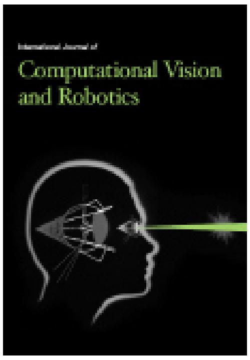


International Journal of Computational Vision and Robotics

ISSN online: 1752-914X - ISSN print: 1752-9131
https://www.inderscience.com/ijcvr

An investigation into automated age estimation using sclera images: a novel modality

Sumanta Das, Ishita De Ghosh, Abir Chattopadhyay

DOI: 10.1504/IJCVR.2022.10049572

Article History:

Received:	19 January 2022
Accepted:	24 June 2022
Published online:	01 December 2023

An investigation into automated age estimation using sclera images: a novel modality

Sumanta Das*

University of Engineering and Management,
University Area, Plot No. III,
B/5, New Town Rd.,
Action Area III, Newtown,
Kolkata, West Bengal,
700160, India
Email: sumanta.research@gmail.com
*Corresponding author

Ishita De Ghosh

Barrackpore Rastraguru Surendranath College,
85, Middle Road, 6,
River Side Rd., Kolkata,
West Bengal, 700120, India
Email: ishitade.ghosh@gmail.com

Abir Chattopadhyay

University of Engineering and Management,
University Area, Plot No. III,
B/5, New Town Rd.,
Action Area III, Newtown,
Kolkata, West Bengal,
700160, India
Email: abir.uem@gmail.com

Abstract: Automated age estimation attracts attention due to its potential application in fields like customer relationship management, surveillance, and security. Ageing has a significant effect on human eye, particularly in the sclera region, but age estimation from sclera images is a less explored topic. This work presents a comprehensive investigation on automated human age estimation from sclera images. We employ light-weight deep learning models to identify the changes in the sclera colour and texture. Extensive experiments are conducted for three related tasks: estimation of exact-age of a subject, categorical classification of subjects in different age-groups, and binary classification of adult and minor subjects. Results demonstrate good performance of the proposed models against the state-of-the-art methods. We have obtained mean-absolute-error of 0.05 for the first task, accuracy of 0.92 for the second task, and accuracy of 0.89 for the third task.

Keywords: human age estimation; age-group classification; adult-minor binary classification; sclera images; deep learning; MASDUM; SBVPI.

Reference to this paper should be made as follows: Das, S., De Ghosh, I. and Chattopadhyay, A. (2024) 'An investigation into automated age estimation using sclera images: a novel modality', *Int. J. Computational Vision and Robotics*, Vol. 14, No. 1, pp.42–62.

Biographical notes: Sumanta Das has received his Master's in Computer Science from Barrackpore Rastraguru Surendranath College, West Bengal, India. He is presently a PhD Scholar at the University of Engineering and Management, Kolkata, India. His research interests include image retrieval, pattern recognition, artificial intelligence and their applications. His research topics includes image retrieval, auto detection of diabetic retinopathy, sclera biometrics, auto human age estimation, anti-spoofing strategies in biometrics and advanced deep learning models. He has published research papers in multiple conferences and international journals and contributed chapters in books published by reputed publishers. He is an active member of the ACM community.

Ishita De Ghosh is a senior member of IEEE and ACM from 2013. She received her MTech and PhD both in Computer Science from Indian Statistical Institute, Kolkata. Presently, she is an Associate Professor in the Department of Computer Science, Barrackpore Rastraguru Surendranath College, West Bengal, India. Her current research interests includes pattern recognition, artificial intelligence, and their applications. She has published research papers in several international journals, and contributed chapters in books published by reputed publishers.

Abir Chattopadhyay completed his Master's in Solid State Electronics from Jadavpur University in 1988. He holds a PhD in Low Temperature Electronics from Jadavpur University. He mainly concentrated on implementation of Hartee-Fock potential in place of max born potential. He had been associated with many reputed colleges and universities, research projects, CSIR fellowships and also worked as a junior scientific officer. Presently, he is working as the Dean Research in the University of Engineering and Management, Kolkata, India. His areas of research interests include computer communications, mobile computing, cognitive radio, sensor networks, distributed systems, performance evaluation and network security, semiconductor physics, and VLSI design. He is an author of many textbooks, research papers published in international conferences and journals. He serves as a review committee member of IEEE sponsored ICCCCM; co-convener and TPC Co-Chair of Springer sponsored OPTRONIX and is doing regular review work in the various international reputed journals.

This paper is a revised and expanded version of a paper entitled 'Deep age estimation using sclera images in multiple environment' presented at ICCET-2021: 6th International conference on Computing in Engineering & Technology-2021, Dr. Babasaheb Ambedkar Technological University, Lonere, India, 30–31 January 2021.

1 Introduction

Images of human body or its parts characterising the structural forms or behavioural traits are popular for *automated age estimation* research. It has wide applicability in fields like electronic customer relationship management (ECRM), security and surveillance, and biometrics (Angulu et al., 2018). Some other applications are detection of specific age groups for prioritisation or prohibition. Examples are identification of senior citizen age-group for prioritisation in certain situations, or detection of minor age-group for prohibition in other situations (like combating child sexual exploitation material or CSEM). Ageing has a significant effect on human eye, particularly on sclera, which is the white region surrounding the iris region. For most of the people, sclera colour slowly turns yellowish with age, due to deposition of fat globules between the collagen fibres (Watson and Young, 2004). Yellowness of the sclera is also associated with liver pathology. Though there are many recent work exploring sclera images as a biometric modality (Rot et al., 2020; Das et al., 2020), its potential as age estimation modality is less explored. Images of face, fingerprints, palm-veins, gait, peri-ocular region of eye, retina or iris are used commonly for automated age estimation. Now we discuss why sclera images are more advantageous. Before that, other modalities are discussed briefly.

There are several recent works on age estimation by face images (Jana et al., 2015; Liu et al., 2020; Hiba and Keller, 2021). One of the reasons for popularity of face image is that, it can be acquired easily in visible light without any physical contact. However, it has certain drawbacks, e.g., there are great variations in facial appearances for diverse ethnic groups and communities of human race (Angulu et al., 2018). Other factors that may lead to wrong perceptions are diet, skin infections, use of cosmetics, etc. Moreover, face age estimation requires good quality images covering the complete face area which may not be feasible all the time, e.g., when the subject uses a face-cover or a face-mask (Thorley et al., 2022; Zhang et al., 2021; Wang and Chen, 2021).

Fingerprint images are used effectively for age estimation (Falohun et al., 2016; Saxena and Chaurasiya, 2017; Galbally et al., 2019). However, it also has some limitations. Firstly, in this case, image acquisition requires touching the sensor since modern touchless devices are not commonly available. Nonetheless, physical contact should be avoided to minimise the chance of infection in several situations. Secondly, fingerprint sensors embedded in modern computing systems may help in combating CSEM to some extent, but data acquisition is done during login time only. Thus, continuous monitoring of the user is not possible. Finally, spoofing and degradation of fingerprint quality with age may affect performance of the system (Galbally et al., 2019). Palm vein images are also used for age estimation (Damak et al., 2020). In general, demerits of fingerprint images are also applicable to palm vein images.

Gait images are used successfully for age detection in some recent work (Zhou et al., 2020; Islam et al., 2021). These images can be acquired from a distance, but loose dresses, like body covered gowns, or burkhas, that hide body-structure, may affect the performance. However, a heat map may overcome the issue with slightly higher cost. It is also noted that the estimated age is a mere assumption for application like ECRM but not decisive with practical confidence (Islam et al., 2021).

