2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops

AgeDB: the first manually collected, in-the-wild age database

Stylianos Moschoglou
Imperial College London
s.moschoglou@imperial.ac.uk

Athanasios Papaioannou
Imperial College London
a.papaioannou11@imperial.ac.uk

Christos Sagonas
Imperial College London, Onfido
c.sagonas@imperial.ac.uk

Jiankang Deng
Imperial College London
j.deng16@imperial.ac.uk

Irene Kotsia
Middlesex University London
i.kotsia@mdx.ac.uk

Stefanos Zafeiriou
Imperial College London
s.zafeiriou@imperial.ac.uk

Abstract

Over the last few years, increased interest has arisen with respect to age-related tasks in the Computer Vision community. As a result, several "in-the-wild" databases annotated with respect to the age attribute became available in the literature. Nevertheless, one major drawback of these databases is that they are semi-automatically collected and annotated and thus they contain noisy labels. Therefore, the algorithms that are evaluated in such databases are prone to noisy estimates. In order to overcome such drawbacks, we present in this paper the first, to the best of knowledge, manually collected "in-the-wild" age database, dubbed AgeDB, containing images annotated with accurate to the year, noise-free labels. As demonstrated by a series of experiments utilizing state-of-the-art algorithms, this unique property renders AgeDB suitable when performing experiments on age-invariant face verification, age estimation and face age progression "in-the-wild".

1. Introduction

Adamantly, one of the most challenging and important tasks in Computer Vision throughout the years has been face recognition. This spark of interest in face recognition was firstly attributed to security reasons, since automatic facial analysis would assist security agencies in detection of passport frauds, identification of criminals or missing children, restriction of identity thefts, etc. The first algorithm for face recognition was introduced in mid 1960's [3] and since then numerous approaches have been proposed in the literature. Before the advent of deep learning in 2006 [16], construction of algorithms that were used in facial analysis tasks required huge amount of time, domain specific knowledge and a delicate engineering process in order to transform the raw facial data in a feature space. The derived features would then be utilized to produce the final classification result [24, 12].

Nevertheless, successful implementation of face recognition applications also requires sufficiently large facial datasets that will be utilized in order to train the algorithms. For decades and prior to the proliferation of deep learning, various facial datasets were introduced in the literature. One common feature of all these datasets was that they contained images which were captured under controlled conditions (e.g., common background, controlled lighting setting, etc.). This restriction was imposed due to the fact that the feature extractors did not perform well on "in-the-wild" datasets (i.e., datasets that included images captured in uncontrolled conditions). Some of the most widely used databases which included images captured under controlled conditions were the XM2VTS database [30], the Multi-PIE database [13, 14], the AR face database [27], the Caltech faces database [2], the FERET database [35, 34, 38], the Yale face database [11].

During the last few years, an explosion in scientific research with respect to the development of deep learning architectures for face recognition has been witnessed. Deep learning comprises a set of methods that are able to automatically discover the patterns that may exist in raw data and, as a result, feature extractors which were utilized to transform the raw data are no longer required [24, 12]. This competitive advantage of deep learning methods against the conventional algorithms led to the introduction of several "in-the-wild" databases in the literature. The term "in-the-wild" is used to refer to databases that contain images which

2160-7516/17 $31.00 © 2017 IEEE
DOI 10.1109/CVPRW.2017.250

have been captured under completely uncontrolled conditions (e.g., varying backgrounds, existence of noise in the pictures, existence of occlusions in the faces depicted in various images, different types of cameras used to capture the images, etc.).

Over the past decade, various "in-the-wild" facial databases became publicly available. More specifically, in 2007 the Labeled Faces "in-the-wild" (LFW) database [17, 23] was introduced. LFW contains13,233 images of 5,749 individuals, where 1,680 subjects are depicted in two or more images and the rest appear only in one image. In 2009, PubFig database [21] was introduced. PubFig is an "in-the-wild" database, containing 58,797 images of 200 subjects. In 2011, YouTube Faces (YTF) database [51] was introduced. YTF contains 3,425 videos of 1,595 individuals, where the average length of each video is 181.3 frames. In 2012, WDRef database [6] was introduced. WDRef contains 99,773 images of 2,995 individuals, where 2,995 subjects have 15 or more images. In 2014, CelebFaces database [26, 48, 47] was introduced. CelebFaces contains 202,599 images of 10,177 identities (celebrities), where each identity has about 20 images. In 2014, CASIA-WebFace database [52] was introduced. CASIA-WebFace contains 494,414 images pertaining to 10,575 subjects. In 2015, VGG Face dataset [33] was introduced. VGG Face dataset contains 2.6M image of 2,622 distinct individuals. Moreover, in 2015, the IARPA Janus Benchmark A (IJB-A) [20] was introduced. IJB-A contains 5,712 images and 2,085 videos from 500 subjects. In 2015 and 2016, the MegaFace database [18, 31] was introduced and extended, respectively. MegaFace contains 4.7M images of 672K individuals. A table providing an overview of the most recently introduced aforementioned databases along with their essential statistics is presented in Table 1.

