



# Age Prediction from Sclera Images using Deep Learning

P. O. Odion<sup>a</sup>, M. N. Musa<sup>b,\*</sup>, S. U. Shuaibu<sup>a</sup>

<sup>a</sup>Computer Science Dept, Nigerian Defence Academy Kaduna, Nigeria

<sup>b</sup>Cyber Security Dept, Nigerian Defence Academy Kaduna, Nigeria

---

## Abstract

Automatic age classification has drawn the interest of many scholars in the fields of machine learning and deep learning. In this study, we looked at the problem of estimating age groups using different biometric modalities of human beings. We looked at the problem of determining age groups in humans using various biometric modalities. Specifically, we focused on the use of transfer learning for sclera age group classification. 2000 Sclera images were collected from 250 individuals of various ages, and Otsu thresholding was used to segment the images using morphological processes. Experiment was conducted to determine how accurately the age group of a person can be classified from sclera images using pre-trained CNN architectures. The segmented images were trained and tested on four different pre-trained models (VGG16, ResNet50, MobileNetV2, EfficientNet-B1), which were compared based on different performance metrics in which ResNet-50 was shown to outperform the others, resulting in an accuracy, precision, recall and F1-score of 95% while VGG-16, EfficientNetB1, and MobileNetV2 had 94%, 93%, and 91%, respectively. The findings from the study showed that there is an aging template in the sclera that can be utilized to classify age.

DOI:10.46481/jnsps.2022.787

**Keywords:** sclera images, pre-trained CNN, age estimation, deep learning, segmentation, SBVPI dataset

### Article History:

Received: 26 April 2022

Received in revised form: 30 June 2022

Accepted for publication: 15 July 2022

Published: 15 August 2022

© 2022 The Author(s). Published by the Nigerian Society of Physical Sciences under the terms of the Creative Commons Attribution 4.0 International license (<https://creativecommons.org/licenses/by/4.0>). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Communicated by: T. Latunde

---

## 1. Introduction

Every human being possesses distinct biometric traits, which are employed in individual identification, recognition, verification, age estimation, and authentication due to their uniqueness. Face recognition, ear recognition, voice recognition, iris recognition, retina scan, finger recognition, vein recognition, and deoxyribonucleic acid (DNA) matching are just a few of the biometric technologies that have become a hot topic in academia. Facial recognition is the most prominent area of human recognition since the human face reveals many features such as age, gender, race, emotion, expression, and identity [1].

Age, as a human characteristic, plays a significant role in facilitating or limiting communication. It functions as a barrier, affecting both how we interact with one another and how we comprehend what others are saying, much as language, culture, beliefs, and experience do. However, as people age, their faces change and their skin thickens, as well as their colour and texture. The tissue composition begins to be more sub-cutaneous and the facial skeleton lines or wrinkles appear. The ability of a machine to recognize and interpret faces or facial traits such as age, emotion, or gender in real time gave rise to the concept of computer vision [2].

Machine learning (ML) algorithms have been used to address issues in several domains, by analysing and understanding vast amounts of data [3-6]. Machine Learning has benefited

---

\*Corresponding author tel. no: +234(0)8166046507

Email address: muhammadmusa2502@nda.edu.ng (M. N. Musa)

in the detection and identification of facial images, as well as the age estimation [7]. The method employs classic computer vision techniques that employ handcrafted features, as well as a step framework that include an aging learning pattern [2].

Most facial recognition algorithms that use ML lack generalizability when applied to unseen photos [7]. Factors that affect facial recognition and facial age estimation include race, beards, plastic surgery, skin tone, etc. This has led to age estimation research using facial features being extensively studied with the aim of finding out aging patterns and variations and how to best characterize an aging face for accurate age estimation [8].

The failings in using facial images for age estimation led to research in other areas, among which include eye regions like the iris, sclera, and retina, in which iris recognition is the predominant technology in the area [9]. The problem with using ML method in eye region recognition for age estimation occurs due to the eye region been so small and extracting the features manually using Histogram of Oriented Gradients (HOG), Principal Component Analysis (PCA) is undesirable as we might miss important features during the cause of building the model [10].

