An investigation on local wrinkle-based extractor of age estimation

An Investigation on Local Wrinkle-based Extractor of Age Estimation

Choon-Ching Ng, Moi Hoon Yap, Nicholas Costen and Baihua Li
School of Computing, Mathematics & Technology, Manchester Metropolitan University, Manchester, U.K.
{*choon.c.ng, m.yap, n.costen, b.li*}*@mmu.ac.uk*

Keywords: Age Estimation, Facial Wrinkles, Canny Edge Detection, FG-NET.

Abstract: Research related to age estimation using face images has become increasingly important due to its potential use in various applications such as age group estimation in advertising and age estimation in access control. In contrast to other facial variations, age variation has several unique characteristics which make it a challenging task. As we age, the most pronounced facial changes are the appearance of wrinkles (skin creases), which is the focus of ageing research in cosmetic and nutrition studies. This paper investigates an algorithm for wrinkle detection and the use of wrinkle data as an age predictor. A novel method in detecting and classifying facial age groups based on a local wrinkle-based extractor (LOWEX) is introduced. First, each face image is divided into several convex regions representing wrinkle distribution areas. Secondly, these areas are analysed using a Canny filter and then concatenated into an enhanced feature vector. Finally, the face is classified into an age group using a supervised learning algorithm. The experimental results show that the accuracy of the proposed method is 80% when using FG-NET dataset. This investigation shows that local wrinkle-based features have great potential in age estimation. We conclude that wrinkles can produce a prominent ageing descriptor and identify some future research challenges.

1 INTRODUCTION

Age estimation is an important processing task that serves many purposes. In marketing, companies may increase their profits by measuring the demographics of groups interested on their billboard or street advertising through age estimation. In security control and surveillance monitoring, an age estimation system, with the input of a monitoring camera, can warn or stop under-age drinkers from entering wine shops; prevent minors from purchasing tobacco products from vending machines; refuse the aged when he or she wishes to try a roller coaster at an amusement park; and deny children access to adult websites or restricted films (Lanitis et al., 2004). In addition, estimated age also provides a type of soft biometric information which provides ancillary parameters for user identity. It can be used to complement primary biometric features, such as face, fingerprint, iris, and hand geometry, to improve the performance of a primary (hard) biometrics system. Face-based authentication systems which typically compare age-separated face images, are also bound to benefit from facial ageing models and from faces (Lanitis et al., 2004).

The process of age estimation attempts to label a face image automatically with the exact age (year) or the age group (year range) of the individual face. By deriving significant features from faces of known ages, the age of an individual face can be estimated by solving the inverse problem using the same feature-extraction technique.

Age estimation methods can be broadly classified as global feature or local features approaches. In the former, a common way is to capture the variability in the facial appearance due to all systematic sources of variability. The Active Appearance Model (AAM) (Cootes et al., 2001) is a well-known model which represents faces with statistical appearance and shape models built using Principal Component Analysis.The first age estimation algorithm to use AAM features and regression methods (Lanitis et al., 2004) defined the relationship between age and features via quadratic ageing functions, and the facial age was then estimated using ageing features. Subsequently, AAM-based features have been widely investigated in age estimation, for example AGES (Geng et al., 2007), manifold learning (Yan et al., 2009), Gaussian process regression (Zhang and Yeung, 2010). Among recent approaches, the best mean absolute error (MAE) for the FG-NET database using a leave one person out (LOPO) approach was 4.25 when using AAM and biologically inspired features (BIF) (Chao et al., 2013).

Although the AAM-based features provide sufficient information for detailed age estimation, they do not include a comprehensive characterization of wrinkles or quantifiable wrinkles. The skin changes associated with ageing are the focus of many surgical and non-surgical procedures aimed to improve the appearance of skin (Khavkin and Ellis, 2011). It is also useful in age synthesis and skin rejuvenation. Knowledge of skin histology will deepen the understanding of cutaneous changes associated with ageing and will promote optimal cosmetic and functional patient outcomes. AAM features do not include information about wrinkles and skin, due to the dimensionality reduction done by the PCA (Choi et al., 2011). Moreover, the computation cost of global features is higher than that for equivalent local features. Therefore, it is worthwhile to study the effects of local wrinkle-based features on age estimation.

