https://doi.org/10.2352/ISSN.2470-1173.2017.17.COIMG-432

Evaluating Age Estimation Using Deep Convolutional Neural Nets

C. Belver [a]***, I. Arganda-Carreras*** [a,b]***, and F. Dornaika*** [a,b]
[a] ***University of the Basque Country UPV/EHU, San Sebastian, Spain***
[b] ***IKERBASQUE, Basque Foundation for Science, Bilbao, Spain***

Abstract

Age estimation from a facial image is still a challenge due to the variations caused by different aging processes, face appearance, human expression, and face pose. In this paper, we provide a comparative study for age estimation using classic image features as well as deep image features that are provided by pre-trained deep Convolutional Neural Networks. The presented work compares several image features. The experiments are conducted on two face datasets: MORPH II and PAL. In the light of the conducted experiments, image features that are providing the best performances can be highlighted.

Introduction

In the last decade, with the increasing interest in social robotics and video-based security systems, research on the numerical analysis of human faces (including face detection, face recognition, classification of gender, and recognition of facial expression) has attracted attention in the communities of computer vision and pattern recognition [9, 26, 20, 25, 29]. In connection with these investigations, estimating the age of a person from the numerical analysis of his face image is a relatively new topic. Age estimation by numerical analysis of the face image has many potential applications such as the development of intelligent human-machine interfaces and improving safety and protection of minors in various and diverse sectors (transport, medicine, etc.). It can be very useful for advanced video surveillance, demographic statistics collection, business intelligence and customer profiling, and search optimization in large databases. The age attribute could also be used in the verification of the face and enriching the tools used in police investigations. In general, automatic age estimation by a machine is useful in applications where the objective is to determine the age of an individual without identifying him. The age estimator can use a machine learning approach to train a model for extracted features and make age prediction for query faces with the trained model. Generally speaking, age estimation can be viewed as a multi-class classification problem, a regression problem or a composite of these two.

The anthropometry-based approach mainly depends on measurements and distances of different facial landmarks. Kwon and Lobo [24] published the earliest paper on age classification based on facial images by computing ratios to distinguish babies from others. In [12], the authors designed a neural network to locate facial features and calculate several geometric ratios and differences which are used for estimating the age from a facial image. The anthropometry-based approaches might be useful for babies, children, and young adults, but they are impractical for adults since their facial skin appearance is the main source of information about ethnicity, gender, and age. Estimating human age from a facial image requires a great amount of information from the input image. Extraction of these features is important since the performance of an age estimation system will heavily rely on the quality of extracted features. Lots of research on age estimation has been conducted towards aging feature extraction. Examples include: the active appearance model (AAM) [7], age manifold [9], AGing pattern Subspace (AGES) [11], biologically inspired features (BIF) [17]. Image-based age estimation approaches view the face image as a texture pattern. Many texture features have been used like Local Binary Patterns (LBP) [3], Histograms of Oriented Gradients (HOG) [8], BIF, Binarized Statistical Image Features (BSIF) [21] and Local Phase Quantization (LPQ) in demographic estimation works. BIF and its variants are widely used in age estimation works such us [18, 16, 19]. Han et al. [19] used selected BIF features on order to estimate the age, gender and ethnicity attributes.

Due to their significant performance improvement in facial recognition domain, deep learning approaches have been recently proposed for age estimation (e.g. [20, 29]). Deep learning approaches claim to have the best performances in demographic classification (ethnicity, gender and age). However, this claim cannot be always true. It is known, that deep learning can provide impressive results within a single database. However when another database is used with the trained deep net, the age estimation performance can drop significantly. In this paper, we provide a comparative study of age estimation based on hand-crafted and deep features. The paper also shows that the full power of a pre-trained net can be exploited by simply using its deep features and only training an age regressor. This regressor training is much more efficient than retraining or fine-tuning the whole deep net using a sheer number of images. Experiments will show that this scheme for age estimation can be more accurate than the one obtained by the end-to-end deep net solution.

The rest of the paper is organized as follows. Firstly, 2D face alignment is briefly introduced. Secondly, we summarize three classic image features and five deep features. Thirdly, the experimental setup and results are given.

