v. 3, n. 16, 2023

Journal of
Engineering
Research

UNCOVERING DEMOGRAPHIC INFORMATION ON DEEP-FACE FEATURES

Rafael Oliveira Ribeiro

Abstract: Facial recognition is one of the most successful applications of Deep Learning, with the advent of Convolution Neural Networks (CNN) being associated with break-through results in the last few years. The novel aspect of the use of CNNs has not changed the basic facial recognition pipeline, though: facial detection, pre-processing, feature extraction, and comparison/recognition. This paper investigates the possibility of inferring demographic information from facial features generated by CNNs trained for facial recognition. Pre-trained models on three different architectures (ArcFace, DeepFace and FaceNet) are used to extract features from faces of five distinct datasets: Fairface, UTKFaces, Labeled Faces in the Wild, Racial Faces in the Wild, and CelebA. Features and labels from the Fairface dataset are used to train neural networks classifiers for gender (female and male) and ethnicity (african, asian, caucasian and indian). Differences in the performance of the classifiers were observed, depending on which facial recognition model/ architecture is used as feature extractor. Some of the trained classifiers showed an improvement in performance compared to results in the literature and very low variance among demographic groups.

INTRODUCTION AND PREVIOUS WORKS

This paper investigates the possibility of obtaining demographic information from features extracted from facial images. It focuses on gender and ethnicity, but previous works demonstrated the feasibility of inferring more information from these features. One recent study (TERHÖRST et al., 2020) showed that facial templates can be used to accurately predict up to 74 attributes, with many of these being of non-permanent nature (e.g. hairstyle, use of heavy make-up, presence of accessories). This is undesirable not only from the perspective of privacy but also from the perspective of performance since facial recognition systems should be robust against such variations.

Another motivation for these experiments is understanding which information is revealed by the facial features only. This has direct implications for the research that depends on facial images since many large datasets cannot be shared due to privacy reasons (e.g. civil identification datasets from governments). The development of facial recognition models that generates features that are "privacy-aware" could, therefore, make it viable to share the features computed from such datasets with the research community.

The remarkable improvement in the performance of facial recognition models during the last few years was accompanied by a similar improvement in the tasks related to facial attribute analysis, e.g., gender, age, emotion, and expression estimation. It is noteworthy that the two frameworks used in this paper[1] also provide pre-trained models for gender, age, and emotion estimation. These attribute estimation models were not evaluated in this report, though.

PREVIOUS WORKS

We briefly review some works related to attribute analysis based on deep neural networks. In 2014, (ZHANG et al., 2014) hair style, clothes style, expression, action proposed a method that combines part-based models and deep learning by training pose-normalized CNNs to estimate various attributes, such as gender, hair style, clothes style, expression, and action, using the whole image as input. (LIU et al., 2015) introduced LNet+ANet in 2015, with a focus on facial attributes - it takes only facial images as input. This work also introduced two new datasets,

1 Deepface (gender, age, emotion), available at https://github.com/serengil/deepface
Insightface (gender, age), available at https://github.com/deepinsight/insightface

LFWA and CelebA, with 40 labeled attributes, which contributed to the development of new methods, such as (KALAYEH; GONG; SHAH, 2017; MAHBUB; SARKAR; CHELLAPPA, 2020), which are part-based methods, and (CAO; LI; ZHANG, 2018; HAND; CHELLAPPA, 2017; RUDD; GÜNTHER; BOULT, 2016), which are holistic methods. For a more detailed review of facial attribute estimation methods and the taxonomy of part-based versus holistic methods, the reader is referred to the survey presented in (ZHENG et al., 2020).

It should be noted that all the works mentioned in the previous paragraph take an image as input and use CNNs to explicitly learn features to estimate attributes. This paper uses a different approach, which is based on the use of CNNs trained for facial recognition for the extraction of features, which are then used as input to attribute classifiers.

Using a similar approach, (PARDE et al., 2017) demonstrated that face representations (akin to the features used in this report) could reveal information on the yaw and pitch angles of the face, and also if the facial image originates from a still image or video. Other works (BEST-ROWDEN; JAIN, 2018; HERNANDEZ-ORTEGA et al., 2019; TERHORST et al., 2020) showed that facial image quality could also be inferred from these features.

In 2016, (ZHONG; SULLIVAN; LI, 2016a, 2016b) demonstrated that mid-level features from CNNs trained for facial recognition could also be used to estimate facial attributes.

