Freeman, "80 million tiny images: A large data set for nonparametric object and scene recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 11, pp. 1958–1970, 2008.

[2] T. Swearingen and A. Ross, "Lookalike disambiguation: Improving face identification performance at top ranks," in *2020 25th International Conference on Pattern Recognition (ICPR)*, 2021, pp. 10 508–10 515.

[3] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.

[4] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Semantic image segmentation with deep convolutional nets and fully connected crfs," *arXiv preprint arXiv:1412.7062*, 2014.

[5] R. Ranjan, S. Sankaranarayanan, C. D. Castillo, and R. Chellappa, "An all-in-one convolutional neural network for face analysis," in *2017 12th IEEE International Conference on Automatic Face Gesture Recognition (FG 2017)*, 2017, pp. 17–24.

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

[7] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 815–823.

[8] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in *European conference on computer vision*. Springer, 2016, pp. 21–37.

[9] J. Redmon and A. Farhadi, "Yolo9000: Better, faster, stronger," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 6517–6525.

[10] P. Hu and D. Ramanan, "Finding tiny faces," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 951–959.

[11] R. Wang, S. Shan, X. Chen, Q. Dai, and W. Gao, "Manifold–manifold distance and its application to face recognition with image sets," *IEEE Transactions on Image Processing*, vol. 21, no. 10, pp. 4466–4479, 2012.

[12] R. Wang and X. Chen, "Manifold discriminant analysis," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 429–436.

[13] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of cognitive neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

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

[15] H. Lamba, A. Sarkar, M. Vatsa, R. Singh, and A. Noore, "Face recognition for look-alikes: A preliminary study," in *2011 International Joint Conference on Biometrics (IJCB)*. IEEE, 2011, pp. 1–6.

[16] Y. Sun, X. Wang, and X. Tang, "Deep learning face representation by joint identification-verification," *arXiv preprint arXiv:1406.4773*, 2014.

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

[18] C. Ding and D. Tao, "Trunk-branch ensemble convolutional neural networks for video-based face recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 4, pp. 1002–1014, 2017.

[19] Y. Sun, X. Wang, and X. Tang, "Deeply learned face representations are sparse, selective, and robust," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 2892–2900.

[20] J. Zheng, R. Ranjan, C.-H. Chen, J.-C. Chen, C. D. Castillo, and R. Chellappa, "An automatic system for unconstrained video-based face recognition," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 2, no. 3, pp. 194–209, 2020.

[21] P. K. Chandaliya and N. Nain, "Child face age progression and regression using self-attention multi-scale patch gan," in *2021 IEEE International Joint Conference on Biometrics (IJCB)*, 2021, pp. 1–8.

[22] Y. Martínez-Díaz, H. Méndez-Vázquez, L. S. Luevano, L. Chang, and M. Gonzalez-Mendoza, "Lightweight low-resolution face recognition for surveillance applications," in *2020 25th International Conference on Pattern Recognition (ICPR)*, 2021, pp. 5421–5428.

[23] J. Guo, X. Zhu, C. Zhao, D. Cao, Z. Lei, and S. Z. Li, "Learning meta face recognition in unseen domains," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 6163–6172.

[24] J. Chang, Z. Lan, C. Cheng, and Y. Wei, "Data uncertainty learning in face recognition," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 5710–5719.

[25] E. Zangeneh, M. Rahmati, and Y. Mohsenzadeh, "Low resolution face recognition using a two-branch deep convolutional neural network architecture," *Expert Systems with Applications*, vol. 139, p. 112854, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417419305561

[26] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499–1503, 2016.

[27] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "Arcface: Additive angular margin loss for deep face recognition," 2019.

[28] Y. Sun, Y. Chen, X. Wang, and X. Tang, "Deep learning face representation by joint identification-verification," in *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*, ser. NIPS'14. Cambridge, MA, USA: MIT Press, 2014, p. 1988–1996.

[29] Y. Wen, K. Zhang, Z. Li, and Y. Qiao, "A discriminative feature learning approach for deep face recognition," in *European conference on computer vision*. Springer, 2016, pp. 499–515.

[30] 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.

[31] Z. Cheng, X. Zhu, and S. Gong, "Low-resolution face recognition," 2018.

[32] D. Yi, Z. Lei, S. Liao, and S. Z. Li, "Learning face representation from scratch," *arXiv preprint arXiv:1411.7923*, 2014.