Face Alignment

Face alignment is one of the most important stages in image-based age estimation. In our experiments, the eyes of each face are detected using the Ensemble of Regression Trees (ERT) algorithm [22] which is a robust and very efficient algorithm for facial landmarks localization. Once we have the 2D positions of the two eyes, we use them to compensate for the in-plane rotation of the face. To this end, within the detected face region, the po-

Original MORPH-II image

Face alignment + cropping

Figure 1. Face alignment and cropping associated with one original image in MORPH II database.

sitions of right and left eyes are located as (R_x, R_y) and (L_x, L_y), respectively. Then, the angle of in-plane rotation is calculated by $\theta = artan(\frac{R_y - L_y}{R_x - L_x})$, and the input face region is rotated by the that angle. After rotation correction, we use a global scale for the face image, this scale normalizes the inter-ocular distance to a fixed value l. After performing the rotation and rescaling, the face region should be cropped (aligned face). To this end, a bounding box is centered on the new eyes location (on the transformed face image) and then stretched to the left and to the right by $k_0 \cdot l$, and to top by $k_1 \cdot l$ and to bottom by $k_2 \cdot l$. Finally, in our case, k_0, k_1, k_2 and l are chosen such that the final face image has a size of 50×50 pixels for the MORPH II database and 200×200 for the PAL database.

Face Features

In order to make the paper self-contained, this section will briefly describe some features that are very often used for extracting face features. We present three hand-crafted features as well as five deep features that can be obtained from pre-trained deep CNNs.

Local Binary Patterns (LBP)

The original LBP operator labels the pixels of an image with decimal numbers, which are called LBPs or LBP codes that encode the local structure around each pixel [1, 3]. The basic operator proceeds as follows. Each pixel is compared with its eight neighbors in a neighborhood by subtracting the central pixel value; the resulting strictly negative values are encoded with 0, and the others with 1. For each given pixel, a binary number is obtained by concatenating all these binary values in a clockwise direction, which starts from the one of its top-left neighbor. The corresponding decimal value of the generated binary number is then used for labeling the given pixel. The histogram of LBP labels (the frequency of occurrence of each code) calculated over a region or an image can be used as a texture descriptor. It should be noticed the LBP descriptors can be either an LBP image or a histogram of that image.

In our work, we used the classic LBP operator that provides a histogram of 256 bins for a given face image.

Histogram of Oriented Gradients (HOG)

The essential thought behind the histogram of oriented gradients descriptor [8] is that local object appearance and shape within an image can be described by the distribution of intensity gradients or edge directions. The image is divided into small connected regions called cells, and for the pixels within each cell, a histogram of gradient directions is compiled. The descriptor is then the concatenation of these histograms. For improved accuracy, the local histograms can be contrast-normalized by calculating a measure of the intensity across a larger region of the image, called a block, and then using this value to normalize all cells within the block. This normalization results in better invariance to changes in illumination and shadowing.

Binarized Statistical Image Features (BSIF)

This descriptor [21] can be used in texture recognition tasks in a similar manner as LBPs. Each element (i.e. bit) in the binary code string is computed by binarizing the response of a linear filter with a threshold at zero. Each bit is associated with a different filter and the desired length of the bit string is determined by the number of filters used. The set of filters is learnt from a training set of natural image patches by maximizing the statistical independence of the filter responses. Hence, statistical properties of natural image patches determine the BSIF descriptors.

Visual Geometry Group (VGG) Face features

This CNN comprises 11 blocks, each containing a linear operator followed by one or more non-linearities such as ReLU and max pooling [28]. The first eight such blocks are said to be convolutional as the linear operator is a bank of linear filters (linear convolution). The last three blocks are instead called Fully Connected (FC); they are the same as a convolutional layer, but the size of the filters matches the size of the input data, such that each filter senses data from the entire image. All the convolution layers are followed by a rectification layer (ReLU). The first two FC layers output are 4,096 dimensional vectors. This multi-way CNN is trained to discriminate between the 2,622 identities using about 2.6M images. The deep features of this network are extracted by taking the 4K dimensional features and removing the last classification layer. The resulting vector is L2 normalized.

ImageNet VGG-F features

The Fast (VGG-F) architecture [5] is similar to the one used by Krizhevsky *et al.* [23]. It comprises 8 learnable layers, 5 of which are convolutional, and the last 3 are fully-connected. The input image size is 224×224. Fast processing is ensured by the 4 pixel stride in the first convolutional layer. The main differences between this architecture and that of Krizhevsky are the reduced number of convolutional layers and the dense connectivity between convolutional layers (Krizhevsky used sparse connections to enable training on two GPUs). The network was trained on ILSVRC-2012 using gradient descent with momentum. The hyper-parameters are the same as used by Krizhevsky. The authors applied data augmentation in the form of random crops, horizontal flips, and RGB color jittering. We extracted the deep features from the 4K dimensional feature vector after removing the last classification layer. The resulting vector is L2 normalized. The only image pre-processing consists on resizing the input images to the network input size and subtracting the average image (provided by the authors in the network metadata).