In 2018, (ALVI; ZISSERMAN; NELLÅKER, 2019) presented an algorithm that could remove biases from features obtained from CNNs trained for image classification and demonstrated an improvement in accuracy for ethnicity estimation in a novel dataset presented in the same work.

NEURAL NETWORKS

As previously stated, this report uses features obtained from facial recognition models as input to classifiers built on fully connected neural networks. In this kind of network, each layer i receives input from all neurons in layer $(i-1)$, and the value of its neurons can be modeled as $h_i=\sigma(W_{i-1}h_{i-1}+B_i)$, where W is a matrix of weights, B is the bias vector and σ is a non-linearity. The weights and biases are the parameters that can be trained in a neural network.

Such training is performed using the back-propagation algorithm (RUMELHART; HINTON; WILLIAMS, 1986). Quoting the authors, this algorithm *"repeatedly adjusts the weights of the connections in the network so as to minimize a measure of the difference between the actual output vector of the net and the desired output vector"*. This description makes it clear that a measure of the difference between the model output and the desired output is necessary. Such measure is called *loss function* (L) and one fundamental requirement of this function is that it must be *differentiable*, at least in the range of outputs being considered.

The training process of a neural network aims to minimize the loss function[2], which is done by backpropagating the gradient of the loss function with respect to the weights (hence the requirement of the loss function being differentiable). The gradient describes the direction of greatest increase of a function, while the negative of the gradient provides the direction of "steepest" descent (or minimization). The chain rule of differential calculus is used to compute the gradients of the loss with respect to all neurons in every layer. The weights are then updated using the negative of the gradient multiplied by a factor, called the learning rate η - see Eq. (1), and the process is repeated with another sample,

2 In most cases the loss function is a proxy for the desired performance metric that one seeks to optimize.

in the case of stochastic gradient descent, with another batch (a subset of the training dataset), or with the entire training dataset (in the case of gradient descent). Additional techniques for updating the weights and biases can be used, like variable learning rate and regularization.

$$\Delta w_{ij} = -\eta \frac{\partial L}{\partial w_{ij}} \quad (1)$$

METHODS AND EXPERIMENTAL SETUP

The experiments conducted for this paper consisted in (i) extracting features/ representations from facial images using pre-trained facial recognition models; (ii) training gender and ethnicity classifiers using the features extracted in the previous stage; and (iii) test the trained classifiers on data from multiple datasets. Figure 1 presents a diagram with the main stages of this pipeline.

We discard the last stage of the facial recognition models and use the features computed from such systems as input to gender (female and male) and ethnicity (african, asian, caucasian and indian) classifiers.

Although these categories may not represent all possibilities for both gender and ethnicity (MERLER et al., 2019; ROBINSON et al., 2020), they were chosen for the practical reason that multiple datasets are labeled with these categories, which allows for a comprehensive evaluation.

This scheme can be interpreted as an application of transfer learning, where the initial layers from the facial recognition model are kept fixed and the new layers, corresponding to the classifiers, are trained. We consider this pipeline differently because the features obtained from facial recognition models are usually stored in databases in the form of facial templates, that can be more efficiently stored and compared than images itself. It makes sense, then, to consider the classifiers independently from the facial recognition layers.

FACIAL IMAGE DATASETS

The datasets used in this paper are detailed in Table 1. Gender labels for the LFW (HUANG et al., 2007) dataset were obtained from (AFIFI; ABDELHAMED, 2017) and the ethnicity labels for the Fairface (KARKKAINEN; JOO, 2021) dataset were treated as follows: *Latino* and *Middle Eastern* were discarded, while *Southeast Asian* and *East Asian* were merged and labeled as *Asian*. This was done to allow for comparison with the RFW (WANG et al., 2019) and UTKFaces (ZHANG; SONG; QI, 2017) datasets, both of which only have four categories for ethnicity: *African*, *Asian*, *Caucasian* and *Indian*. All classifiers were trained using the *train* partition of the FairFace dataset.

FACIAL RECOGNITION MODELS

Three pre-trained facial recognition models are used in accordance with the pipeline in Figure 1: DeepFace (TAIGMAN et al., 2014), FaceNet (SCHROFF; KALENICHENKO; PHILBIN, 2015) and Arcface (DENG et al., 2019). Both DeepFace and FaceNet were heavily influenced by the architecture of AlexNet (KRIZHEVSKY; SUTSKEVER; HINTON, 2012), winner of the 2012 ImageNet classification challenge, that was based on five convolutional layers followed by three fully connected layers.