Deep learning using Convolutional Neural Networks (CNN) has recently piqued the interest of many computer vision researchers due to its superior ability to learn a series of non-linear features directly from raw pixels [11]. This prompted researchers to concentrate their efforts on various eye region experiments using deep learning to observe template changes. Images of the retina reveal considerable age-related changes. Bruch's membrane, which is present in the retina, is especially prone to aging [12]. High-resolution retinal images, on the other hand, are taken under limited imaging settings by professional specialists using sophisticated imaging devices. There is also evidence of studies into the use of the iris as a method of determining age [13]. Infrared cameras are typically used to capture iris images from a close range. Retina or iris images are not an obvious choice for age estimation because image gathering is becoming more difficult and expensive. This made researchers consider the sclera for age estimation because the sclera is the white visible portion of the eye image and thus it remains visible for various gaze directions [14]. It can also be captured with hand-held cameras or mobile cameras. Even though the sclera is relatively stable over time, its colour changes with age and health [15, 16].

Due to the fact that training deep learning models consumes a lot of resources and is computationally expensive [17], coupled with the inability to gather millions of eye images to train deep learning models from scratch, This motivates us to explore the use of some of the best pre-trained models that were state-of-the-art in the ImageNet contest (VGG-16 [18], ResNet-50 [19], MobileNetv2 [20], EfficientNet [21]) to predict human age using the sclera of the eye. To achieve this, we were able to acquire eye images from individuals between the ages of 5 to 30 years old, apply Otsu thresholding to execute segmentation of the sclera from the acquired eye images, train, validate and interpret the result generated using different performance metrics.

## 2. Literature Review

The first research conducted on age estimation was in 1994 by Kwon and da Vitoria Lobo, in which the age was simply divided into numerous ranges [11]. Many studies have focused on age estimation using facial images, with few studies focusing on eye regions [10]. There is evidence of research using the iris as a modality for age estimation [14]. Some of the research that used iris for classification includes [22] who conducted an experiment using 596 iris images, where 300 consisted of young subjects and two hundred and 296 elderly subjects by employing a random forest algorithm to extract features and classify the images, resulting in a classification rate of 64.68%. [23] suggested an alternative method that relies on five geometric features extracted from the human iris using a total of 210 respondents, spanning from 18 to 73, and were classified as young (25), adult (25–60), and senior (> 60). The authors used segmentation to aid in efficiently detecting iris and pupil boundaries. They were able to extract 12 geometric features from the iris and pupil parameters, which were trained and tested on different classifiers. They achieved a 75% accuracy rate. Again, [24] proposed a technique that used the iris structure to estimate a person's age group. The image input was obtained from an iris database and was divided into three age groups. The iris boundaries were localized using a circular Hough Transform technique for image pre-processing and segmentation. They used five different classifiers (K nearest neighbour, fine Gaussian support vector machine, decision tree, bagged ensemble, and linear discriminant) were used, with the bagged ensemble classifier outperforming the others because it reduces the problem of overfitting of training data and reduces the variance of the estimate. The suggested model attained an overall accuracy of 83.7% and outperformed prior state-of-the-art models. Researchers have recently begun to use deep learning methods in the field of eye region biometrics for age prediction, as shown in the research by [25]. They used deep learning methods to test the iris for gender and age classification. They began by estimating gender using deep CNN models like AlexNet and GoogleNet, which were trained on a real-time database of 213 people spanning 3-73 years. The features extracted from the human iris were fed into a multi-class SVM using these CNN models. In comparison to GoogleNet, AlexNet performed better, with an overall accuracy level of 95.34%. Similarly, their trained model for age prediction indicates that the anticipated age was correct for virtually all of the subjects. They also indicate that gender classification performs better than age classification.