ImageNet VGG-verydeep-16 features

This network is part of the evaluation of networks of increasing depth carried out by Simonyan and Zisserman [32] that proved to be very performant at the ImageNet 2014 challenge. The configuration is quite different from the ones used in the top-performing entries of the 2012 and 2013 competitions. Rather than using relatively large receptive fields in the first convolutional layers, they used very small 3×3 receptive fields throughout the whole net, which are convolved with the input at every pixel. More specifically, the convolution stride is fixed to 1 pixel; the spatial padding of convolutional layer input is such that the spatial resolution is preserved after convolution, i.e. the padding is 1 pixel for 3×3 convolutional layers. Spatial pooling is carried out by 5 max-pooling layers, which follow some of the convolutional layers (not all the convolutional layers are followed by max-pooling). Max-pooling is performed over a 2×2 pixel window, with stride 2. In total, the network we used has 13 convolutional layers and 3 FC. The only preprocessing we do is subtracting the mean RGB value of the input image. The 4K features are collected from the last FC layer and L2 normalized.

DEX-IMDB-WIKI and DEX-ChaLearn-ICCV2015 features

The Deep EXpectation (DEX) on apparent age method [30, 31] uses the VGG-16 architecture for its networks, which are pre-trained on ImageNet for image classification. In addition, the authors explored the benefit of fine-tuning over crawled Internet face images with available age. In total, they collected more than 500,000 images of celebrities from IMDb and Wikipedia. The networks of DEX were fine-tuned on the crawled images and then on the provided images with apparent age annotations from the ChaLearn LAP 2015 challenge on apparent age estimation. We extracted the features provided by two networks: DEX-IMDB-WIKI and DEX-ChaLearn-ICCV2015. The first one was trained on real age estimation using the cropped and aligned faces of the IMDB-WIKI dataset, while the second one is a fine-tuned version of the previous model, trained on apparent age using the challenge images. An ensemble of these models led to 1st place at the challenge (115 teams). The 4K features are collected from the previous to the last FC layer.

Experimental Results

Table 1 illustrates the MAE obtained on the MORPH II database using the eight face features. In this table, we considered two cases: original images and the aligned/cropped images. We can observe that with face alignment and cropping the performances obtained with the hand-crafted features have increased. This is very intuitive since the hand-crafted features need to focus on the face region only. On the other hand, for the last two deep features, the use of the original images provided better performance. This can be explained by the fact that these ones were trained on face images having significant background. Whether the original images or the aligned and cropped images were used, the deep features provided by DEX-IMDB-WIKI and DEX-ChaLearn-ICCV2015 nets provided the best performances. Moreover, we can observe that among deep features the best performances were obtained with nets that were trained on face images, i.e. VGG-Face, DEX-IMDB-WIKI and DEX-ChaLearn.

The performances of some state-of-the-art approaches are

(a)

(b)

Figure 2. *Database images.* ***(a)*** *Sample images from MORPH II database.* ***(b)*** *Sample images from PAL database.*

shown in Table 2. As can be seen, our deep feature results are comparable to the performance obtained by the work of Han *et al.* (2015) [19]. The latter uses coarse-to-fine and hierarchical age estimation via binary decision trees for classifying non-overlapping age groups and within-group age regressors. In our case, only one single regressor is used. Figure 3 represents the cumulative score associated with the eight face features. As can be seen, for some face features the cumulative scores are similar.

Table 3 illustrates the MEA obtained on the PAL database using the eight face features. In this table, we considered three cases: (i) original images, (ii) aligned images with loose crop (face plus some background), and (iii) aligned/cropped images. We can observe that with face alignment and cropping the performances obtained with the hand-crafted features have increased. The performances obtained by the last two deep features were the best for all three types of cropping. In general, the deep features gave their best results when loose cropping is adopted.

The performances of some state-of-the-art approaches are shown in Table 4. As can be seen, by adopting the proposed scheme, we got a significant improvement in performance. The best state-of-the art MAE was 5.4 years, whereas the best MAE obtained by our adopted scheme was 3.79 years. Figure 4 represents the cumulative score associated with the eight face features. As can be seen, for some face features the cumulative scores are very similar.