Due to the recent surge of deep learning, age estimation from facial images has gradually gathered increased interest in the community. As of today, various deep learning architectures have been proposed and several databases annotated with regard to the age attribute have been made publicly available. The first database which contained images annotated with respect to the age attribute was FG-NET [22], introduced in 2002. FG-NET includes 1,002 images, captured under controlled conditions, pertaining to 82 subjects with accurate to the year age annotations. In 2006, MORPH database [37] was introduced. MORPH contains 1,724 images, pertaining to 464 subjects with accurate to the year age annotations. In 2008, the UIUC-IFP-Y Internal Aging Database [9, 15] was introduced, containing 8,000 images, captured under controlled conditions, pertaining to 1600 subjects with accurate to the year age annotations. In 2009, Gallagher group photos [10] was introduced. Gallagher group photos is an "in-the-wild" database containing 28,231 images of 5,080 subjects. As far as the annotation with the regard to the age attribute is concerned, 7 distinct age groups are utilized. Each image is annotated with a unique group identifier. In 2011, VADANA database [46] was introduced. VADANA is an "in-the-wild" database containing 2,298 images pertaining to 43 subjects. In total 4 age groups are utilized and each image is annotated with a unique group identifier. In 2014, AdienceFaces database [8, 25] was introduced, containing 26,580 "in-the-wild" images of 2,984 subjects. In total 8 age groups are utilized and each image is annotated with a unique group identifier. In 2014, Cross-Age Celebrity Dataset (CACD) [4, 5] was introduced, containing 163,446 "in-the-wild" images of 2,000 celebrities. Images are annotated with a specific age label which is semi-automatically estimated. In 2015, IMDB-WIKI database [39, 40] was introduced, containing 523,051 "in-the-wild" images pertaining to 20,284 celebrities. Images are annotated with a specific age label which is semi-automatically estimated. A table providing an overview of the aforementioned databases along with the newly introduced in this paper AgeDB database is presented in Table 2.

Nevertheless, despite of the increased interest in age estimation "in-the-wild" and the several databases that came into existence to tackle this task over the last years, no manually collected "in-the-wild" database with accurate to the year age annotations has been introduced in the literature. In order to fill this gap in the literature, we present in this paper the first, to the best of our knowledge, manually collected "in-the-wild" age database, dubbed AgeDB. AgeDB contains images of several subjects annotated with accurate to the year age labels. The fact that AgeDB is manually collected to ensure the accuracy of the age labels comes with several advantages:

- AgeDB can be used in age-invariant face verification "in-the-wild" experiments, i.e., the sensitivity in the performance of face recognition algorithm can be measured as the age gap between instances (images) of the same subject increases. Since the age labels are clean, AgeDB ensures a noise-free evaluation of the various face recognition algorithms.
- AgeDB can be used in age estimation "in-the-wild" experiments. Since the age labels are clean, AgeDB may be utilized as a benchmark database for such tasks.
- AgeDB can be utilized in face age progression "in-the-wild" experiments, since it is a manually collected database with large range of ages for each subject. This property renders AgeDB highly beneficial when training models for age progression experiments.

The structure of the paper is summarized as follows. In Section 2, we provide all the necessary details pertaining to the AgeDB database. In Section 3, we present the various

Table 1: Concise overview of the most broadly used "in-the-wild" facial datasets in the Computer Vision community since 2007 onwards.