Ageing has its effect on every organ of our body, and, eye is no exception (Cavallotti and Cerulli, 2008). Peri-ocular region of an eye image is a good indicator of age (Kamarajugadda and Polipalli, 2019). But the problem is that skin and eyebrow texture may change drastically within a short period of time due to use of cosmetics or plucking of hair. Geographical origin or ethnicity may also result in variation in appearance within the same age-group. Iris and retina image are also used for age detection (Rajput and Sable, 2020; Gerrits and et al., 2020). Both iris and retina image acquisition require specialised imaging devices under constrained environment, making acquisition a costly affair. Moreover, the device is to be placed very close to the eye, which can merely be called contactless.

With this background, we now discuss the motivation of using sclera image as an age estimation modality.

- Potential use of sclera images for age estimation is not available in literature except for an introductory work (Das et al., 2022a). Nevertheless, that work deals with exact-age estimation. Present work deals with three different aspects, namely, exact-age estimation, age-group classification and detection of adult and minor subjects. Moreover, this is a comprehensive study with extensive experiments.
- Partial occlusion of face does not restrict use of sclera image for age estimation. The entire face image is not required; rather a portion of the image containing sclera is enough. Thus, small angle view can be used for image acquisition, resulting in less memory as well as less bandwidth requirement.
- Sclera images can be acquired without physical contact. A normal digital camera fitted in a desktop, laptop, or smart phone can continuously capture the eye images for sclera age estimation. Hence, control and prevention of CSEM is possible with the help of sclera age estimation.
- There are many recent papers on biometric recognition (Jammoussi et al., 2016; Vitek et al., 2020a; Bokade and Kanphade, 2021; Das et al., 2021, 2022b). However, the pertinent question for any biometric modality remains relevant here also. Is the characteristic sufficiently invariant over a period? To answer this question, it is required to find out the effects of age and ageing on sclera images.

In the present work, the potential of sclera images for automated age estimation is explored using three different tasks, viz.,

a exact age estimation formulated as a regression problem

b age group estimation as a categorical classification problem

c adult-vs-minor as a binary classification problem to differentiate between adults and minors.

Lighter deep models similar to VGG-16 are proposed to accomplish the tasks. Extensive experiments are conducted to examine how the proposed models perform for images acquired under constrained or unconstrained environments using different imaging conditions and devices. Two sclera datasets, viz., *SBVPI* and *MASDUM* are used in our work. Results obtained by training and testing with each dataset are presented and analysed. Combined training and testing utilising both datasets are also analysed.

Contribution of the work is summarised below:

- *Exact age estimation by regression:* In order to estimate exact age of a subject in years, a light-weight variant of deep learning model VGG-16 is employed

(Simonyan and Zisserman, 2014). The best mean-absolute-error (MAE) obtained by the model is 0.05.

- *Age-group estimation by categorical classification:* Human beings as well as machines seem to predict the age-group of a person more accurately, than his or her exact age. A deep learning model is employed to perform the categorical classification task of estimating age-groups. More than 92% accuracy is achieved.
- *Detection of adult subjects by binary classification:* The problem of differentiating adult and minor (children) subjects is formulated as a binary classification problem. Two separate binary classification problems are solved to identify subjects less than 18 years and 21 years of age, since majority of the countries choose one of these two as the age of attaining adulthood. We achieve more than 87% and 89% accuracy for both these classification tasks respectively.
- *Validation of proposed models:* All the three deep models were trained and tested on two datasets built in completely different set-ups and environments. Results on individual datasets as well as that obtained by combined training are also presented. Good performance measures establish the applicability of sclera images as a potential age estimation modality.

The work is organised in the following sections, a brief review is given in Section 2, proposed methods are described in Section 3, experimental results are presented in Section 4, analysis of results is presented in Section 5, and concluding notes with future scope are given in Section 6.

2 Review: age estimation from images

A brief review on state-of-the-art work for automated estimation of human age using images captured in visible light is given now. Structural or behavioural characteristics of the human body or its parts, such as, face, fingerprint, gait, peri-ocular region of eye, iris, or, retina are used commonly for *automated age estimation*.

Age estimation from face images is being studied for more than two decades. Angulu et al. (2018) published a survey on the topic. It shows that feature extraction techniques employed commonly for the purpose include Gabor filters, linear discriminant analysis, local binary, ternary and directional patterns, grey-level co-occurrence matrix, spatially flexible patch, biologically inspired features, Grassmann manifold, deep learning classifiers, regression analysis and hybrid approaches. The best performance reported in the work is achieved by using a cascaded CNN with Gaussian error of 0.297.

In 2020, Liu et al. proposed a method using Mixed Attention-ShuffleNetV2 (MA-SFV2) network, achieving MAE of 3.81 for FG-NET dataset and 2.68 for MORPH-II dataset (Liu et al., 2020). In 2021, Hiba and Keller proposed hierarchical attention-based age estimation along with bias analysis involving a VGG-16 model achieving MAE of 1.13 using the same dataset (Hiba and Keller, 2021). Wrinkles around the eye region are seen to be useful in producing competitive results when compared with the results obtained by the entire face image, achieving MAE of 3.71 (Jana et al., 2015). Very recently in 2022, Yi published a work that emphasise more on eye than entire face, and uses deep models to achieve better prediction accuracy of 94% on 6 years error margin (Yi, 2022).

Table 1 Characteristics of age estimation techniques using common modalities

Modality ↓	*Source of light*	*Contactless*	*Probable occlusion*	*Reliability*	*Cost of installation*	*Performance*
Palm veins	Visible	No	Minimal	High	High	F1-score = 0.94 (Damak et al., 2020)
Iris	NIR	No	Lenses	High	High	Accuracy = 0.74 (Rajput and Sable, 2020)
Fingerprint	Visible	No	Minimal	High	High	Accuracy = 0.82 (Falohun et al., 2016)
Retina	NIR	No	Minimal	High	High	Accuracy = 0.89 (Gerrits and et al., 2020)
Gait	Visible and NIR	Yes	Dresses	Low	Medium	Accuracy = 0.90 (Zhou et al., 2020)
Face	Visible	Yes	Masks, clothing	Medium	Low	Accuracy = 0.91 (Yi, 2022)
Peri-ocular	Visible	Yes	Spectacle	Medium	Low	Accuracy = 0.96 (Kamarajugadda and Polipalli, 2019)
Sclera	Visible	Yes	Lenses	High	Low	Accuracy = 0.92, *(Proposed, 2022)*

Age estimation from fingerprint images is being studied during the last ten years. In 2013, Arulkumaran et al. proposed a technique using DWT and PCA, which shows an average success rate of 68% for an in-house dataset (Arulkumaran et al., 2013). Saxena et al. proposed a technique using curvelet features in 2015, which achieves the best standard error of 0.14 for three age groups within the range 0 to 18 years (Saxena et al., 2015). Later in 2017 they obtained an accuracy of 0.80 for children below 14 years (Saxena and Chaurasiya, 2017). In 2016, Falohun et al. reported an accuracy of 0.82 for an in-house dataset using DWT and PCA (Falohun et al., 2016). In 2019, Galbally et al. published a work which studies the effect of age and ageing on fingerprint biometrics (Galbally et al., 2019). *Palm vein images* are also used for age estimation. In 2020, Damak et al. proposed a method for classifying children, young people and adults using weighted K-nearest neighbour classifier with an F-measure of 94.2% using the VERA palm vein database (Damak et al., 2020).