Local features or distinct age features such as shape-based cues (anthropometric) and texture-based cues (wrinkles, skin textures, geometric features) can be used in estimating one's age (Fu et al., 2010). For example, anthropometric distances extracted from different regions of the face and wrinkle density can help characterize facial growth (Kwon and da Vitoria Lobo, 1999). The advantage of using local features is its ability to describe ageing features without an influence of personal characteristics. It is better suited to age group classification than detailed age estimation (Choi et al., 2011). In this paper, we proposed a new local wrinkle-based method of age group classification. This method employs an Otsu's method threshold (Otsu, 1979) and Canny edge detection (Canny, 1986) to extract a suitable wrinkle representation for the estimation of facial ageing. A set of shape free images constructed from mean shape is used along with standardized landmarks to allow region of interest extraction and centre point localization. Finally, wrinkle length and the amount in a region are utilized as the feature pattern for age estimation.

The paper is organised into the following sections: Section II provides a brief overview of previous work relevant to the proposed method; Section III outlines the proposed method; Section IV presents the results and discussion; finally, conclusions are summarized in Section V.

2 RELATED WORK

In general, an age estimation algorithm can be divided into two steps: feature extraction and age prediction. In the first step, facial features related to human age or facial appearance changes caused by ageing are extracted from human faces to form a compact representation; in the second step, an age prediction function is built to estimate the age based on the extracted features.

Among all kinds of facial features, the first one utilized in age estimation is the anthropometric model, which is based on the domain knowledge of facial ageing processes, such as occurrence of wrinkles and the change of face shapes. In early work, the snake algorithm was exploited for wrinkle detection and this information was combined with some measures of facial geometry for an age range classification (Kwon and da Vitoria Lobo, 1999). The wrinkles were computed in several regions, such as on the forehead, next to the eyes, and near the cheek bones. The presence of wrinkles in a region was based on the detection of curves in that region. However, recently the anthropometric model was claimed not to be suitable for actual age estimation (Fu et al., 2010), as this model can only deal with younger ages since human head shape does not change very much in its adult period. In addition, the head profile is difficult to measure from 2-D face images. Therefore, it motivates our further investigation on local wrinkle-based extractor for age estimation.

AAM decouples and models two parts of an object: shape and texture. Adopting the AAM allowed exploration of combined shape and intensity model to represent face images (Lanitis et al., 2004). Face images were represented by means of lower dimension model parameters giving weightings on the principal components from the eigenspaces that correspond to facial shape and intensity. A number of approaches have been proposed to estimate one's age from such low-dimensional representations of faces. An 'ageing pattern' subspace ('AGES') has been proposed (Geng et al., 2007), defining the ageing pattern as a sequence of personal face images ordered in time, which further considers the identity information and the ordinal relationship of ages during feature extraction. When an novel face image, $\mathbf{I}$, is presented, its feature vector, $\mathbf{b}$, is first extracted by the feature extractor. Then, the procedure of age estimation involves two steps. First, an ageing pattern suitable for $\mathbf{I}$ is selected; this is achieved by finding a projection, $\mathbf{y}$, in the ageing pattern subspace. Then, the age associated with the position ω is returned as the estimated age of $\mathbf{I}$. Although the AAM-based approaches can in general deal with any age, but there are some limitations. For example, AAM features may involve many outliers in the age labelling space or high dimensional parameters. In addition, . some crucial ageing features such as wrinkles and skin textures are removed from

the AAM features in the dimension reduction process performed by the PCA and therefore, individual characteristics are stronger than the ageing characteristics (Choi et al., 2011). For that reason, this paper explores a simple and intuitive feature extractor as described in the following section.

3 PROPOSED METHOD

This work presents an effective local wrinkle-based extractor (LOWEX) which extracts wrinkles of interest from the facial images. Given a training set and an unknown image, the wrinkle features are extracted by LOWEX. Using that feature patterns, a multilayer perceptron (MLP) classifier is implemented to predict the age groups.