Deepface was introduced in 2014 and was the first model to achieve near human performance (97.35% accuracy) on the LFW benchmark. Composed by nine layers, its pre-processing stage is relatively complex, with 3D alignment of face images and the penultimate layer, from which we obtain the features, has a relatively large size (4096 dimensions).

FaceNet was introduced in 2015 and achieved 99.63% accuracy on LFW. It applies

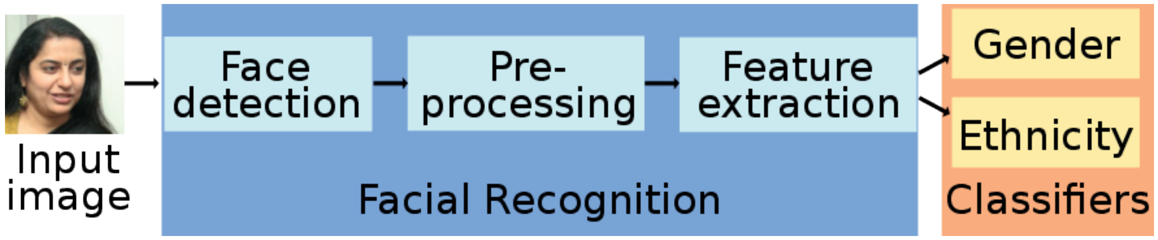

Figure 1 - The pipeline with the stages of the experiments in this paper.

Dataset	# of images	Gender	Ethnicity
FairFace	108 K	Yes	Yes
RFW	40 K	No	Yes
LFW	13K	Yes	No
UTKFaces	23 K	Yes	Yes
CelebA	202 K	Yes	No

Table 1 - Datasets used in this paper.

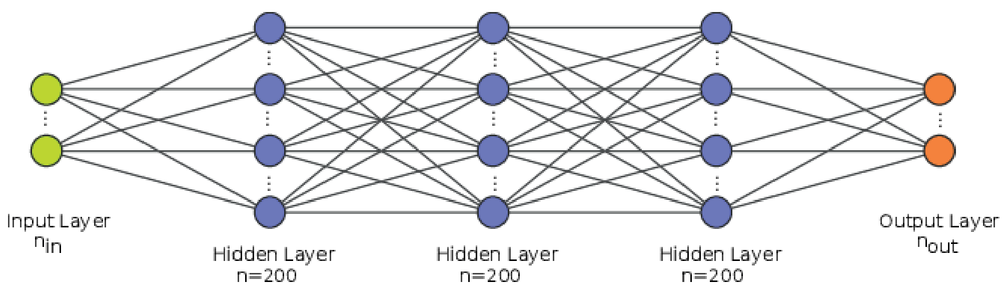

Figure 2 - Network architecture for the classifiers.

Face rec. model	LFW			CelebA			UTKFaces			FairFace (val.)		
	F	M	F+M	F	M	F+M	F	M	F+M	F	M	F+M
ArcFace	.865	.970	.946	.778	.970	.890	.877	.910	.894	.940	.953	.947
DeepFace	.864	.883	.879	-	-	-	.752	.841	.799	.802	.832	.818
FaceNet	.975	.966	.968	.975	.960	.966	.903	.900	.901	.884	.868	.876

Table 2 - Accuracies for gender. F = females, M = males, F+M = all

Face rec. model	UTKFaces					FairFace (val.)					RFW				
	A	Af	As	Ca	In	A	Af	As	Ca	In	A	Af	As	Ca	In
ArcFace	.810	.772	.873	.843	.712	.876	.866	.944	.863	.774	.754	.699	.901	.817	.607
DeepFace	.614	.529	.616	.762	.332	.685	.717	.726	.625	.656	.713	.817	.799	.731	.512
FaceNet	.846	.809	.873	.926	.661	.815	.855	.890	.804	.644	.894	.953	.927	.935	.762

Table 3 - Accuracies for ethnicity. A = all, Af = african, As = asian, Ca = caucasian, In = indian.