Image of human gait or *walking style* is a popular modality used in ECRM and similar applications (Islam et al., 2021). Many datasets are available and accuracy rates are around 0.80 for several methods with a mean-absolute-error (MAE) of around 5 years. A recent ANN-based technique published in 2020 by Zhou et al. utilises an in-house dataset and achieves accuracy of 0.90 and AUC of 0.86 (Zhou et al., 2020). Gait images can be acquired from a distance, but loose dresses like body covered gowns or burkhas that hide the structure may affect performance. However, a heat map may overcome the issue with slightly higher cost. It is also noted that the estimated age is a mere assumption for application like ECRM, but not decisive with practical confidence.

Peri-ocular regions cropped from eye images are increasingly being used for recognition purpose. In 2019, Kamarajugadda and Polipalli proposed a technique in which feature extraction was done using SURF and classification was done by a combination of SVM and K-NN (Kamarajugadda and Polipalli, 2019). The method achieved an accuracy of 0.96 for an in-house dataset. *Iris images* are also used for age detection. A recent work published in 2020, reports accuracy of 0.74 and F1-score of 0.83, using near infrared eye images (Rajput and Sable, 2020). A small number of age detection work is available using *retina images*. In 2020, best reported age prediction accuracy is 0.89 with MAE of 0.0278 using deep learning techniques (Gerrits and et al., 2020).

Scleral changes due to age and/or disease have been studied extensively for clinical/medical purposes (Beattie et al., 2011; Coudrillier et al., 2012; Chakraborty et al., 2016). In 2015, Russell et al. studied effects of ageing on colour of iris and sclera region (Russell et al., 2014). Colours were analysed in $L^*a^*b^*$ colour-space and the attributes were estimated/evaluated by a number of human examiners. It was concluded that sclera colour is a good indicator of human age, health, and attractiveness. In 2021, an introductory work on *age estimation using sclera images* was published by Das et al. (2022a). Exact age estimation was explored with a regression technique employing a modification of VGG-16 model. It was trained and tested by SBVPI dataset and the experimental results show an MAE of 0.06.

Various modalities for age estimation are listed in Table 1. Emphasis is given on techniques that use images acquired in visible spectrum. For each modality, the table shows:

1 source of light

2 whether body contact is required or not during acquisition

3 probable occlusions

4 reliability

5 installation cost

6 the best performance obtained till date.

Though the dataset and evaluation metrics may vary for every modality, we think the last column is important for completeness of the table.

3 Proposed work

The primary step is to extract the region of interest, which is sclera, from eye images. Sclera segmentation has been widely explored in literature (Vitek et al., 2020a; Das et al., 2020). We utilise sclera mark-up images available in the datasets for this step. Thus for each RGB image, the corresponding binary mark-up image is used to guide segmentation of sclera portion. Example of an eye image, its sclera mark-up image, sclera segmented image, and, cropped sclera patch is shown in Figure 1. Sample sclera patches along with age of subjects are shown in Figure 2.

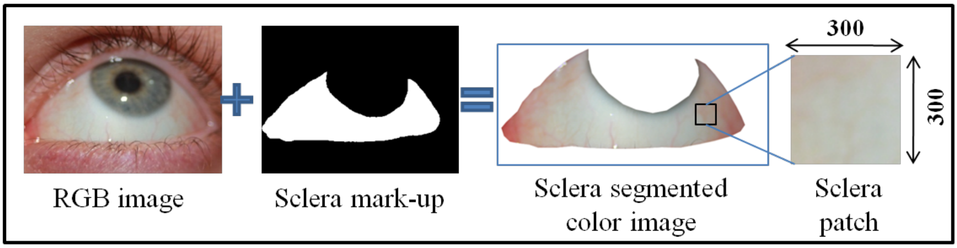
Three different tasks for age estimation using sclera images, viz., estimation of exact-age of subjects, categorical classification of subjects among various age-groups, and binary classification of children and adult subjects. We experimented on standard VGG-16 or UNet and SegNet deep models, but due to over-fitting, we had to reduce model complexity for proper convergence. This indicates low complexity of the given problem. They are elaborated one by one now.

3.1 Exact-age estimation

Given the sclera images, objective is to estimate the exact-age of the subjects in years. A preliminary work based on regression analysis using unbalanced training (unequal number of images from every age) was reported in Das et al. (2022a). Balanced training requires equal or nearly equal number of images from each age, so that the prediction is not biased towards a single age.

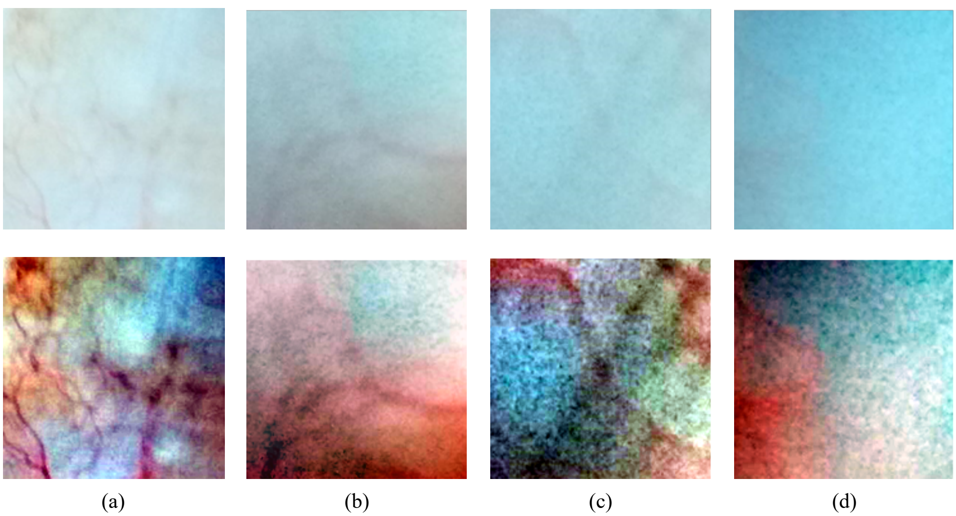
In this work, we propose a regression analysis for *exact age estimation* using a light-weight deep learning model based on VGG-16 (Simonyan and Zisserman, 2014). The proposed model with layer parameters is given in Figure 3. The model includes four convolution-layers, three max-pooling layers and three fully connected dense layers, producing a single output. The output is a floating-point number within the range [0.01–0.99], both inclusive. Estimated age of the subject is obtained on multiplying the output by 100. The model is trained by small square-shaped coloured patches chosen randomly from sclera segmented RGB eye images. All training patches are augmented by random contrast variation of up to 20%. The model is given a balanced training with equal number of patches extracted from images belonging to subjects of all age groups. Absolute error is the absolute difference between the actual age and the predicted age. Mean absolute error (MAE) is calculated as the average of absolute errors obtained for all the subjects.

Figure 1 Sclera segmented colour image obtained from the original RGB image (see online version for colours)



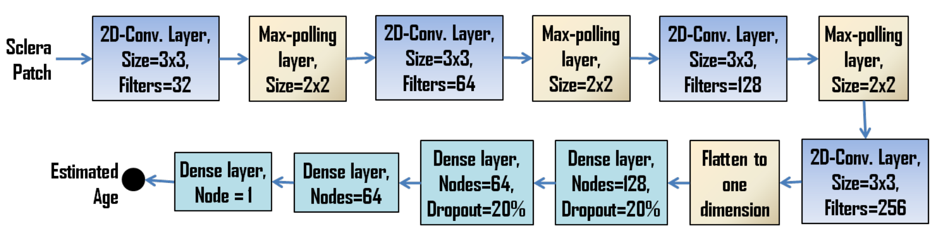
Note: Example of a sclera patch cropped from the segmented image is also shown.