Dataset	Year	# Images	# Videos	# Subjects	"In-the-wild"
LFW [17, 23]	2007	13,233	–	5,749	Yes
PubFig [21]	2009	58,797	–	200	Yes
YTF [51]	2011	–	3,425	1,595	Yes
WDRef [6]	2012	99,773	–	2,995	Yes
CelebFaces [26, 48, 47]	2014	202,599	–	10,177	Yes
CASIA-WebFace [52]	2014	494,114	–	10,575	Yes
VGG Face [33]	2015	2.6M	–	2,622	Yes
IJB-A [20]	2015	5,712	2,085	500	Yes
MegaFace [18, 31]	2016	4.7M	–	672K	Yes

Table 2: Concise overview of the most broadly used age datasets in the Computer Vision community since 2002 onwards.

Dataset	Year	# Images	# Subjects	Age labels	Noise-free labels	"In-the-wild"
FG-NET [22]	2002	1,002	82	Accurate to the year	Yes	No
MORPH [37]	2006	1,724	464	Accurate to the year	Yes	No
IFP-Y [9, 15]	2008	8,000	1600	Accurate to the year	Yes	No
Gallagher [10]	2009	28,231	5,080	7 age groups	No	Yes
VADANA [46]	2011	2,298	43	4 age groups	Yes	Yes
AdienceFaces [8, 25]	2014	26,580	2,984	8 age groups	Yes	Yes
CACD [4, 5]	2014	163,446	2,000	Accurate to the year	No	Yes
IMDB-WIKI [39, 40]	2015	523,051	20,284	Accurate to the year	No	Yes
AgeDB	2017	16,516	570	Accurate to the year	Yes	Yes

experiments we performed on AgeDB. More specifically, we perform: i) various age-invariant face verification "in-the-wild" experiments utilizing state-of-the-art pre-trained deep networks and report their performance on AgeDB, ii) various age estimation "in-the-wild" experiments utilizing state-of-the-art pre-trained deep networks and report their performance on AgeDB, iii) several face age-progression "in-the-wild" experiments and report their results on AgeDB.

2. The AgeDB database

In this section we thoroughly discuss all the details pertaining to the collection of the AgeDB database as well as provide the necessary statistics related to the database. The database is publicly available at: https://ibug.doc.ic.ac.uk/resources/agedb/.

As aforementioned, AgeDB is a *manually* collected database to ensure that the age labels are clean as opposed to other age databases which have been semi-automatically collected utilizing crawlers and in which the age labels may be noisy. In order to achieve noise-free annotation with respect to the age labels, we manually searched for images through Google Images and subsequently kept only images where the exact, accurate to the year age of each depicted subject is *explicitly* mentioned in the accompanied caption of the image. An indicative example of the process followed is provided in Fig. 1.

Moreover, AgeDB is an "in-the-wild" database, meaning it contains images captured under completely uncontrolled, real-world conditions (i.e., having different poses, containing noise, bearing various expressions, containing occlu-

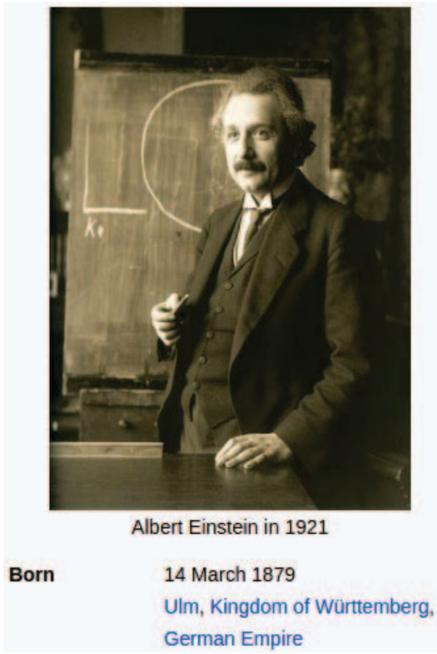

Figure 1: Screenshot captured from the Wikipedia lemma about Albert Einstein (http://bit.ly/AgeDB_fig1). Since the exact year the image was captured is available and also the birth year is provided, the age label can be subsequently calculated. For the image depicted in the screenshot, the age label is 42.

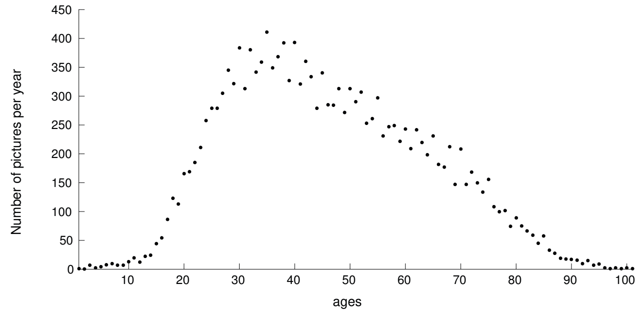

Figure 2: Scatter plot depicting the age distribution in the AgeDB "in-the-wild" database.