3.1 Local Wrinkle-based Extractor

Figure 1 illustrates the flow of the proposed method, LOWEX. All images are converted to a grey-scale level (D_g), reducing dimensionality and avoiding issues concerning colour balance. The shape is then modelled based on manually labelled landmarks. In 2-D cases, for instance the shape is represented by concatenating n point vectors, (x_i, y_i), into

$$\tilde{\mathbf{s}} = (x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n)^T, \quad \{\tilde{\mathbf{s}}\} \in D \quad (1)$$

The shapes are then normalized by the Procrustes analysis (Goodall, 1991) and projected onto the shape subspace created by PCA

$$\tilde{\mathbf{s}} = \tilde{\mathbf{s}}_0 + \mathbf{P}_{\tilde{\mathbf{s}}} \cdot \tilde{\mathbf{c}} \quad (2)$$

where $\tilde{\mathbf{s}}_0$ denotes the mean shape, $\mathbf{P}_{\tilde{\mathbf{s}}} = \tilde{\mathbf{s}}_j$ is the matrix consisting of a set of orthonormal base vectors $\tilde{\mathbf{s}}_j$, which describe the modes of variability derived from the training set, and $\tilde{\mathbf{c}}$ includes the shape parameters in the shape subspace. Subsequently, based on the corresponding points, images in the training set are warped to the mean shape (D_m) to produce 'shape-free patches'. Figure 2 shows the mean shape sample of the FG-NET dataset.

In the next step, the Otsu algorithm is applied to each region of interest (ROI) denoted as b_k ($\mathbf{b} \in D_m$) in order to discover an appropriate threshold (t) for that region.

From the mean shape face, a selected triangular region (b_k) which reflects to the wrinkle region is identified, as previously suggested (Kwon and da Vitoria Lobo, 1999). In each triangle, a square image is cropped, denoted as c_k ($\mathbf{c} \in \mathbf{b}$) with the size $l \times l$ and the centre point as the corresponding triangle centroid. Two reasons for using the square image are

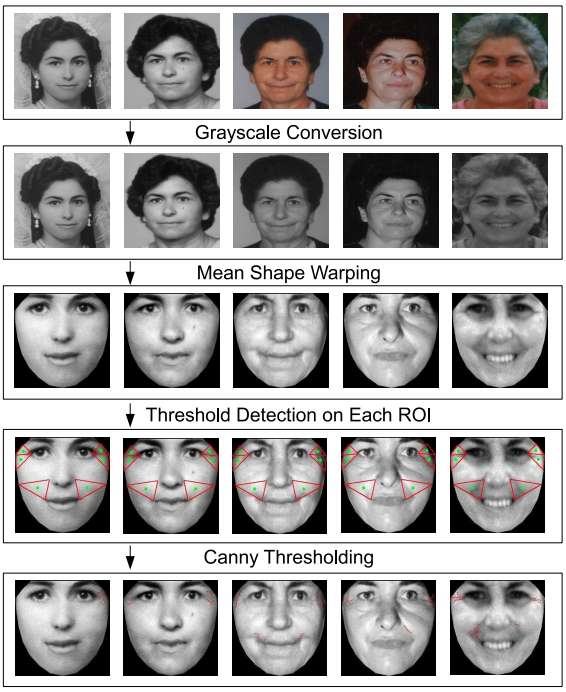

Figure 1: Information flow of the local wrinkle-based extractor for a single individual.

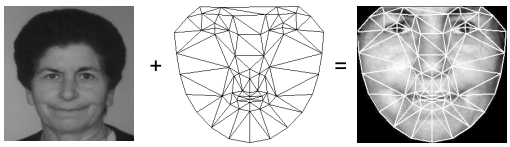

Figure 2: FG-NET facial model mapping and image to the mean shape.

to avoid noise around the triangle corners and allow derivation of stable thresholds for wrinkles. If the l value is either too large or too small, it will exceed the ROI or lost the focus of detection. Hence, he square image size l is set to 20 pixels.

Finally, Canny edge detection (Canny, 1986) is utilised $f(b_k)$ to determine the wrinkle lines $\mathbf{w}$. The wrinkle lines total length ($\mathbf{e}$) and amount ($\mathbf{u}$) are combined into a feature pattern,

$$\mathbf{f} = [e_1, \cdots, e_k, u_1, \cdots, u_k], \quad \{\mathbf{e}, \mathbf{u}\} \in \mathbf{w}. \quad (3)$$

3.2 MLP Classifier

Artificial Neural Networks (ANN) are fundamentally parallel processors. An ANN is a computer program which is biologically inspired to simulate some of the ways in which the human brain processes information. They have been applied to many applications including language identification due to their fascinating features, such as learning, generalizing, fast real-time computation and their modelling and clas-

sification capabilities (Rumelhart et al., 2002). The use of Multilayer Perceptrons (MLPs) with the back propagation learning algorithm for estimating the age of a subject given a set of face parameters has been investigated (Lanitis et al., 2004) . Based on the training set, each type of network was evaluated in order to establish the optimal architecture and optimal parameters. In each case, the generalization capability of the neural network was assessed as a function of the initial parameters of the respective network.