Figure 2 Sample sclera patches are along with the age of the subjects, (a) 60 (b) 40 (c) 20 (d) 5 (see online version for colours)



Note: Original patch images are given in the first row. Enhanced patch images are given in the second row.

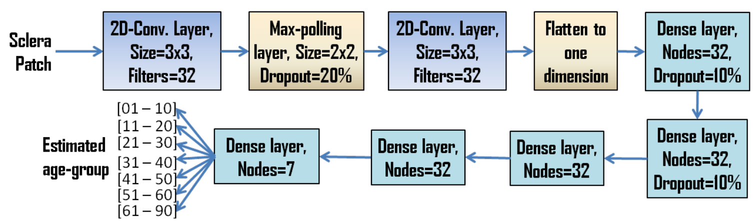
Figure 3 Proposed model for *exact-age estimation* (see online version for colours)



3.2 Age-group estimation

Categorical classification of subjects in different age-groups is an important task often found in literature. However, for sclera modality, this approach is not used before. Therefore, this is a pioneering work on age-group estimation using sclera images. For our work, the entire dataset is partitioned into *seven disjoint sets* based on seven age-groups, viz., [1–10], [11–20], [21–30], [31–40], [41–50], [51–60], [61–90]. For training, a set of 100 sclera patches from each age-group are selected randomly, and for testing another random set of 100 patches from each group are selected. Thus, a sum of 700 patches is used for training and another 700 patches for testing. These patches are fed one-by-one into a deep model for categorical age estimation. In this model, two convolution layers with a max-pooling layer in between are followed by fully connected dense layers to produce seven categories at output for seven age-groups. The output is represented by $7 \times N$ Boolean matrix, where N represents the number of samples. High value in a matrix-row indicates the age-group of the subject. The model and layer parameters are given in Figure 4.

Figure 4 Proposed model for *age-group classification* (see online version for colours)



3.3 Classification of adults and minors

The need for binary classification of adults and minors arises from several application fields like:

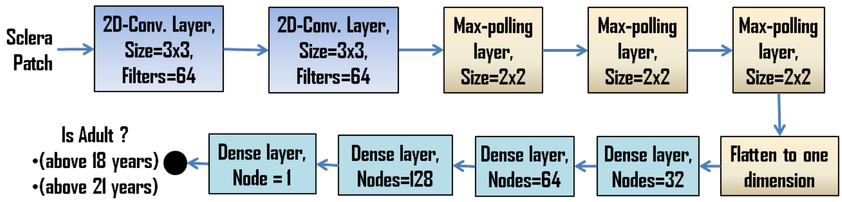
1 access control to places like malls, airports, public clubs, liquor shops

2 CSEM that aims at restriction of certain online content from under-age users of the system

3 visual surveillance and monitoring.

However, for sclera modality, this approach is new. Therefore, this is a pioneering work on binary classification of adults and minors using sclera images. A simple model is built for the work. The model with layer details is shown in Figure 5. Two convolution layers followed by three max-pooling layers are used to reduce the layer size. Few fully connected dense layers are used to classify the samples. The model is trained and tested separately for identifying subjects below 18 and 21 years of age, since these two are the most frequently used ages of attaining adulthood. In some countries like India, people of age 18 years and above are considered as adults, whereas in some other countries, the adult age starts from 21 years (Wikipedia, n.d.). In testing phase, ten random sclera patches from each sclera image are used to predict the class. Median value of ten outputs is taken as the final output for the image. Since the goal here is binary classification, accuracy, recall and precision values are used to evaluate the performance.

Figure 5 Proposed model to *filter-adult* subjects (see online version for colours)



4 Experimental results

Experiments are performed using a system with Intel-i7 9th gen. processor, and NVIDIA GPU GTX1660Ti with CUDA capability 7.5; 12-GB RAM and, 512-GB SSD. The program along with the dataset is stored in SSD. Tensor-flow 2.3 along with Python 3.6 is used for all the experiments. How the proposed models perform on different datasets, especially on datasets built in separate environments, with different kinds of acquisition devices? That was a question of interest, so we experimented with MASDUM and SBVPI datasets. *Sclera blood vessels, peri-ocular and iris (SBVPI)* dataset is a publicly available sclera dataset with age information. It contains high quality images acquired with a standard camera in restricted environment, whereas *Multi angle sclera dataset in unrestricted mobile-environment (MASDUM)* contains images acquired in unrestricted settings with two smart phone cameras. The dataset is publicly available at https://sites.google.com/view/sclerabiometrics/home to facilitate further research. Capturing devices and other specifications of the two datasets are given in Table 2.

Table 2 Main characteristics of datasets

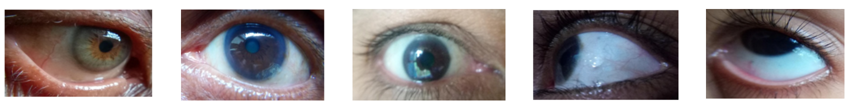
Characteristics	*MASDUM*	*SBVPI*
Type of camera	Mobile + macro lens	Standard + macro lens
Device name	Yu 5530 and Samsung G-J4	Canon EOS 60D
Environment	Unconstrained	Constrained
Lighting	Indoor, outdoor	Indoor
No. of images, (size)	1,055, (700 × 1,500)	1,856, (1,700 × 3,000)
No. of persons, eyes	48, 96	55, 110
Range of age	4 to 68	15 to 80
Gender	19 male, 29 female	29 male, 26 female
Geographical origin	Bengalis from West Bengal, India	Caucasian

Results of experiments done separately on each dataset are provided. Results of experiments conducted in a combined fashion on the two datasets together are also provided. In addition to that, detailed analyses of results are provided. MASDUM dataset has been built by us and the details of building procedure are given in the next subsection.

4.1 MASDUM dataset

Eye images for four gaze directions (left, right, front and up) were acquired in *indoor* as well as *outdoor* lighting conditions. Reflections and noise are present in the images, as happens in a real-life scenario under unconstrained environment. Images were captured by two separate mobile handsets, YUNICORN 5530 and SAMSUNG GALAXY J4. Mobile handset cameras were mounted with macro lenses for better focus on sclera area. Some of the recently launched mobile handsets come with such macro lens embedded in them. The capturing device was at a distance of 5 to 15 cm from the eye. Images were captured from 48 individuals, resulting in 96 eyes. The dataset contains 1,055 images in total. Sample images from the dataset are shown in Figure 6.

Figure 6 Sample images from MASDUM dataset, (a) left, 74 (b) front, 66 (c) front, 50 (d) right, 31 (e) up, 5 (see online version for colours)



(a) (b) (c) (d) (e)

Note: Gaze direction and age of the subject are indicated below each image.

Randomly cropped sclera patches were used for training and testing the deep models. Experiments suggest that larger patch area yields better results. We fixed the patch area at 300×300 pixels to comply with hardware limitations. A patch containing at least 90% of sclera area is considered as an input to the proposed models. Results of the proposed methods are given now.