Figure 3: Random images from the AgeDB "in-the-wild" database.

sions, etc.). This feature may prove highly advantageous, since most of the state-of-the-art deep networks are trained and evaluated in "in-the-wild" databases.

AgeDB contains $16,488$ images of various famous people, such as actors/actresses, writers, scientists, politicians, etc. Every image is annotated with respect to the identity, age and gender attribute. There exist a total of 568 distinct subjects. The average number of images per subject is 29. The minimum and maximum age is 1 and 101, respectively. The average age range for each subject is 50.3 years. A scatter plot depicting the age distribution of the database is presented in Fig. 2. Samples from the AgeDB "in-the-wild" database along with their labels are provided in Fig. 3.

3. Experiments

In this section we present various experiments we performed on AgeDB. More specifically, we utilize state-of-the-art algorithms and conduct experiments in tasks as age-invariant face verification "in-the-wild", age estimation "in-the-wild", face age progression "in-the-wild" and subsequently show that AgeDB constitutes a proper benchmark for evaluating state-of-the-art algorithms for the previously mentioned tasks.

3.1. Age-invariant face verification "in-the-wild"

AgeDB may be utilized for age-invariant face verification "in-the-wild" experiments. To this end, following the procedure of the verification protocol described in the LFW [17] database, we developed four new age-invariant face verification protocols based on the AgeDB "in-the-wild" database. More specifically, for each protocol we split AgeDB in 10 folds, with each fold consisting of 300 intra-class and 300 intra-class pairs. The main difference among the protocols is that in each protocol the age difference of each pair's faces is equal to a fixed, predefined value, i.e., $5, 10, 20$ and 30 years. In Table 4, we report the results from the utilization of the VGG Face deep network. In Table 3, we report the results from the Center Loss [50] and Marginal Loss [7] methods. Thorough details of the experimental process that was followed are provided in [7]. It should be noted that this series of experiments was conducted on a subset of the final version of the AgeDB, as AgeDB was further extended by the time it became publicly available.

3.2. Age estimation "in-the-wild"

Age estimation "in-the-wild" is a challenging problem in Computer Vision and has recently gained huge interest in the community, mainly due to the increased penetration of deep learning techniques in the literature. The challenging nature of this task is primarily attributed to the fact that the databases which are publicly available and utilized for age estimation have been semi-automatically collected and thus contain age annotations that are noisy. As a result, age prediction based on such databases cannot be accurate, since the algorithms are trained on data where the labels in the first place are not accurate. To overcome the said disadvantage, we introduce the AgeDB "in-the-wild" database, the first manually collected age database.

The state-of-the-art publicly available pre-trained deep network for age estimation "in-the-wild" is DEX (Deep EXpectation of apparent age from a single image) [39, 40], winner of the LAP challenge 2015 on apparent age estimation [39]. We hence utilized the publicly available DEX pre-trained deep network [39] and performed age estimation in the totality of pictures included in AgeDB.

Preprocessing

In order to feed the images of AgeDB in the pre-trained network, we followed the preprocessing process described in [39]: We firstly employed the Mathias et al. face detection algorithm [28] and then used an extra 40% margin on the initial bounding box boundary. Moreover, we discarded the images in which the face detector was not able to extract a proper bounding box containing the face. We then used the cropped images as input in the pre-trained deep network.

Age estimation

As mentioned in [39], the output layer of the deep network corresponds to a 101 dimensional vector $\mathbf{v}$, representing softmax output probabilities, one for each age between 0 and 100 included. The final age estimation is given by the softmax expected value, i.e.:

$$\mathbb{E}\left[\mathbf{v}\right] = \sum_{n=0}^{100} y_n \cdot v_n, \tag{1}$$

where y_i are the years corresponding to the n-th class. A graphical overview of the processed followed in provided in Fig. 4.

Evaluation

As evaluation protocol we used the standard Mean Absolute Error (MAE), which is defined as the average of the absolute errors between the estimated age and the ground truth age. For the AgeDB "in-the-wild" database, the MAE for the pre-trained DEX deep network [39] is 13.1 years.