Table 5 illustrates a comparison between the MAE of the end-to-end CNNs and that obtained by the use of deep features. The table corresponds to the MORPH II database with two different types of images. For each CNN, the upper row illustrates the MAE obtained by applying the net in order to estimate the age. The lower row depicts the MAE where the net is used to provide only the deep features. As can be seen, by adopting the deep features the obtained MAE was better than that of the end-to-end CNN. Table 6 illustrates a comparison between the MAE of the end-to-end CNNs and that obtained by the use of deep features for PAL database. The table corresponds to the database with three different types of images. We can observe a similar behavior to that obtained with the MORPH II database. This tends to confirm that by only retraining the regressor, we are able to transfer the power of the pre-trained CNN without having to retrain the whole network.

Conclusion

The paper has addressed the issue of comparing several face features for the task of age estimation from facial images. In the study, we have considered three hand-crafted image features as well as five deep features provides by pre-trained CNNs. The comparison shown is the paper yields several conclusions. First, the solution adopted in the paper shows that efficient and stable age estimation can be obtained from deep features on the premise that the age regressor is retrained. The last process is by far more efficient than re-training the whole deep CNN on the new set of images. Second, the use of deep features gave better results than using hand-crafted features. Third, task accuracy obtained by deep features can be highly correlated to the deep net context (training imaged objects, training objective).

Table 1. Mean Age Error (years) obtained with different face features on MORPH II database.

Face features	Orig. images	Aligned+cropped
LBP	7.20	6.53
HOG	6.26	4.84
BSIF	7.34	6.69
VGG-FACE	4.72	4.79
IMAGENET-VGG-F	5.11	5.04
IMAGENET-VERY-DEEP-16	5.53	5.47
DEX-CHALEARN	**3.67**	4.77
DEX-IMDB-WIKI	3.77	**4.76**

Table 2. MAE (years) obtained with different state-of-the-art approaches on MORPH II database.

Publication	Approach	MAE
Guo and Mu (2011) [14]	*+KPLS†	4.2
Chang *et al.* (2011) [4]	BIF*	6.1
Geng *et al.* (2013) [10]	BIF*	4.8
Guo and Mu (2013) [15]	BIF*	4.0
Huerta *et al.* [20]	CNN‡	3.9
Han *et al.* (2015) [19]	DIF††	**3.6**
Proposed scheme	Deep features + transfer	3.67

* Biologically Inspired Features
† Kernel Partial Least Squares
‡ Convolutional Neural Networks
†† Demographic Informative Features

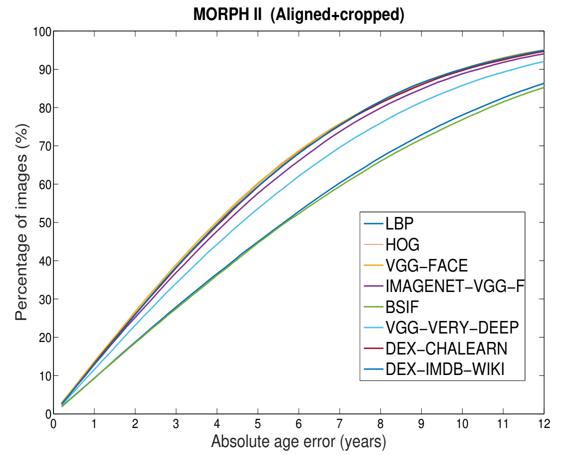

Figure 3. *Cumulative scores obtained with eight face features for MORPH II database (aligned and cropped images).*

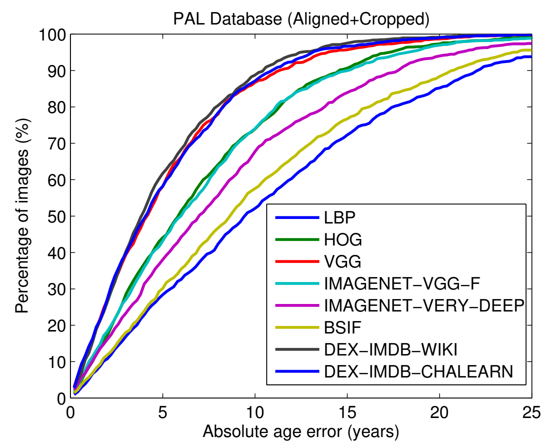

Figure 4. *Cumulative scores obtained with eight face features for PAL database (aligned and cropped images).*

References

[1] T. Ahonen, A. Hadid, and M. Pietikainen. Face description with local binary patterns: Application to face recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(12):2037–2041, Dec 2006.