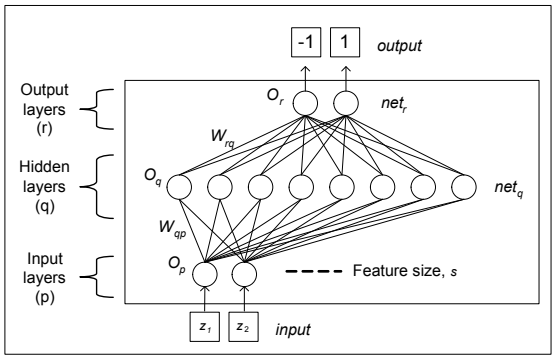

Figure 3: Back-propagation neural networks architecture.

Figure 3 shows the example of architecture of Back-propagation Neural Network (BPNN). This BPNN consists of one input layer (p), one hidden layer (q) and one output layer (r). The total number of nodes in the input layer is determined according to the feature size, s, used for capturing input patterns. If the feature size, s, is 24 then the number of input layers will be set correspondingly. There are two hidden units. The number of output layer units which is based on the number of outputs required, here again two. The neural networks parameters are the iteration number ℓ, a learning rate η and the momentum rate Γ. O_p is the output on unit p, O_q is the output on unit q, O_r is the output on unit r, W_{qp} is a q weight to the unit p, W_{rq} is an r weight to the unit q, net_q is the first transfer function at hidden layer q, net_r is the second transfer function at output layer r, θ_q is a bias on hidden unit q, θ_r is a bias on output unit r, δ_q is the generalized error through a layer q, and δ_r is the generalized error through a layer q and r. The input values for the back-propagation neural network are represented as $z_1, z_2, ..., z_s$, where each input is between -1 and 1 ($\mathbf{z} \in [-1,1]$), and s is the number of features which have been selected. The output values to the back-propagation neural network are represented by the binary code $[-1,1]$ which correspond to the age groups.

Table 1 shows the parameters setting on the BPNN. The input node, hidden node and output node are $\mathbf{s}$, $(\mathbf{s}+2)/2$, and 2, respectively; learning rate is 0.3; momentum rate is 0.2; epochs is 500; and features and output are normalized between -1 and 1.

Table 1: The BPNN structure.

Description	Value
Function	Backpropagation
Input Node	**s**
Hidden Node	(**s** + output node) / 2
Output Node	Group A / Group B
Learning Rate	0.3
Momentum Rate	0.2
Epochs	500
Features Normalized	-1 to 1
Output Normalized	-1 to 1

4 EXPERIMENTAL SITUATION

The FG-NET ageing database was used to evaluate the performance of the proposed method. This is one of databases used more frequently in previous work for estimating age, as it is publicly available. The database has 1002 images comprising 82 participants with an the age range of 0-69 years. All individuals in the database have more than one image included with different ages. Each image has 68 annotated facial feature points; these were used for the shape features. Images were obtained by scanning photographs, unlike other databases such as BERC (Choi et al., 2011) were captured using a digital camera with fixed light and position conditions. As a result, there are extreme variations in lighting, expression, background, pose, resolution and noise from scanning. Figure 4 illustrates some samples of the type of variation seen in the FG-NET dataset. Based on human observation, the first row are clear images and the second row is blurred images which do not have sufficient texture information such as wrinkles. Thus, clear images were randomly selected from the FG-NET database as experiment samples.

Normal ageing of the facial soft tissues begins in the 20s with the fine facial lines appearing horizon-

Figure 4: The FG-NET dataset (FG-NET aging database, 2000).