4.2 Exact-age estimation results for MASDUM

Results of exact (chronological) age estimation method proposed in Subsection 3.1 are presented now. Two sclera patches are randomly selected from each image of the dataset to increase the number of training samples. The validation set contains 30% of patches and rest is used for training. *Adam* optimiser is used with *learning rate* of 0.00001. *Binary cross-entropy* loss function is used with batch size 16. *Relu* activation function is used at each layer with kernel initialiser *glorot-uniform*. The output layer uses *sigmoid* activation function. Data augmentation is done by randomising image contrast up to 20%. *Mean absolute error (MAE)* is used to evaluate the results. It is obtained by calculating the absolute difference of the actual age and the predicted age of a subject.

In a previous work, an MAE of 0.10 was reported for exact-age estimation on MASDUM dataset (Das et al., 2022a). In this dataset, elderly subjects are less in number, as there are only 6 subjects with age more than 60 years. Poor representation of elderly population and unbalanced training may result in increase of overall MAE, and, thereby decrease the performance. Removing the six elderly subjects resulted in reduction of *MAE to 0.06* (termed as *MAE-R* in the cited work).

In this work, we try to overcome the problem of unbiased training by selecting 100 sclera patches from each of the age groups: [1–10], [11–20], [21–30], [31–40], [41–50], [51–70]. An improved *MAE of 0.08* and *MAE-R of 0.06* is obtained for all subjects, which is obtained after removal of 6 elderly subjects during experimentation. We note a greater number of subjects in the age group [21–30], and the MAE obtained for the group is in the range of [0.01-0.05]. Higher range of MAE, for example, [0.31–0.40] is obtained for the age-group [51–70]. Hence, it is concluded that the model is still biased towards young subjects. Comparative study of the results with a previous work is presented in Table 3. Figure 9 illustrates the variation of number of images with respect to MAE in specific age-groups. We found no significant effect due to gender of subjects in our results.

4.3 Age-group classification results for MASDUM

Prediction of age-group seems to be an easier job than predicting the exact age. As before, sclera patches are randomly selected for input. Data is augmented by random contrast variation of up to 20%. Each layer uses the *Relu* activation function except the last layer which uses the *softmax* activation function. Kernel is initialised using *he-normal*. Optimiser used is *stochastic gradient decent*. Loss function is *categorical-cross-entropy*. Batch size used is 64. *Accuracy* is calculated as the number of correctly classified predictions to the total number of predictions. Average accuracy obtained for the two datasets are trained and tested separately, and retrained in combination, as given in Table 4. Accuracy obtained for each age-group along with its graphical representation is given in Figure 10(b). For MASDUM dataset, on an average, the accuracy of age-group classification is 0.920.

4.4 Adult-minor binary classification results for MASDUM

For this work, 1,000 sclera patches are selected randomly from both classes (minor and adult) within 70% of subjects. Another 1,000 sclera patches are used for validation testing from the remaining 30% of the dataset. Data is augmented by random contrast variation of 10%. Activation function *Relu* is used with kernel initialiser *he-normal*. The output layer uses *sigmoid activation function*. Optimiser used is stochastic gradient decent with loss function *binary cross entropy* monitoring validation-loss. Batch size used is 64. *F1-score* is used as the measure of evaluation, which is the harmonic mean of *recall* and *precision*.

Average accuracy metrics obtained for the two datasets are trained and tested separately, and in combination are given in Table 5 for 18 years and 21 years respectively. Graphical plots showing percentage of minor and adult images obtained for actual and predicted samples are given in Figure 11(b). For MASDUM dataset, on an average, the accuracy of identifying adult-subjects for 18 years and 21 years are 0.891 and 0.872 respectively.

4.5 Experiments with SBVPI dataset

Extensive experiments are performed with SBVPI dataset alone, and in combination with MASDUM dataset. Sclera segmentation, model training and testing on SBVPI dataset are performed as proposed in previous sections, experimental results are presented in this section, and the analysis of results are presented in the next section.

Table 3 Results of exact-age prediction using SBVPI and MASDUM

Training	*Testing*	*MAE*	*MAE-R*
MASDUM + SBVPI	SBVPI	0.104	0.081
MASDUM + SBVPI	MASDUM	0.093	0.079
SBVPI	SBVPI	0.107	0.080
MASDUM	MASDUM	0.080	*0.052*

Table 4 Results of age-group classification using SBVPI and MASDUM

Training	*Testing*	*Accuracy*
SBVPI	SBVPI	0.838
MASDUM + SBVPI	MASDUM + SBVPI	0.873
MASDUM	MASDUM	*0.920*

Table 5 Results of classifying adults above 18 years and 21 years using SBVPI and MASDUM

Training	*Testing*	*Accuracy*	
		18 years	*21 years*
SBVPI	SBVPI	0.500	0.487
MASDUM + SBVPI	SBVPI	0.686	0.741
MASDUM + SBVPI	MASDUM	0.822	0.800
MASDUM	MASDUM	*0.891*	*0.872*

Results for exact-age estimation process are presented in Table 3. MAE is the mean-absolute-error and MAE-R is mean-absolute-error obtained after removing subjects above 60 years of age, which is age group with the smallest number of subjects in the dataset. Age-group classification results are tabulated in Table 4. Age-group 1–10 contains zero subjects. Adult-minor binary classification results are presented in Table 5 for 18 years and 21 years age margin. It is seen that training and testing with MASDUM dataset alone give better results for all the three tasks. Probable causes are discussed in the next section.

Table 6 gives an idea about the time requirements.

Table 6 Time requirement of the proposed deep models

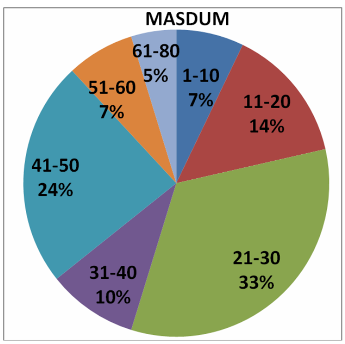
Model	*#FLOPS*	*#Param.*	*Exec.time*
Exact-age estimation	22 m	10 m	13 ms
Age-group class	46 m	23 m	25 ms
Adult subject identification	5.7 m	2.8 m	67 ms

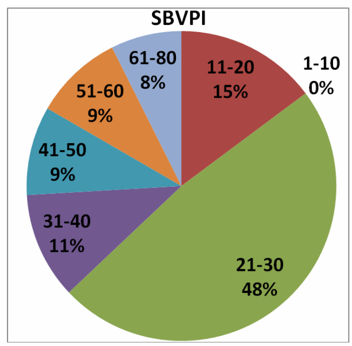
The time required by each model is suggestive of their light-weight nature and cost effective deployment. ‘FLOPS in millions’ is the number of floating point operations in a single forward pass, ‘#Param’ is the number of trainable parameters in millions, and ‘exec. time’ is the execution time in milliseconds required for prediction in a batch. It is observed that increase in the number of fully connected dense layers increase the execution time.

5 Discussion

Composition of the datasets with respect to the origin of subjects is important in this kind of work. All the subjects in MASDUM dataset are of Bengali origin from West Bengal, India. Bengalis are ethno linguistic group of people with diverse ethnic origin (Chattopadhyay, 2004). All the subjects in SBVPI database are of Caucasian origin (Vitek et al., 2020b). Sclera colour may vary due to this reason. Secondly, we find that the age distribution of subjects plays a major role because it is essential for proper training of models. Distribution of subjects in various age-groups for the two datasets is shown in Figure 7.