3.3. Face age progression "in-the-wild"

Face age progression consists in synthesizing plausible faces of subjects at different ages. It is considered as a very challenging task due to the fact that the face is a highly deformable object and its appearance drastically changes under different illumination conditions, expressions and poses. As mentioned in the introductory section, various databases that contain faces at different ages have been collected in the last couple of years. Although some of these databases possess huge number of images, they have some limitations including limited images for each subject that cover a narrow range of ages and noisy age labels, since most of them have been collected by employing automatic procedures (crawlers). AgeDB overcomes the aforementioned shortcomings.

We performed face age-progression "in-the-wild" experiments utilizing the Robust Joint and Individual Variance Explained (RJIVE) model [42] and other state-of-the-art algorithms for face progression, as mentioned below. In order to train RJIVE, AgeDB was split into $M = 10$ age groups: $0-3, 4-7, 8-15, 16-20, 21-30, 31-40, 41-50, 51-60, 61-70$ and $71-100$. As a next step, in order to effectively recover the joint and common components of the images, the faces of each dataset should be put in correspondence. Therefore, their $N = 68$ facial landmarks points are localized using the face detector from [1], trained with images provided from 300-W challenge [44, 41, 43] and subsequently employed to compute a mean reference shape. Moreover, the faces of each dataset are warped into corresponding reference shape by using the piecewise affine warp function $\mathcal{W}$ [29]. RJIVE was then employed to extract the joint and common components from the warped images. The performance of RJIVE in face age progression "in-the-wild" is qualitatively assessed conducting experiments on images from the FG-NET dataset [22]. To this end, we compare the performance of RJIVE with the Illumination Aware Age Progression (IAAP) method [19], Coupled Dictionary Learning (CDL) method [45], Deep Aging with Restricted Boltzmann Machines (DARB) method [32], CG [36], and Recurrent Face Aging (RFA) method [49]. In Fig. 5, progressed images produced by the compared methods are depicted. Note that all progressed faces have been warped back and fused with the actual ones.

4. Conclusion

In this paper we introduced the AgeDB database, the first *manually* collected, "in-the-wild" age database, which contains noise-free identity, gender and accurate to the year age labels. Moreover, we utilized AgeDB along with state-of-the-art algorithms and performed a series of experiments in

Table 3: Age-invariant face verification in AgeDB utilizing the Centre Loss [50] and Marginal Loss [7] methods.

Age gap	Accuracy per method	
	Center Loss	Marginal Loss
5	0.959	0.981
10	0.951	0.979
20	0.931	0.971
30	0.907	0.957

Table 4: Age-invariant face verification in AgeDB utilizing the VGG Face [33] deep network.

Age gap	Accuracy per layer			
	33-rd	34-th	35-th	36-th
5	0.912	0.931	0.911	**0.934**
10	0.900	**0.922**	0.902	0.914
20	0.879	**0.891**	0.879	0.888
30	0.839	**0.851**	0.842	0.840

tasks such as age-invariant face verification "in-the-wild", age estimation "in-the-wild", face age progression "in-the-wild". Finally, we showed that AgeDB can be utilized as a benchmark for evaluating state-of-the-art algorithms that aim to tackle the aforementioned tasks.

5. Acknowledgements

The authors would like to thank Mr. Polychronis Zafeiriou for the meticulous collection and annotation of the AgeDB. S. Moschoglou was supported by the EPSRC DTA studentship from Imperial College London. A. Papaioannou was funded by the European Community Horizon 2020 [H2020/2014-2020] under grant agreement no. 688520 (TeSLA). J. Deng was supported by the President's Scholarship of Imperial College London. S. Zafeiriou was partially funded by the EPSRC Project EP/N007743/1 (FACER2VM).