In recent years, sclera images obtained in visible light have been used in biometric recognition systems [26, 27]. However, [27] used the VGG architecture to create a SegNet and ScleraNet model for sclera segmentation and recognition, respectively. They conducted rigorous experiments and tested their findings using Sclera Blood Vessel, Periocular, and Iris (SBVPI), a new public dataset that captures diverse gaze directions of the eye to comprehensively represent all parts of the sclera. The SBVPI collection contains images of 55 distinct people looking at four different directions, with the iris, sclera,

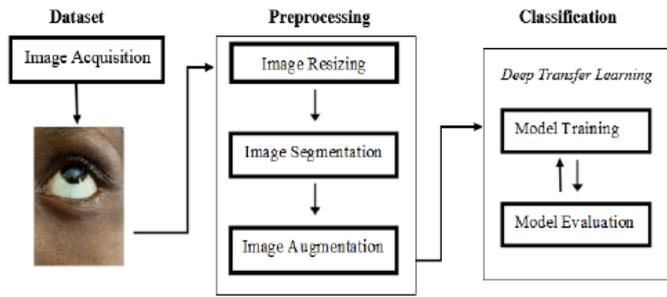


Figure 1. Proposed Methodology



Figure 2. Image Captured at Different Gaze Direction

| Age Group | Age Range | Number of subjects | Number of images |
|-----------|-----------|--------------------|------------------|
| Kids      | 5 - 13    | 84                 | 672              |
| Teens     | 14 - 20   | 85                 | 680              |
| Adult     | 21 - 30   | 80                 | 640              |
|           | Total     | 250                | 2000             |

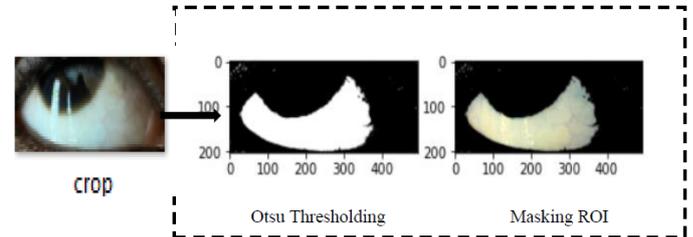


Figure 3. Image Segmentation using Otsu thresholding algorithm

blood vessels, eyelashes, and periocular segment. Other models (U-net, Refine-Net) were compared to SegNet, and it surpassed the preceding model in terms of processing speed and accuracy, resulting in highly competitive recognition performance. [14] recently published the first study on the analysis of sclera for age estimation, in which they employed a modified VGG-16 network on a custom dataset. They approached it as a regression problem, resulting in a mean absolute error (MAE) of 0.06. This indicates that there is an aging template in the sclera and can be used for age prediction. Hence, this research employed a custom dataset for use in age prediction using a classification technique. The newly captured images are then compared using some state-of-the-art pre-trained models. This research is novel, as to the best of our knowledge, there is no work that uses sclera to predict age using a classification technique.

### 3. Methodology

The proposed sclera-based age group classification method is shown in Figure 1.

#### 3.1. Image Acquisition

This is the first step towards putting the proposed approach into action. This was done with an iPhone 11 pro max set at the highest resolution and quality. The photos were taken in an unstructured setting. A total of 2000 photographs were taken of 250 people, ranging in age from 5 to 30, and were divided into three categories: children, teenagers, and adults, as in Table 1. The image acquisition setup was inspired by [27]. In their work, they generated the first standardized dataset that consists of images captured in different gaze directions for sclera segmentation and recognition. Each subject was asked to change gaze direction four times, that is, straight, upward, left, and right for both left and right eye as shown in Figure 2.

#### 3.2. Pre-processing

Several undesirable facial features, such as the eyebrows, a portion of the nose, the cheeks, and so on, are included in the captured image. As a result, the images were manually cropped and resized to reduce the amount of noise and focus just on the region of interest, which is the eye, while maintaining the aspect ratio. Since 10 of the initial collected images were fuzzy and the region of interest was obscured, they were not used for segmentation. As a result, we now have a final image of 1990. The Otsu thresholding technique was used to segment the images, as well as perform morphological operations. It is thought to be one of the simplest and most successful binary image segmentation methods [28]. It divides the image into two classes, white for the foreground and black for the background, based on its grayscale characteristics as seen in Figures 3 and 4, and then we mask the region of interest. According to [29], the algorithm for the Otsu method is outlined below:

1. Compute histogram and probabilities of each intensity level
2. Initialize  $w_i$  and  $\mu_i$  to zero (0) respectively
3. Iterate over all possible threshold  $t = 0 \dots \text{max-intensity}$ 
  - (a) Update  $w_i$  and  $\mu_i$
  - (b) Compute the between the class variance  $\sigma_b^2(t)$
4. The final threshold is the maximum  $\sigma_b^2(t)$

To increase the size of the training dataset, image augmentation was performed, which involves applying geometric transformation techniques to the training images, such as shearing, contrasting, horizontally flipping, spinning, zooming, and blurring. The dataset is then downsized to  $224 \times 224$  pixels and converted to array format, with the pixel intensities adjusted to the range  $[-1, 1]$ . This has proved to be a successful approach [30].