Figure 7 Percentage of subjects distributed in different age-groups in SBVPI and MASDUM datasets (see online version for colours)





The age span of MASDUM dataset is 4 years to 68 years, whereas the same for SBVPI dataset is 15 years to 80 years. Both the datasets show higher number of subjects in the age group [21–30], and less number of subjects in the elderly group [61–80]. SBVPI contains no subject in the age-group [1–10]; whereas MASDUM contains 7% of all the subjects in that age-group. SBVPI contains 19 subjects in the age-group [15–22]. Exact-age wise frequency distribution of the subjects is given now as *age (frequency)*: 15 (1), 16 (1), 18 (1), 19 (5), 21 (2), 22 (9). Since most of the subjects in the age-group are in the borderline of 18 and 21 years, the model for adult-minor classification may get confused during training. This may affect the accuracy of results. Both MASDUM and SBVPI datasets are prepared with unbalanced number of subject-ages. Further scope is to obtain a larger dataset with equal or near-equal representation of subjects from all ages. A brief discussion on experimental results is given now.

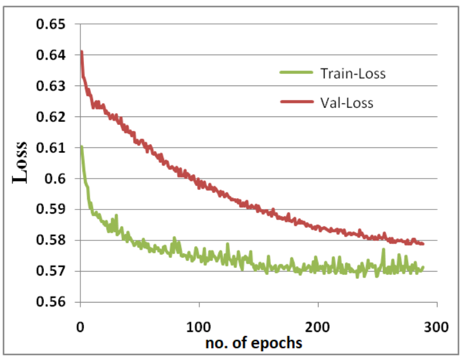
5.1 Exact-age estimation

It aims to estimate the exact age of the subject in years. We obtain converging graphs during training as shown in Figures 8(b) and 8(a).

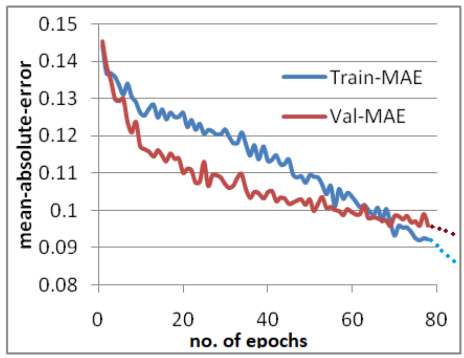
Both validation-loss and MAE keep decreasing with number of epochs. This confirms use of sclera modality for age estimation. At higher epochs, the convergence fails suggesting a model with better design and optimal parameters. Further optimisation of the model remains a future work. With respect to age, subjects are better distributed (balanced) in MASDUM than in SBVPI, leading to better results for MASDUM, refer Figure 9.

The figure shows how less the mean-absolute-error (MAE) is, for the middle aged population. Big red and green colour bars are seen for age group [21–30] indicating low MAE, whereas orange colour bars appear in [51–80], indicating larger error possibility. This relates to large number of subjects in datasets within [21–30].

Figure 8 Graph depicts the training convergence and results for the MASDUM dataset, (a) convergence (loss) during training for proposed *exact-age* estimation (b) convergence (MAE) during training for proposed *exact-age* estimation (see online version for colours)

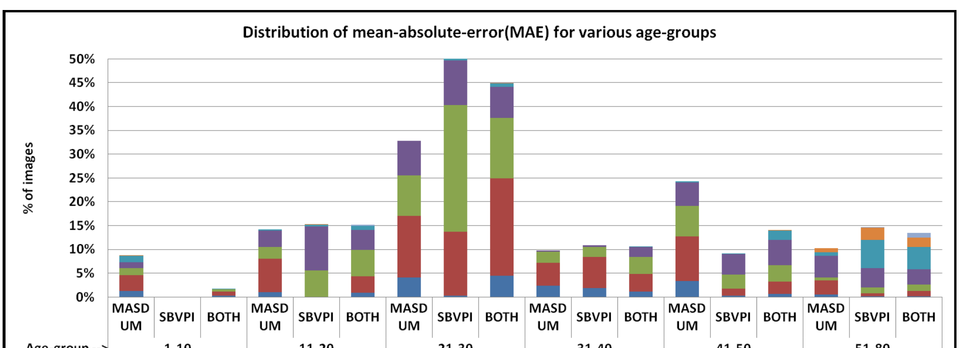


(a)



(b)

Figure 9 Comparison of experimental results obtained for exact-age estimation (see online version for colours)



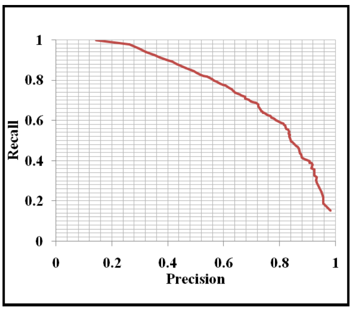
MAE / Age-group -->	1-10 MASDUM	1-10 SBVPI	1-10 BOTH	11-20 MASDUM	11-20 SBVPI	11-20 BOTH	21-30 MASDUM	21-30 SBVPI	21-30 BOTH	31-40 MASDUM	31-40 SBVPI	31-40 BOTH	41-50 MASDUM	41-50 SBVPI	41-50 BOTH	51-80 MASDUM	51-80 SBVPI	51-80 BOTH
0.41-0.50	0.00%	0.00%	0.00%	0.00%	0.00%	0.15%	0.00%	0.00%	0.04%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.22%	0.90%
0.31-0.40	0.12%	0.00%	0.00%	0.00%	0.11%	0.11%	0.00%	0.00%	0.19%	0.00%	0.00%	0.00%	0.00%	0.00%	0.04%	0.83%	2.48%	2.02%
0.21-0.30	1.43%	0.00%	0.11%	0.24%	0.38%	0.82%	0.00%	0.38%	0.75%	0.00%	0.00%	0.07%	0.24%	0.05%	1.98%	0.71%	5.99%	4.71%
0.11-0.20	1.19%	0.00%	0.00%	3.45%	9.28%	4.19%	7.26%	9.28%	6.54%	0.24%	0.38%	2.13%	4.88%	4.26%	5.31%	4.64%	4.10%	3.18%
0.06-0.10	1.43%	0.00%	0.49%	2.50%	5.56%	5.49%	8.45%	26.66%	12.67%	2.38%	2.10%	3.51%	6.43%	3.02%	3.40%	0.60%	1.19%	1.31%
0.01-0.05	3.33%	0.00%	0.93%	6.90%	0.00%	3.51%	12.98%	13.38%	20.37%	4.76%	6.53%	3.70%	9.40%	1.46%	2.65%	2.98%	0.59%	1.16%
0.00-0.009	1.31%	0.00%	0.26%	1.07%	0.00%	0.86%	4.05%	0.32%	4.52%	2.38%	1.83%	1.16%	3.33%	0.27%	0.64%	0.48%	0.16%	0.15%

Notes: 1 Trained and tested separately on MASDUM and SBVPI.
2 Combined experiment with both datasets.

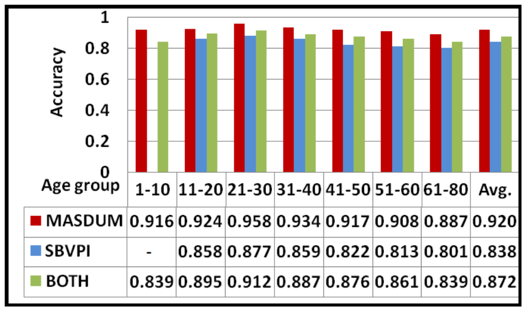
5.2 Age-group estimation

It performs better than exact-age estimation. Recall-precision graph obtained on experiments follows a regular pattern as shown in Figure 10(a). We have used a simple model for the novel work, which suggests sclera images for age estimation; however, there is scope for further improvement.