References

[1] J. Alabort-i Medina, E. Antonakos, J. Booth, P. Snape, and S. Zafeiriou. Menpo: A comprehensive platform for parametric image alignment and visual deformable models. In *Proceedings of the ACM International Conference on Multimedia*, pages 679–682. ACM, 2014. 5

[2] A. Angelova, Y. Abu-Mostafam, and P. Perona. Pruning training sets for learning of object categories. In *Proceedings of IEEE International Conference on Computer Vision & Pattern Recognition (CVPR)*, volume 1, pages 494–501. IEEE, 2005. 1

[3] W. W. Bledsoe and H. Chan. A man-machine facial recognition system: some preliminary results. *Panoramic Research, Inc, Palo Alto, California., Technical Report PRI A*, 19:1965, 1965. 1

[4] B.-C. Chen, C.-S. Chen, and W. H. Hsu. Cross-age reference coding for age-invariant face recognition and retrieval. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 768–783. Springer, 2014. 2, 3

[5] B.-C. Chen, C.-S. Chen, and W. H. Hsu. Face recognition and retrieval using cross-age reference coding with cross-age celebrity dataset. *IEEE Transactions on Multimedia*, 17(6):804–815, 2015. 2, 3

[6] D. Chen, X. Cao, L. Wang, F. Wen, and J. Sun. Bayesian face revisited: A joint formulation. *Proceedings of the European Conference in Computer Vision (ECCV)*, pages 566–579, 2012. 2, 3

[7] J. Deng, Y. Zhou, and S. Zafeiriou. Marginal loss for deep face recognition. In *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition (CVPR-W), Faces "in-the-wild" Workshop/Challenge*, 2017. 4, 6

[8] E. Eidinger, R. Enbar, and T. Hassner. Age and gender estimation of unfiltered faces. *IEEE Transactions on Information Forensics and Security*, 9(12):2170–2179, 2014. 2, 3

[9] Y. Fu, Y. Xu, and T. S. Huang. Estimating human age by manifold analysis of face pictures and regression on aging features. In *Proceedings of the IEEE International Conference on Multimedia and Expo*, pages 1383–1386. IEEE, 2007. 2, 3

[10] A. C. Gallagher and T. Chen. Understanding images of groups of people. In *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 256–263. IEEE, 2009. 2, 3

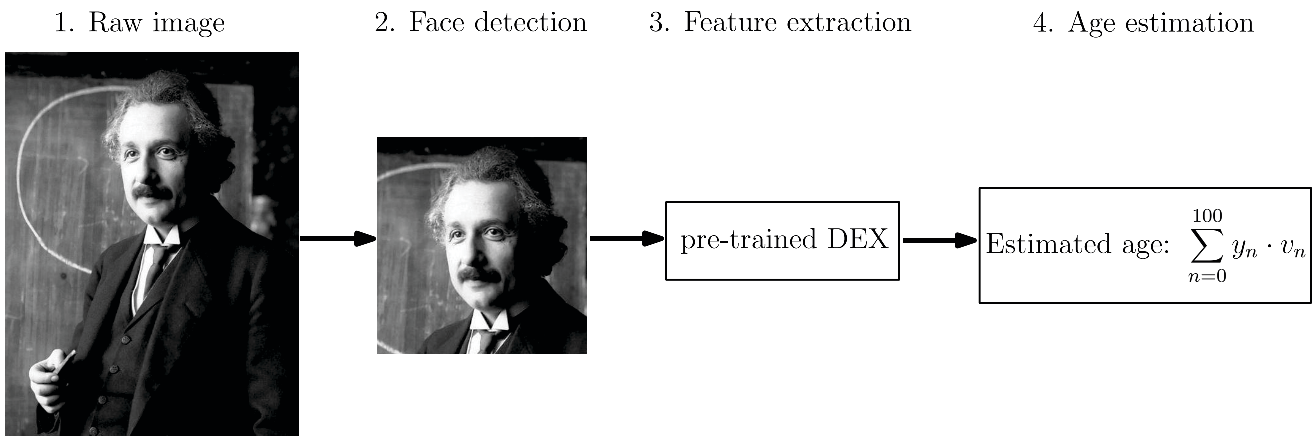

Figure 4: Visualization of the complete process for age estimation utilizing the pre-trained DEX deep network [39].