Face rec. model	Gender avg (std. dev.)	Ethnicity avg (std. dev.)
ArcFace	.919 (.027)	.813 (.050)
DeepFace	.832 (.034)	.671 (.042)
FaceNet	.928 (.040)	.852 (.032)

Table 4 - verage accuracies for gender and ethnicity.

a simpler pre-processing stage, with only a tight crop of the face area and uses more layers than DeepFace, but its main novelty was the introduction of *triplet-loss*, which optimizes the network for minimizing the distance between positive pairs and maximizing the distance between negative pairs. Features obtained from the penultimate layer of FaceNet have 128 dimensions.

ArcFace, introduced in 2019, uses a ResNet (HE et al., 2016) backbone and introduced a new loss function *additive angular margin*), further improving the results on popular benchmarks (99.83% on LFW). The features generated by ArcFace have 512 dimensions.

NEURAL NETWORK CLASSIFIERS

The classifiers for gender and ethnicity were built as fully connected neural networks, with the input layer of size n_{in} = *dimension of features*, followed by three hidden layers, each with 200 neurons and the output layer of the same size as the number of categories (two for gender and four for ethnicity). Rectified linear unit (ReLU) was used as activation function (σ). The architecture is depicted in Figure 2.

For training, the Cross-Entropy loss function (see Eq. 2) is used for both the gender and ethnicity classifiers. The choice of Cross-Entropy as the loss function is justified by the faster convergence that it allows, when compared to other loss functions, like Mean-Squared Error. A discussion of this property of Cross-Entropy is available in the Chapter 3 of (NIELSEN, 2018).

$$CE = -\sum_{i}^{C} y_i \, log(\hat{y}_i) \qquad (2)$$

In this equation, C is the number of classes, y_i is the label for class $\hat{y}_i$ and is the model prediction for class i. When using batches of size m, the average cross-entropy over the batch samples is used for the optimization step.

The adaptive moment estimation - Adam (KINGMA; BA, 2015) - algorithm is used as optimizer, with default parameters β_1 (0.9) and β_2 (0.999). This algorithm "computes individual adaptive learning rates for different parameters from estimates of first and second moments of the gradients" (quoting the authors of Adam). The initial learning rate and the weight decay hyper-parameters were determined by trial and error, based on the average accuracy in a 5-fold cross-validation strategy, considering the different combinations of features (ArcFace, FaceNet, DeepFace) and attributes (gender, ethnicity). The value of $1x10^{-3}$ was used for both parameters for all six different classifiers that were trained. Although this is not optimal, we choose to use the same value for all six classifiers for practical reasons.

According to the theorem of universal approximation (HORNIK, 1991), a simpler network could be used to approximate any continuous function, but we opted to use more layers in order to have a similar network architecture as (TERHÖRST et al., 2020). Different architectures (varying the number of hidden layers and the number of neurons per layer) were experimented with, but, in general, the results were very similar to each other, with one or two percent differences in accuracy, depending on the tested dataset. This invariance in performance with relation to the network architecture was also observed in (TERHÖRST et al., 2020).

RESULTS

We evaluate the classifiers using the accuracy metric, since three of the five tested datasets are well balanced. Only LFW and CelebA are heavily unbalanced for gender, with the majority of the images being from males. Accuracy is also the metric used in the work (KARKKAINEN; JOO, 2021) that introduced the FairFace dataset, which we used for training the classifiers. A direct comparison with some of the results in that work is then possible.

Table 2 summarizes the results for the gender attribute, while Table 3 presents the results for ethnicity. For each dataset, both the overall and per-class accuracies are presented. Values in bold represent improvements when compared with the results from Tables 2 and 3 from (KARKKAINEN; JOO, 2021). The classifier trained with DeepFace features was not evaluated on the CelebA dataset due to hardware limitations.

The results are in agreement with previous works, like (TERHÖRST et al., 2020), that showed that demographic information could be inferred from deep facial features with a high degree of accuracy. It is also not surprising that gender and ethnicity information are encoded in these features, since facial appearance is used, both by humans and image-based classifiers, to estimate gender and ethnicity.

DISCUSSION AND FUTURE WORKS

We observe that features extracted from models with very similar performance for facial recognition yield very different performance for gender and ethnicity estimation. This suggests that some facial recognition models "leak" more demographic information into its features than others. From the three facial recognition models used, FaceNet seems to be the one with more demographic information encoded in its features. This result is also in accordance with the findings of (TERHÖRST et al., 2020), which used features from FaceNet and ArcFace. Table 4 summarizes the average performance for the gender and ethnicity classifiers, over the tested datasets, according to the features employed.