[2] S. Bekhouche, A. Ouafi, A. Taleb-Ahmed, A. Hadid, and A. Benlamoudi. Facial age estimation using BSIF and LBP. In *International Conference on Electrical Engineering*, 2014.

[3] M. Bereta, P. Karczmarek, W. Pedrycz, and M. Reformat. Local descriptors in application to the aging problem in face recognition. *Pattern Recognition*, 46:2634–2646, 2013.

[4] K. Y. Chang, C. S. Chen, and Y. P. Hung. Ordinal hyperplanes ranker with cost sensitivities for age estimation. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 585–592, June 2011.

[5] K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman. Return of the devil in the details: Delving deep into convolutional nets. In *British Machine Vision Conference*, 2014.

[6] S. E. Choi, Y. J. Lee, S. J. Lee, K. R. Park, and J. Kim. A comparative study of local feature extraction for age estimation. In *International Conference on Control Automation Robotics* Vision, pages 1280–1284, 2010.

[7] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):681–685, 2001.

Table 3. Mean Age Error (years) obtained with different face features on PAL database.

Face features	Original images	Aligned+Loose crop	Aligned+crop
LBP	11.40	11.16	10.99
HOG	8.68	7.61	7.00
BSIF	10.71	11.26	10.09
VGG-FACE	5.91	5.13	5.23
IMAGENET-VGG-F	6.89	6.81	7.14
IMAGENET-VERY-DEEP-16	8.04	8.64	8.41
DEX-CHALEARN	**3.97**	**3.79**	5.12
DEX-IMDB-WIKI	4.04	**3.79**	**4.90**

Table 4. Mean Age Error (years) obtained with different state-of-the art approaches on PAL database.

Publication	Approach	MAE
Gunay and Nabiyev (2016) [13]	AAM+GABOR+LBP	**5.4**
Nguyen *et al.* (2014) [27]	MLBP+GABOR+SVR	6.5
Bekhouche *et al.* (2014) [2]	LBP+BSIF+SVR	6.2
Choi *et al.* (2010) [6]	GHPF*+SVR	8.4
Luu *et al.* (2011) [26]	CAM†+SVR	6.0
Proposed scheme	Deep features+transfer	**3.79**

* Gaussian High Pass Filter
† Contourlet Appearance Model
‡ Multi-Quantized Local Binary Patterns

Table 5. Mean Age Error (years) obtained with two deep CNNs on MORPH II database. For each CNN, the upper row illustrates the MEA obtained by applying the CNN as an end-to-end solution. The lower row depicts the MEA where the net is used to provide only the deep features.

CNN	Scheme	Original	Aligned+crop
DEX-CHALEARN	End-to-end	5.34	11.1
	Deep features	**3.67**	**4.77**
DEX-IMDB-WIKI	End-to-end	5.77	11.6
	Deep features	**3.77**	**4.76**

[8] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2005.

[9] Y. Fu and T. S. Huang. Human age estimation with regression on discriminative aging manifold. *IEEE Transactions on Multimedia*, 10(4):578–584, June 2008.

[10] X. Geng, C. Yin, and Z. H. Zhou. Facial age estimation by learning from label distributions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(10):2401–2412, Oct 2013.

[11] X. Geng, Z. H. Zhou, and K. Smith-Miles. Automatic age estimation based on facial aging patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(12):2234–2240, Dec 2007.

[12] A. Gunay and V. V. Nabiyev. Automatic detection of anthropometric features from facial images. In *2007 IEEE 15th Signal Processing and Communications Applications*, pages 1–4, June 2007.

[13] A. Günay and V. V. Nabiyev. *Information Sciences and Systems 2015: 30th International Symposium on Computer and Information Sciences (ISCIS 2015)*, chapter Age Estimation Based on Hybrid Features of Facial Images, pages 295–304. Springer International Publishing, Cham, 2016.

[14] G. Guo and G. Mu. Simultaneous dimensionality reduction and human age estimation via kernel partial least squares regression. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 657–664, June 2011.

[15] G. Guo and G. Mu. Joint estimation of age, gender and ethnicity: CCA vs. PLS. In *IEEE International Conference and Workshop on Automatic Face and Gesture Recognition*, pages 1–6, April 2013.