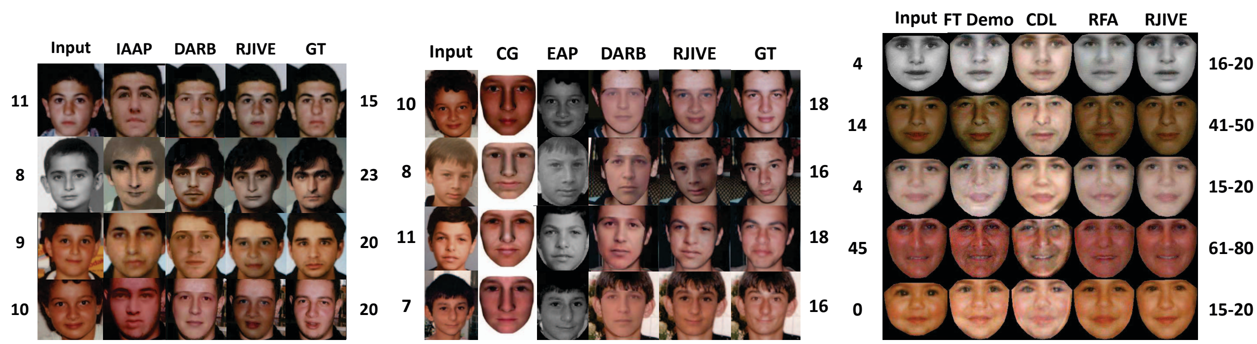

Figure 5: Progressed faces produced by the compared methods on the FG-NET [22] dataset. RJIVE is trained on the AgeDB database.

[11] A. Georghiades, P. Belhumeur, and D. Kriegman. Yale face database. *Center for computational Vision and Control, Yale University*, 2, 1997. 1

[12] I. Goodfellow, Y. Bengio, and A. Courville. *Deep learning*. MIT Press, 2016. 1

[13] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker. The cmu multi-pose, illumination, and expression (multi-pie) face database. *Technical report, Carnegie Mellon University Robotics Institute. TR-07–08*, 2007. 1

[14] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker. Multi-pie. *Image and Vision Computing (IMAVIS)*, 28(5):807–813, 2010. 1

[15] G. Guo, Y. Fu, C. R. Dyer, and T. S. Huang. Image-based human age estimation by manifold learning and locally adjusted robust regression. *IEEE Transactions on Image Processing*, 17(7):1178–1188, 2008. 2, 3

[16] G. E. Hinton. Learning multiple layers of representation. *Trends in cognitive sciences*, 11(10):428–434, 2007. 1

[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. Technical report, Technical Report 07-49, University of Massachusetts, Amherst, 2007. 2, 3, 4

[18] I. Kemelmacher-Shlizerman, S. M. Seitz, D. Miller, and E. Brossard. The megaface benchmark: 1 million faces for recognition at scale. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4873–4882, 2016. 2, 3

[19] I. Kemelmacher-Shlizerman, S. Suwajanakorn, and S. M. Seitz. Illumination-aware age progression. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3334–3341, 2014. 5

[20] B. F. Klare, B. Klein, E. Taborsky, A. Blanton, J. Cheney, K. Allen, P. Grother, A. Mah, and A. K. Jain. Pushing the frontiers of unconstrained face detection and recognition: Iarpa janus benchmark a. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1931–1939, 2015. 2, 3

[21] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar. Attribute and simile classifiers for face verification. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 365–372. IEEE, 2009. 2, 3

[22] A. Lanitis and T. Cootes. Fg-net aging data base. *Cyprus College*, 2002. 2, 3, 5, 7

[23] E. Learned-Miller, G. B. Huang, A. RoyChowdhury, H. Li, and G. Hua. Labeled faces in the wild: A survey. In *Advances*

in Face Detection and Facial Image Analysis, pages 189–248. Springer, 2016. 2, 3

[24] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015. 1

[25] G. Levi and T. Hassner. Age and gender classification using convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR-W), Workshop on Analysis and Modeling of Faces and Gestures*, pages 34–42, 2015. 2, 3

[26] Z. Liu, P. Luo, X. Wang, and X. Tang. Deep learning face attributes in the wild. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 3730–3738, 2015. 2, 3

[27] A. M. Martinez. The ar face database. *CVC technical report*, 24, 1998. 1

[28] M. Mathias, R. Benenson, M. Pedersoli, and L. Van Gool. Face detection without bells and whistles. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 720–735. Springer, 2014. 5

[29] I. Matthews and S. Baker. Active appearance models revisited. *International Journal of Computer Vision (IJVC)*, 60(2):135–164, 2004. 5

[30] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre. Xm2vtsdb: The extended m2vts database. In *Second international conference on audio and video-based biometric person authentication*, volume 964, pages 965–966, 1999. 1

[31] D. Miller, E. Brossard, S. Seitz, and I. Kemelmacher-Shlizerman. Megaface: A million faces for recognition at scale. *arXiv preprint arXiv:1505.02108*, 2015. 2, 3