Table 2. Comparison with different pre-trained CNN model with respect to performance metrics

| layer name | output size | 18-layer  | 34-layer  | 50-layer  | 101-layer  | 152-layer  |
|------------|-------------|---|---|---|--|--|
| conv1      | 112 × 112   |   |   | 7 × 7, 64, stride 2   |  |  |
|            |             |   |   | 3 × 3, max pool, stride 2   |  |  |
| conv2_x    | 56 × 56     | $\begin{bmatrix} 3 \times 3, & 64 \\ 3 \times 3, & 64 \end{bmatrix} \times 2$   | $\begin{bmatrix} 3 \times 3, & 64 \\ 3 \times 3, & 64 \end{bmatrix} \times 3$   | $\begin{bmatrix} 1 \times 1, & 64 \\ 3 \times 3, & 64 \\ 1 \times 1, & 256 \end{bmatrix} \times 3$    | $\begin{bmatrix} 1 \times 1, & 64 \\ 3 \times 3, & 64 \\ 1 \times 1, & 256 \end{bmatrix} \times 3$     | $\begin{bmatrix} 1 \times 1, & 64 \\ 3 \times 3, & 64 \\ 1 \times 1, & 256 \end{bmatrix} \times 3$     |
| conv3_x    | 28 × 28     | $\begin{bmatrix} 3 \times 3, & 128 \\ 3 \times 3, & 128 \end{bmatrix} \times 2$ | $\begin{bmatrix} 3 \times 3, & 128 \\ 3 \times 3, & 128 \end{bmatrix} \times 4$ | $\begin{bmatrix} 1 \times 1, & 128 \\ 3 \times 3, & 128 \\ 1 \times 1, & 512 \end{bmatrix} \times 4$  | $\begin{bmatrix} 1 \times 1, & 128 \\ 3 \times 3, & 128 \\ 1 \times 1, & 512 \end{bmatrix} \times 4$   | $\begin{bmatrix} 1 \times 1, & 128 \\ 3 \times 3, & 128 \\ 1 \times 1, & 512 \end{bmatrix} \times 8$   |
| conv4_x    | 14 × 14     | $\begin{bmatrix} 3 \times 3, & 256 \\ 3 \times 3, & 256 \end{bmatrix} \times 2$ | $\begin{bmatrix} 3 \times 3, & 256 \\ 3 \times 3, & 256 \end{bmatrix} \times 6$ | $\begin{bmatrix} 1 \times 1, & 256 \\ 3 \times 3, & 256 \\ 1 \times 1, & 1024 \end{bmatrix} \times 6$ | $\begin{bmatrix} 1 \times 1, & 256 \\ 3 \times 3, & 256 \\ 1 \times 1, & 1024 \end{bmatrix} \times 23$ | $\begin{bmatrix} 1 \times 1, & 256 \\ 3 \times 3, & 256 \\ 1 \times 1, & 1024 \end{bmatrix} \times 36$ |
| conv5_x    | 7 × 7       | $\begin{bmatrix} 3 \times 3, & 512 \\ 3 \times 3, & 512 \end{bmatrix} \times 2$ | $\begin{bmatrix} 3 \times 3, & 512 \\ 3 \times 3, & 512 \end{bmatrix} \times 3$ | $\begin{bmatrix} 1 \times 1, & 512 \\ 3 \times 3, & 512 \\ 1 \times 1, & 2048 \end{bmatrix} \times 3$ | $\begin{bmatrix} 1 \times 1, & 512 \\ 3 \times 3, & 512 \\ 1 \times 1, & 2048 \end{bmatrix} \times 3$  | $\begin{bmatrix} 1 \times 1, & 512 \\ 3 \times 3, & 512 \\ 1 \times 1, & 2048 \end{bmatrix} \times 3$  |
|            | 1 × 1       |   |   | average pool, 1000-d fc, softmax  |  |  |
| FLOPs      |             | 1.8 × 10 <sup>9</sup>   | 3.6 × 10 <sup>9</sup>   | 3.8 × 10 <sup>9</sup>   | 7.6 × 10 <sup>9</sup>  | 11.3 × 10 <sup>9</sup>   |