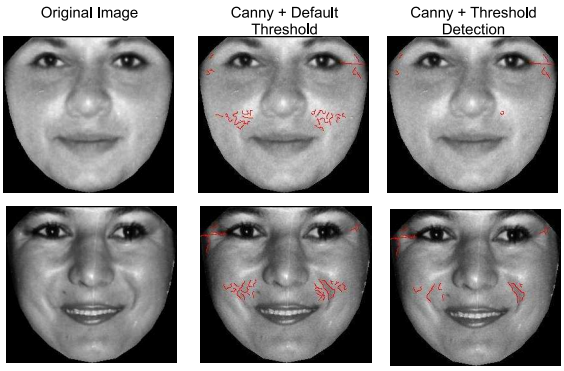

Figure 5: Comparison between default threshold and threshold detection of Canny edge detection.

tally across the forehead, vertical lines emerging between eyebrows, and faint lines developing around the outer corners of the eyes (Albert et al., 2007). With this in mind, the experiment set was divided into two into two groups: group A was between age 0 and 20 while group B was above 20 years old. In the experiment 1 and 2, 20 images were selected from group A and 20 images of group B, all of which had frontal pose and clear texture. In the experiment 3, these images were used as the training set and the remaining images of FG-NET were used for testing. 10-folds cross validation was performed to evaluate the performance of the age group classification. For ROI extraction, all 68 facial feature points were used to allow warping to the mean shape.

5 RESULTS AND DISCUSSION

This paper used Canny edge detector to find the wrinkles in the region of interest. The result of this edge detector is a binary image in which the white pixels closely approximate the true edges of the original image. In Matlab, either the sensitivity thresholds for the Canny method be be specified or the default threshold can be chosen. Figure 5 shows the effects of using different thresholds of Canny method. The first row shows a 17 years old participant with no wrinkles on face, while the second row shows a 22 years old participant with some noticeable wrinkles. It can be seen that the Canny with threshold detection performed better than Canny with default threshold which extracted excessive amounts of detail. The Otsu algorithm decides the best threshold value to detect the major edges. In Experiment 1, threshold detection by using centre point region (c_i) was compared with that from the whole region (b_i). Figure 6 shows the different areas (highlighted) used for threshold detection.

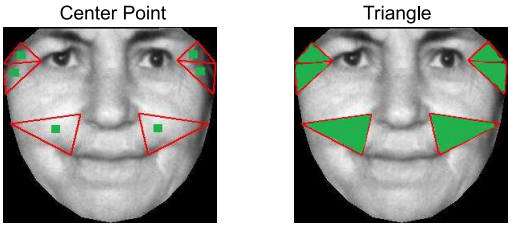

Figure 6: Center point versus triangle region.

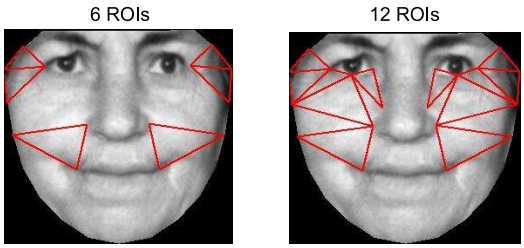

Figure 7: Different wrinkle region amount.

Table 2: Experiment 1, comparison of threshold type.

Threshold Type	Group A	Group B	Accuracy
Centre Point	16	16	80.00%
Triangle	13	14	67.50%

Table 2 shows the comparative results for centre point and triangle region threshold determination. The classification accuracy for triangle threshold is 67.50% and is improved by centre point threshold to an accuracy of 80%. It seems that triangle threshold is less consistent due to noise around the triangle corners which diverts the Otsu algorithm from finding appropriate threshold. From the results, a centre point threshold which focuses on the wrinkle area gives a better result than using the whole region.

In the Experiment 2, effects of ROI size were considered. In general, facial wrinkles are created by repeated facial muscular movements and expressions and are therefore formed increasingly as the person gets older. It was assumed that the amount of wrinkles is a cue that strengthens the feature for discriminating the age groups. The experiment settings are the same as those for Experiment 1 but the number of ROI was increased from 6 to 12, which included the eye bag area. Figure 7 presents the different ROI used in this test.

Table 3: Experiment 2, variation in numbers of ROIs.

Threshold Type	Group A	Group B	Accuracy
6 ROIs	16	16	80.00%
12 ROIs	16	17	82.50%

Table 3 shows the experimental results of using different amount of wrinkle region. The age group classification is affected by the number of wrinkle re-

gion. For the 6 ROIs, the classification accuracy is 80.00% while it is improved by the 12 ROIs with accuracy 82.50%.