[32] C. Nhan Duong, K. Luu, K. Gia Quach, and T. D. Bui. Longitudinal face modeling via temporal deep restricted boltzmann machines. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5772–5780, 2016. 5

[33] O. M. Parkhi, A. Vedaldi, and A. Zisserman. Deep face recognition. In *Proceedings of the British Machine Vision Conference (BMVC)*, volume 1, page 6, 2015. 2, 3, 6

[34] P. J. Phillips, S. Z. Der, P. J. Rauss, and O. Z. Der. *FERET (face recognition technology) recognition algorithm development and test results*. Army Research Laboratory Adelphi, MD, 1996. 1

[35] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss. The feret evaluation methodology for face-recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 22(10):1090–1104, 2000. 1

[36] N. Ramanathan and R. Chellappa. Modeling age progression in young faces. In *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 387–394. IEEE, 2006. 5

[37] K. Ricanek and T. Tesafaye. Morph: A longitudinal image database of normal adult age-progression. In *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*, pages 341–345. IEEE, 2006. 2, 3

[38] S. A. Rizvi, P. J. Phillips, and H. Moon. The feret verification testing protocol for face recognition algorithms. In *Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition*, pages 48–53. IEEE, 1998. 1

[39] R. Rothe, R. Timofte, and L. Van Gool. Dex: Deep expectation of apparent age from a single image. In *Proceedings of the IEEE International Conference on Computer Vision, Looking at People (LAP) Workshop*, pages 10–15, 2015. 2, 3, 5, 7

[40] R. Rothe, R. Timofte, and L. Van Gool. Deep expectation of real and apparent age from a single image without facial landmarks. *International Journal of Computer Vision (IJCV)*, pages 1–14, 2016. 2, 3, 5

[41] C. Sagonas, E. Antonakos, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic. 300 faces in-the-wild challenge: database and results. *Image and Vision Computing (IMAVIS)*, 47:3–18, 2016. 5

[42] C. Sagonas, A. Leidinger, Y. Panagakis, and S. Zafeiriou. Robust joint and individual variance explained. In *Proceedings of IEEE International Conference on Computer Vision & Pattern Recognition (CVPR)*, Hawaii, USA, 2017. 5

[43] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic. 300 faces in-the-wild challenge: The first facial landmark localization challenge. In *Proceedings of IEEE International Conference on Computer Vision (ICCV-W), Workshop on 300 Faces in-the-Wild Challenge (300-W)*, pages 397–403, Sydney, Australia, 2013. 5

[44] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic. A semi-automatic methodology for facial landmark annotation. In *Proceedings of IEEE International Conference on Computer Vision & Pattern Recognition (CVPR-W), 5th Workshop on Analysis and Modeling of Faces and Gestures (AMFG)*, pages 896–903, 2013. 5

[45] X. Shu, J. Tang, H. Lai, L. Liu, and S. Yan. Personalized age progression with aging dictionary. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 3970–3978, 2015. 5

[46] G. Somanath, M. Rohith, and C. Kambhamettu. Vadana: A dense dataset for facial image analysis. In *Proceedings of IEEE International Conference on Computer Vision Workshop (ICCV-W)*, pages 2175–2182. IEEE, 2011. 2, 3

[47] Y. Sun, Y. Chen, X. Wang, and X. Tang. Deep learning face representation by joint identification-verification. In *Advances in neural information processing systems*, pages 1988–1996, 2014. 2, 3

[48] Y. Sun, X. Wang, and X. Tang. Deep learning face representation from predicting 10,000 classes. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1891–1898, 2014. 2, 3

[49] W. Wang, Z. Cui, Y. Yan, J. Feng, S. Yan, X. Shu, and N. Sebe. Recurrent face aging. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2378–2386, 2016. 5

[50] Y. Wen, K. Zhang, Z. Li, and Y. Qiao. A discriminative feature learning approach for deep face recognition. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 499–515. Springer, 2016. 4, 6

[51] L. Wolf, T. Hassner, and I. Maoz. Face recognition in unconstrained videos with matched background similarity. In *Proceddings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 529–534. IEEE, 2011. 2, 3

[52] D. Yi, Z. Lei, S. Liao, and S. Z. Li. Learning face representation from scratch. *arXiv preprint arXiv:1411.7923*, 2014. 2, 3