[16] G. Guo and G. Mu. A framework for joint estimation of age, gender and ethnicity on a large database. *Image and Vision Computing*, 32(10):761 – 770, 2014. Best of Automatic Face and Gesture Recognition 2013.

[17] G. Guo, G. Mu, Y. Fu, and T. S. Huang. Human age estimation using bio-inspired features. In *Computer Vision Pattern Recognition*, 2009.

[18] H. Han and A. K. Jain. Age, gender and race estimation from unconstrained face images. Technical Report MSU-CSE-14-5, Department of Computer Science, Michigan State University, East Lansing, Michigan, July 2014.

[19] H. Han, C. Otto, X. Liu, and A. K. Jain. Demographic estimation from face images: Human vs. machine performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(6):1148–1161, June 2015.

[20] I. Huerta, C. Fernndez, C. Segura, J. Hernando, and A. Prati. A deep analysis on age estimation. *Pattern Recognition Letters*, 68, Part 2:239 – 249, 2015. Special Issue on Soft Biometrics.

[21] J. Kannala and E. Rahtu. BSIF: Binarized statistical image features. In *Pattern Recognition (ICPR), 2012 21st International Conference on*, pages 1363–1366, Nov 2012.

[22] V. Kazemi and J. Sullivan. One millisecond face alignment with an ensemble of regression trees. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1867–1874, 2014.

[23] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[24] Y. H. Kwon and N. da Vitoria Lobo. Age classification from facial images. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 762–767, Jun 1994.

[25] G. Levi and T. Hassncer. Age and gender classification using convolutional neural networks. In *2015 IEEE Conference on Computer*

Table 6. **Mean Age Error (years) obtained with two deep CNNs on PAL database.**

CNN	Scheme	Original	Aligned+Loose crop	Aligned+crop
DEX-CHALEARN	End-to-end	7.12	5.43	8.53
	Deep features	**3.97**	**3.79**	**5.12**
DEX-IMDB-WIKI	End-to-end	6.99	4.72	7.98
	Deep features	**4.04**	**3.79**	**4.90**

Vision and Pattern Recognition Workshops (CVPRW), pages 34–42, June 2015.

[26] K. Luu, K. Seshadri, M. Savvides, T. D. Bui, and C. Y. Suen. Contourlet appearance model for facial age estimation. In *Biometrics (IJCB), 2011 International Joint Conference on*, pages 1–8, Oct 2011.

[27] D. T. Nguyen, S. R. Cho, K. Y. Shin, J. W. Bang, and K. R. Park. Comparative study of human age estimation with or without preclassification of gender and facial expression. *The Scientific World Journal*, 2014:15, 2014.

[28] O. M. Parkhi, A. Vedaldi, and A. Zisserman. Deep face recognition. In *British Machine Vision Conference*, volume 1, page 6, 2015.

[29] R. Ranjan, S. Zhou, J. C. Chen, A. Kumar, A. Alavi, V. M. Patel, and R. Chellappa. Unconstrained age estimation with deep convolutional neural networks. In *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*, pages 351–359, Dec 2015.

[30] R. Rothe, R. Timofte, and L. V. Gool. DEX: Deep expectation of apparent age from a single image. In *IEEE International Conference on Computer Vision Workshops (ICCVW)*, December 2015.

[31] R. Rothe, R. Timofte, and L. V. Gool. Deep expectation of real and apparent age from a single image without facial landmarks. *International Journal of Computer Vision (IJCV)*, July 2016.

[32] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

Author Biography

Carlos Belver obtained his bachelor's degree in Computer Science from the Basque Country University in 2015. He earned his Master degree in Computer Engineering and Intelligent Systems at the Basque Country University, Spain, in 2016.

Ignacio Arganda-Carreras is an Ikerbasque Research Fellow at the Computer Science and Artificial Intelligence department of the Basque Country University, Spain. He earned his Ph. D. in Computer Science and Telecommunications at the Universidad Autonoma de Madrid, Spain, in 2009. He was a postdoctoral fellow at the Massachusetts Institute of Technology, USA from 2009 to 2013, and at INRA, France, from 2013 to 2015.

Fadi Dornaika received his engineer degree in Electrical Engineering from the Lebanese University, in 1990, an M.S. degree in signal, image and speech processing from Grenoble Institute of Technology, France, in 1992, and a Ph.D. degree in computer science from Grenoble Institute of Technology, France and INRIA, in 1995. He has published more than 200 papers in the field of computer vision and pattern recognition. His current research interests include manifold learning, pattern recognition, and data mining.