Figure 10 (a) Recall vs. precision graph for *age-groups classification* with the MASDUM dataset (b) Accuracy plot obtained for three experiments (see online version for colours)



(a)



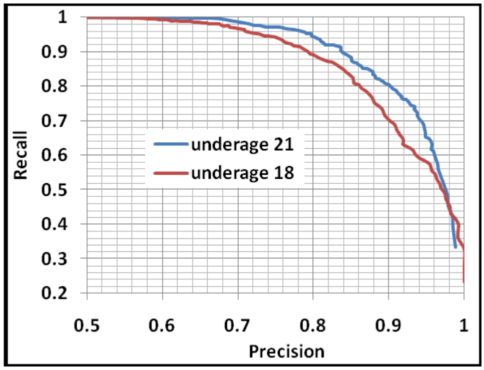
(b)

5.3 Adult-minor classification

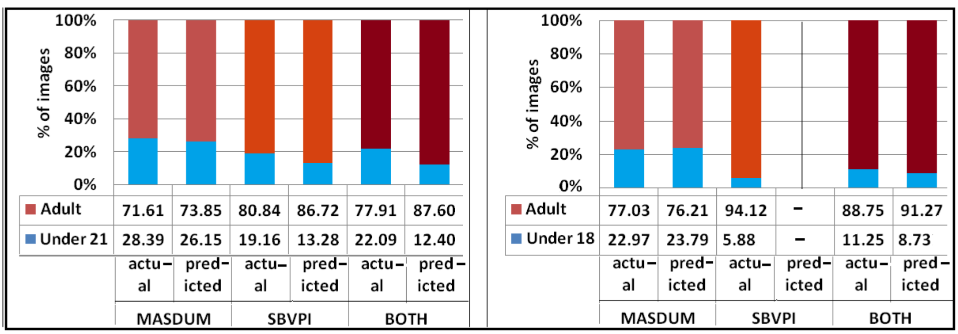
The model accomplishes binary classification for two separate experiments, one for the age of 18 and other for 21. The recall-precision graph is plotted as shown in Figure 11(a). Age-group estimation accuracy obtained by us [shown in Table 4 and Figure 10(b)] is justified with the distribution of subjects in the datasets discussed just now. Blank cell indicates no subject in the age-group. MASDUM has better results. Increase in size of sclera patches enhances results. The deep model needs improvement with balanced training. Estimation for 21 years shows better performance than that for 18 years. This suggests that the trained model is better trained for age-threshold of 21 years. While collecting data, age information was supplied as a self declaration without proof, so, more precise age data from a larger number of subjects may improve the results. Figure 11(b) shows how the number of images in each class (minor and adults) varies when predicted, in comparison to the actual numbers.

The first two bars show that the borderline of red-blue is positioned almost similar which indicates low deviation of predicted and actual numbers in MASDUM. The next two bars show large variation in red-blue borderline indicating larger deviation in predicted and actual numbers for SBVPI. BOTH datasets combined in experiments shows combined effect from both the datasets. SBVPI contains only 5% images below 18 years, which is too low for training the model and hence results are omitted in Figure 11(b). The percentage of actual and predicted correct samples for MASDUM is similar which indicates MASDUM dataset with more stable performance. Deep model could not be trained due to low number of minor subjects in SBVPI. MASDUM performs better with least variation in actual and predicted images. Table 7 compares with existing results from literature. However, adult-minor classification and age-group classification results are novel in this field and hence could not be compared.

Figure 11 (a) Recall vs. precision graph for proposed *adult-minor* classification using MASDUM dataset (b) Variation of the percentage of correctly classified images with respect to actual images in each class for three experiments (see online version for colours)



(a)



(b)

Table 7 Comparison with state-of-the-art results

Method	*Dataset*	*Results*		*Reference*
		MAE	*MAE-R*	
Exact-age estimation	SBVPI	0.117	0.081	Das et al. (2022a)
		0.107	*0.080*	*Proposed*
	MASDUM	0.106	0.067	Das et al. (2022a)
		0.080	*0.052*	*Proposed*

6 Conclusions

The current work is an investigation into automated age estimation by using sclera images. Experimental results suggest in favour of its feasibility in both regression and classification approaches. Potential applications are numerous and easily implementable with low cost mobile handsets. It is advantageous in many practical situations like faces covered by masks because only the eye region is required as input. Wearing normal contact lenses does not affect the colour or texture of sclera and hence it is hard to spoof. Only possibility of sclera occlusion arises from wearing spectacles with dark coloured glasses, or special scleral lenses. In ECRM, consumer data is collected and analysed for market research of the company as well as for enhancing the shopping experience of the consumer (Scullin and Lloyd, 2002). Consumer age-group estimation from sclera images can serve people better by reducing their shopping time. Combating CSEM is another topic of major concern, which can be counteracted by incorporating continuous monitoring of human eye by smart phones, desktops or laptops and sclera age-estimation systems. It has the potential to respond faster due to smaller input size (as eye is a small part of a face, its area is much smaller than face). People with eye disease or disease that affect the sclera appearance are not considered in this study. We did not find any significant impact on sclera due to gender. Larger number of subjects of various age-groups from multiple geographical origins, acquired using multiple devices under variable lighting conditions are essential for further study.

Acknowledgements

We are thankful to Dr. Matej Vitek of University of Ljubljana and his team members for allowing us to use the SBVPI dataset for our research work. We acknowledge the sincere assistance of Puspita Das (Sau) during the preparation of MASDUM dataset. Finally, we thank all the people who volunteered to be subjects during preparation of the dataset.

References

Angulu, R., Tapamo, J.R. and Adewumi, A.O. (2018) 'Age estimation via face images: a survey', *EURASIP Journal on Image and Video Processing*, Vol. 42 [online] https://doi.org/10.1186/s13640-018-0278-6.

Arulkumaran, T., Sankaranarayanan, P.E. and Sundari, G. (2013) 'Fingerprint based age estimation using 2d discrete wavelet transforms and principal component analysis', *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, Vol. 2, No. 3, pp.1060–1066.

Beattie, J.R., Pawlak, A.M., McGarvey, J.J. and Stitt, A.W. (2011) 'Sclera as a surrogate marker for determining age-modifications in Bruch's membrane using a Raman spectroscopy-based index of aging', *Investigative Ophthalmology; Visual Science*, Vol. 52, No. 3, pp.1593–1598.

Bokade, G.U. and Kanphade, R.D. (2021) 'A novel approach for secured multimodal biometric authentication based on data fusion technique', *International Journal of Computational Vision and Robotics*, Vol. 11, No. 2, pp.214–243.

Cavallotti, C. and Cerulli, L. (2008) *Age-Related Changes of the Human Eye*, Humana Press, Spain.

Chakraborty, N., Wang, M., Solocinski, J., Kim, W. and Argento, A. (2016) 'Imaging of scleral collagen deformation using combined confocal raman microspectroscopy and polarized light microscopy techniques', *PLOS ONE*, 2 November, Vol. 11, No. 11, p.e0165520, PMID: 27806070, PMCID: PMC5091908, DOI: 10.1371/journal.pone.0165520.

Chattopadhyay, A. (2004) 'Sectional president's address: Diverse ethno-cultural trends in ancient bengal: a study of processes of acculturation', *Proceedings of the Indian History Congress*, Vol. 65, pp.32–67.

Coudrillier, B., Tian, J., Alexander, S., Myers, K.M., Quigley, H.A. and Nguyen, T.D. (2012) 'Biomechanics of the human posterior sclera: age and glaucoma-related changes measured using inflation testing', *Investigative Ophthalmology & Visual Science*, Vol. 53, No. 4, pp.1714–1728.