We also note the very low variability of the gender estimation using FaceNet features between the females and males sub-groups of each dataset. Since we used features extracted from the same dataset to train all classifiers, we conjecture that this gender balanced performance is due to the facial recognition model, and not due to the facial image dataset. This verification remains as future work.

In spite of this good and stable performance of the classifiers trained with FaceNet features, we should also note that three classifiers trained with ArcFace features achieved superior performance than that presented in the work of (KARKKAINEN; JOO, 2021), which introduced the FairFace dataset. This is a clear indication that ArcFace features also encode a lot of demographic information in its features.

DeepFace features, on the contrary, seemed to be less revealing of demographic attributes. If this is in fact true remains an open question, since we only explored a limited number of classifiers architectures. A possibility of further investigation is to try different classifiers with DeepFace features. Also, the performance of DeepFace for facial recognition is worse than both ArcFace and FaceNet. It may be the case that DeepFace simply encodes less facial information in general.

There is another aspect that prevents further conclusions of which facial recognition architecture encodes more demographic information: the facial recognition models we used for feature extraction were not trained with the same dataset. ArcFace was trained with the MS1M-Arcface dataset, FaceNet was trained with VGGFace2, and the DeepFace model was trained with an unknown dataset. This may have influenced the results, but the extent of such influence, if any, remains as future work, repeating the same procedure employed here, but using facial recognition models trained with the same facial image dataset.

Other possible extensions of this work are a more fine-grained evaluation of ethnicity, including groups that were fused or ignored in this report, the evaluation of the age attribute, and the investigation of temporary attributes, such as facial expressions, hair style, the presence of adorns and the use of makeup.

REFERENCES

AFIFI, M.; ABDELHAMED, A. AFIF4: Deep Gender Classification based on AdaBoost-based Fusion of Isolated Facial Features and Foggy Faces. **CoRR**, v. abs/1706.04277, 2017.

ALVI, M.; ZISSERMAN, A.; NELLÅKER, C. **Turning a Blind Eye: Explicit Removal of Biases and Variation from Deep Neural Network Embeddings**. (L. Leal-Taixé, S. Roth, Eds.)Computer Vision – ECCV 2018 Workshops. **Anais**...Cham: Springer International Publishing, 2019.

BEST-ROWDEN, L.; JAIN, A. K. Learning Face Image Quality From Human Assessments. **IEEE Transactions on Information Forensics and Security**, v. 13, n. 12, p. 3064–3077, 2018.

CAO, J.; LI, Y.; ZHANG, Z. **Partially Shared Multi-task Convolutional Neural Network with Local Constraint for Face Attribute Learning**. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. **Anais**...2018.

DENG, J. et al. **ArcFace: Additive Angular Margin Loss for Deep Face Recognition**. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). **Anais**... Em: 2019 IEEE/CVF CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR). Long Beach, CA, USA: IEEE, jun. 2019. Disponível em: <https://ieeexplore.ieee.org/document/8953658/>. Acesso em: 28 mar. 2022

HAND, E. M.; CHELLAPPA, R. **Attributes for Improved Attributes: A Multi-Task Network Utilizing Implicit and Explicit Relationships for Facial Attribute Classification**. Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence. **Anais**...: AAAI'17.AAAI Press, 2017.

HE, K. et al. **Deep Residual Learning for Image Recognition**. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). **Anais**... Em: 2016 IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR). Las Vegas, NV, USA: IEEE, jun. 2016. Disponível em: <http://ieeexplore.ieee.org/document/7780459/>. Acesso em: 22 jan. 2023

HERNANDEZ-ORTEGA, J. et al. **FaceQnet: Quality Assessment for Face Recognition based on Deep Learning**. 2019 International Conference on Biometrics (ICB). **Anais**...2019.

HORNIK, K. Approximation capabilities of multilayer feedforward networks. **Neural Networks**, v. 4, n. 2, p. 251–257, 1991.

HUANG, G. B. et al. **Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments.** Amherst: University of Massachusetts, out. 2007. Disponível em: <http://vis-www.cs.umass.edu/lfw/lfw.pdf>. Acesso em: 22 mar. 2021.

KALAYEH, M. M.; GONG, B.; SHAH, M. **Improving Facial Attribute Prediction Using Semantic Segmentation**. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). **Anais**...2017.