Table 4: Experiment 3, detailed accuracy by class.

Description	Group A	Group B
TP Rate	0.678	0.551
FP Rate	0.449	0.322
Precision	0.793	0.403
Recall	0.678	0.551
F-measure	0.731	0.466

In the Experiment 3, age group classification was performed on the whole FG-NET dataset. The selected 40 images from Experiment 1 were used as a training set and the remaining images of FG-NET as the test set. Other settings were same as those in Experiment 1. Table 4 illustrates the detailed accuracy by class. In overall, the F-measure of group A was 0.731 and group B was 0.466. This implies that 618 of 962 samples were correctly classified. It was expected that the classification performance might drop significantly due to the very poor quality and imbalanced nature of the FG-NET dataset images. However, the proposed features provide significant ageing characteristics which can be embedded with global features for further enhancement.

In overall, although the accuracy is reasonable, there are a number of issues to be investigated. For the experimental set-up, the age groups were divided into two groups, one is between age 0 and 20 and other is above 20. This arbitrary choice of cut-off point between age groups makes for a highly unbalanced dataset. It would be interesting if the set-up could be narrowed into smaller age interval, say 10 years old. Second, it has been noticed that the manual selection of higher quality images for the training set makes it unrepresentative of the data as a whole. Statistical test would be a better way to show that this dataset is uniform.

6 CONCLUSIONS

In this paper, we proposed a local wrinkle-based extractor namely LOWEX for age estimation. Features produced by LOWEX are wrinkle length and amount. Then, back-propagation neural networks are designed to classify the age groups based on the feature descriptor. The experimental results showed that the proposed method produced effective feature descriptor in age group classification with high accuracy. We conclude that wrinkles can produce prominent ageing descriptor.

In future works, it is planned to extend the investigation to other datasets with high resolution images such as BERC ageing database (Choi et al., 2011). In addition, the number of age groups will be expanded in order to access the discriminative ability of the proposed method. We claimed that the LOWEX is capable of deriving wrinkle pattern from facial images. We also found that local features are useful for age group categorizing instead of detailed age estimation due to different people aged in different ways. In addition, we are also interested to investigate the skin texture changes when human getting older.

ACKNOWLEDGEMENTS

This work was supported by Manchester Metropolitan University's PhD Studentship.

REFERENCES

Albert, A., Ricanek, K., and Patterson, E. (2007). A review of the literature on the aging adult skull and face: Implications for forensic science research and applications. *Forensic Science International*, 172(1):1–9.

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (6):679–698.

Chao, W.-L., Liu, J.-Z., and Ding, J.-J. (2013). Facial age estimation based on label-sensitive learning and age-oriented regression. *Pattern Recognition*, 46(3):628 – 641.

Choi, S. E., Lee, Y. J., Lee, S. J., Park, K. R., and Kim, J. (2011). Age estimation using a hierarchical classifier based on global and local facial features. *Pattern Recognition*, 44(6):1262–1281.

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

FG-NET aging database (2000). http://www-prima.inrialpes.fr/FGnet/. accessed on September 2012.

Fu, Y., Guo, G., and Huang, T. (2010). Age synthesis and estimation via faces: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(11):1955–1976.

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

Goodall, C. (1991). Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 285–339.

Khavkin, J. and Ellis, D. (2011). Aging skin: histology, physiology, and pathology. *Facial plastic surgery clinics of North America*, 19(2):229.

Kwon, Y. and da Vitoria Lobo, N. (1999). Age classification from facial images. *Computer Vision and Image Understanding*, 74(1):1–21.

Lanitis, A., Draganova, C., and Christodoulou, C. (2004). Comparing different classifiers for automatic age estimation. *IEEE Transactions on Systems, Man, and Cybernetics*, 34(1):621–628.

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (2002). Learning representations by back-propagating errors. *Cognitive modeling*, 1:213.

Yan, S., Wang, H., Fu, Y., Yan, J., Tang, X., and Huang, T. (2009). Synchronized submanifold embedding for person-independent pose estimation and beyond. *IEEE Transactions on Image Processing*, 18(1):202–210.

Zhang, Y. and Yeung, D.-Y. (2010). Multi-task warped gaussian process for personalized age estimation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2622–2629. IEEE.