Damak, W., Trabelsi, R.B., Masmoudi, A.D. and Sellami, D. (2020) 'Palm vein age and gender estimation using center symmetric-local binary pattern', *CISIS 2019/ICEUTE 2019, AISC*, pp.114–123.

Das, S., Ghosh, I.D. and Chattopadhyay, A. (2020) 'An efficient deep learning strategy: its aplication in sclera segmentation', *2020 IEEE Applied Signal Processing Conference (ASPCON)*, Kolkata, pp.232–236.

Das, S., Ghosh, I.D. and Chattopadhyay, A. (2021) 'An efficient deep sclera recognition framework with novel sclera segmentation, vessel extraction and gaze detection', *Signal Processing: Image Communication*, September, Vol. 97, Article No. 116349, ISSN: 0923-5965.

Das, S., Ghosh, I.D. and Chattopadhyay, A. (2022a) 'Deep age estimation using sclera images in multiple environment', in Iyer, B., Ghosh, D. and Balas, V.E. (Eds.): *Conference Proceedings of ICCET-2021 in Applied Information Processing Systems*, pp.93–102, Springer Singapore, Singapore.

Das, S., Ghosh, I.D. and Chattopadhyay, A. (2022b) 'Sclera biometrics in restricted and unrestricted environment with cross dataset evaluation', *Displays*, Vol. 74, Article No. 102257, ISSN: 0141-9382.

Falohun, A., Fenwa, O. and Ajala, F. (2016) 'A fingerprint-based age and gender detector system using fingerprint pattern analysis', *International Journal of Computer Applications*, February, Vol. 136, No. 4, pp.43–48, Foundation of Computer Science (FCS), NY, USA.

Galbally, J., Haraksim, R. and Beslay, L. (2019) 'A study of age and ageing in fingerprint biometrics', *IEEE Transactions on Information Forensics and Security*, Vol. 14, No. 5, pp.1351–1365.

Gerrits, N., Elen, B. et al. (2020) 'Publisher correction: age and sex affect deep learning prediction of cardiometabolic risk factors from retinal images', *Scientific Reports*, Vol. 10, No. 9432, Article No. 1198.

Hiba, S. and Keller, Y. (2021) 'Hierarchical attention-based age estimation and bias estimation', *CoRR* [online] https://arxiv.org/abs/2103.09882.

Islam, T.U., Awasthi, L.K. and Garg, U. (2021) 'Gender and age estimation from gait: a review', *International Conference on Innovative Computing and Communications, Advances in Intelligent Systems and Computing*, Springer Nature Singapore Pte Ltd, Vol. 1166, pp.947–961.

Jammoussi, A.Y., Ghribi, S.F. and Masmoudi, D.S. (2016) 'Genetic algorithms-based dominant feature selection for face detection application', *International Journal of Computational Vision and Robotics*, Vol. 6, Nos. 1–2, pp.77–93.

Jana, R., Datta, D. and Saha, R. (2015) 'Age estimation from face image using wrinkle features', *Procedia Computer Science*, Vol. 46, No. 1, pp.1754–1761.

Kamarajugadda, K.K. and Polipalli, T.R. (2019) 'Extract features from periocular region to identify the age using machine learning algorithms', *Journal of Medical Systems*, Vol. 43, No. 1, pp.1–15.

Liu, X., Zou, Y., Kuang, H. and Ma, X. (2020) 'Face image age estimation based on data augmentation and lightweight convolution neural network', *Symmetry*, Vol. 12, No. 146, DOI: 10.3390/sym12010146.

Rajput, M.R. and Sable, G.S. (2020) 'Age group estimation from human iris', in Reddy, V., Prasad, V., Wang, J. and Reddy, K. (Eds.): *Soft Computing and Signal Processing. ICSCSP 2019. Advances in Intelligent Systems and Computing*, Springer, Singapore, Vol. 1118, pp.519–529.

Rot, P., Vitek, M., Grm, K., Emeršič, Ž., Peer, P. and Štruc, V. (2020) 'Deep sclera segmentation and recognition', in Uhl, A., Busch, C., Marcel, S. and Veldhuis, R. (Eds.): *Handbook of Vascular Biometrics*, pp.395–432, Springer, Cham.

Russell, R., Sweda, J.R., Porcheron, A. and Mauger, E. (2014) 'Sclera color changes with age and is a cue for perceiving age, health, and beauty', *Psychology and Aging*, Vol. 29, No. 3, pp.626–635.

Saxena, A.K. and Chaurasiya, V.K. (2017) 'Multi-resolution texture analysis for fingerprint based age-group estimation', *Multimedia Tools and Applications*, Vol. 76, No. 5, pp.3087–3097.

Saxena, A.K., Sharma, S. and Chaurasiya, V.K. (2015) 'Neural network based human age-group estimation in curvelet domain', *Eleventh International Multi-Conference on Information Processing-2015 (IMCIP-2015)*, pp.781–789.

Scullin, S.S. and Lloyd, G.O. (2002) 'Electronic customer relationship management: benefits, considerations, pitfalls and trends', *Proceedings of the IS One World Conference*, Las Vegas, Nevada, 3–5 April.

Simonyan, K. and Zisserman, A. (2014) *Very Deep Convolutional Networks for Large-Scale Image Recognition*, arXiv 1409.1556.

Thorley, C., Acton, B., Armstrong, J. et al. (2022) 'Are estimates of faces' ages less accurate when they wear sunglasses or face masks and do these disguises make it harder to later recognise the faces when undisguised?', *Cogn. Research*, February, Vol. 7, No. 1, Article No. 17.

Vitek, M., Das, A., Pourcenoux, Y., Missler, A., Paumier, C., Das, S., Ghosh, I.D. et al. (2020a) 'SSBC 2020: sclera segmentation benchmarking competition in the mobile environment', *International Joint Conference on Biometrics (IJCB 2020)*.

Vitek, M., Rot, P., Štruc, V. and Peer, P. (2020b) 'A comprehensive investigation into sclera biometrics: a novel dataset and performance study', *Neural Computing & Applications*, Vol. 32, No. 1, pp.1–15.

Wang, M. and Chen, W. (2021) 'Age prediction based on a small number of facial landmarks and texture features', *Technology and Health Care: Official Journal of the European Society for Engineering and Medicine*, Vol. 29, No. S1, pp.497–507.

Watson, P.G. and Young, R.D. (2004) 'Scleral structure, organization and disease. A review', *Experimental Eye Research*, Vol. 78, No. 3, pp.609–623.

Wikipedia (n.d.) *Age of Majority* [online] https://en.wikipedia.org/wiki/Age_of_majority (accessed 14 May 2021).

Yi, T. (2022) 'Estimation of human age by features of face and eyes based on multilevel feature convolutional neural network', *Journal of Electronic Imaging*, Vol. 31, No. 4, pp.1–14.

Zhang, H., Zhang, Y. and Geng, X. (2021) 'Practical age estimation using deep label distribution learning', *Front. Comput. Sci.*, Vol. 15, Article No. 153318 [online] https://doi.org/10.1007/s11704-020-8272-4.

Zhou, Y., Romijnders, R., Hansen, C., van Campen, J., Maetzler, W., Hortobágyi, T. and Lamoth, C.J.C. (2020) 'The detection of age groups by dynamic gait outcomes using machine learning approaches', *Scientific reports nature research*, March, Vol. 10, No. 1, Article No. 4426.