KARKKAINEN, K.; JOO, J. **FairFace: Face Attribute Dataset for Balanced Race, Gender, and Age for Bias Measurement and Mitigation**. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. **Anais**...2021.

KINGMA, D. P.; BA, J. **Adam: A Method for Stochastic Optimization**. (Y. Bengio, Y. LeCun, Eds.)3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings. **Anais**...2015. Disponível em: <http://arxiv.org/abs/1412.6980>

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. **ImageNet Classification with Deep Convolutional Neural Networks**. (F. Pereira et al., Eds.)Advances in Neural Information Processing Systems. **Anais**...Curran Associates, Inc., 2012. Disponível em: <https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>

LIU, Z. et al. **Deep Learning Face Attributes in the Wild**. 2015 IEEE International Conference on Computer Vision (ICCV). **Anais**...2015.

MAHBUB, U.; SARKAR, S.; CHELLAPPA, R. Segment-Based Methods for Facial Attribute Detection from Partial Faces. **IEEE Transactions on Affective Computing**, v. 11, n. 4, p. 601–613, 2020.

MERLER, M. et al. Diversity in Faces. **CoRR**, v. abs/1901.10436, 2019.

NIELSEN, M. A. **Neural Networks and Deep Learning**. Determination Press, , 2018. Disponível em: <http://neuralnetworksanddeeplearning.com/>

PARDE, C. J. et al. **Face and Image Representation in Deep CNN Features**. 2017 12th IEEE International Conference on Automatic Face Gesture Recognition (FG 2017). **Anais**...2017.

PARDE, C. J. et al. Social Trait Information in Deep Convolutional Neural Networks Trained for Face Identification. **Cognitive Science**, v. 43, n. 6, p. e12729, 2019.

ROBINSON, J. P. et al. **Face Recognition: Too Bias, or Not Too Bias?** 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). **Anais**... Em: 2020 IEEE/CVF CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS (CVPRW). Seattle, WA, USA: IEEE, jun. 2020. Disponível em: <https://ieeexplore.ieee.org/document/9150783/>. Acesso em: 28 mar. 2022

RUDD, E. M.; GÜNTHER, M.; BOULT, T. E. **MOON: A Mixed Objective Optimization Network for the Recognition of Facial Attributes**. (B. Leibe et al., Eds.)Computer Vision – ECCV 2016. **Anais**...Cham: Springer International Publishing, 2016.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning Representations by Back-propagating Errors. **Nature**, v. 323, n. 6088, p. 533–536, 1986.

SCHROFF, F.; KALENICHENKO, D.; PHILBIN, J. **FaceNet: A unified embedding for face recognition and clustering**. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). **Anais**...2015.

TAIGMAN, Y. et al. **DeepFace: Closing the Gap to Human-Level Performance in Face Verification**. 2014 IEEE Conference on Computer Vision and Pattern Recognition. **Anais**...2014.

TERHÖRST, P. et al. **Beyond Identity: What Information Is Stored in Biometric Face Templates?** 2020 IEEE International Joint Conference on Biometrics (IJCB). **Anais**...2020.

TERHORST, P. et al. **SER-FIQ: Unsupervised Estimation of Face Image Quality Based on Stochastic Embedding Robustness**. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). **Anais**...Los Alamitos, CA, USA: IEEE Computer Society, jun. 2020.

WANG, M. et al. **Racial Faces in the Wild: Reducing Racial Bias by Information Maximization Adaptation Network**. The IEEE International Conference on Computer Vision (ICCV). **Anais**...out. 2019.

ZHANG, N. et al. **PANDA: Pose Aligned Networks for Deep Attribute Modeling**. Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. **Anais**...: CVPR '14.USA: IEEE Computer Society, 2014.

ZHANG, Z.; SONG, Y.; QI, H. **Age Progression/Regression by Conditional Adversarial Autoencoder**. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). **Anais**...2017.

ZHENG, X. et al. A Survey of Deep Facial Attribute Analysis. **International Journal of Computer Vision**, v. 128, set. 2020.

ZHONG, Y.; SULLIVAN, J.; LI, H. **Face attribute prediction using off-the-shelf CNN features**. 2016 International Conference on Biometrics (ICB). **Anais**...2016a.

ZHONG, Y.; SULLIVAN, J.; LI, H. **Leveraging mid-level deep representations for predicting face attributes in the wild**. 2016 IEEE International Conference on Image Processing (ICIP). **Anais**...2016b.