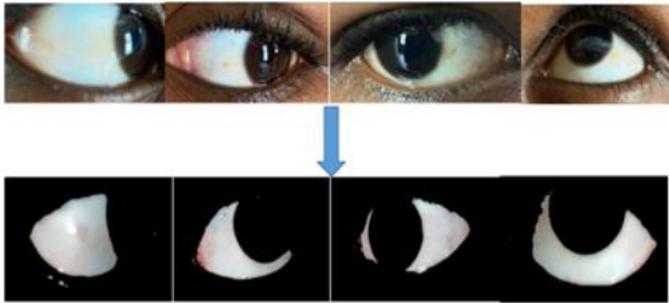


Figure 4. Masking of RGB Images using Otsu thresholding

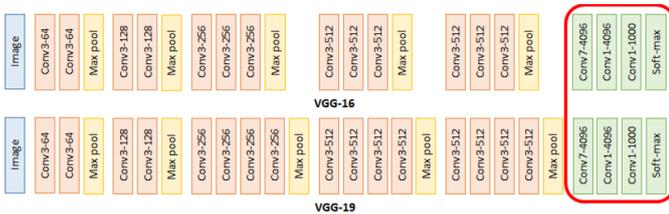


Figure 5. VGG Architecture [18]

size to 32. The model was evaluated on the remaining 20% of the dataset using various performance metrics such as accuracy, precision, recall, and F1 score.

### 3.4. VGG-16 Model

Previous CNN architectures have proven successful in recent years, particularly in image recognition applications, attracting the attention of many academics. [18] presented a simple effective deep CNN architecture known as Visual Geometry Group (VGG) in 2015, with a network of sixteen layers (VGG16) and nineteen layers (VGG19), which was far deeper than previous ZfNet and AlexNet models. It was created to demonstrate that the depth of a network is vital in getting better recognition or classification in CNNs. The VGG replaced the 11\*11 and 5\*5 filters with a stack 3\*3 filter layer, demonstrating that using smaller filters with fewer parameters reduces the computational cost of the network. The network also adds a set of 1\*1 convolutions in between the convolutional layer as shown Figure 5.

VGG16 performed exceptionally well in image classification and localization. Furthermore, it ranked second in the 2014-ILSVRC competition and has become well-known for its simplicity, enhanced depth, and homogenous topology.

### 3.5. ResNet Model

Residual Network was developed by [19] to eliminate the vanishing gradient problem that previous networks experienced. It was declared the 2015-ILSVRC winner. The ResNet architecture introduced the residual learning framework notion. It was created with a variety of layers, including 34, 50, 101, and 152 layers, as shown in Table 2. Despite the increased depth, it has low computational complexity when compared to earlier models.

ResNet established shortcut connections within layers to facilitate cross-layer interconnection, and these residual linkages tend to accelerate deep network convergence, aiding the network in avoiding gradient fading. The ResNet architecture's representational depth is thought to be advantageous for any

Table 3. Comparison with different pre-trained CNN model with respect to performance metrics

| CNN Model       | Accuracy | Precision | Recall | F1 Score |
|-----------------|----------|-----------|--------|----------|
| MobileNetV2     | 0.92     | 0.92      | 0.91   | 0.92     |
| EfficientNet-B1 | 0.93     | 0.93      | 0.93   | 0.93     |
| ResNet-50       | 0.95     | 0.95      | 0.95   | 0.95     |
| VGG-16          | 0.94     | 0.94      | 0.94   | 0.94     |

### 3.3. Age Classification

The goal of image classification is to determine the class of an input image using its attributes [31]. We used state-of-the-art pre-trained CNN models for both feature extraction and classification. The pre-trained models were trained and tested on the dataset. We divided the dataset into 80% for training and set hyperparameter constants such as the initial learning rate of 0.00001, the number of training epochs to 50, and the batch

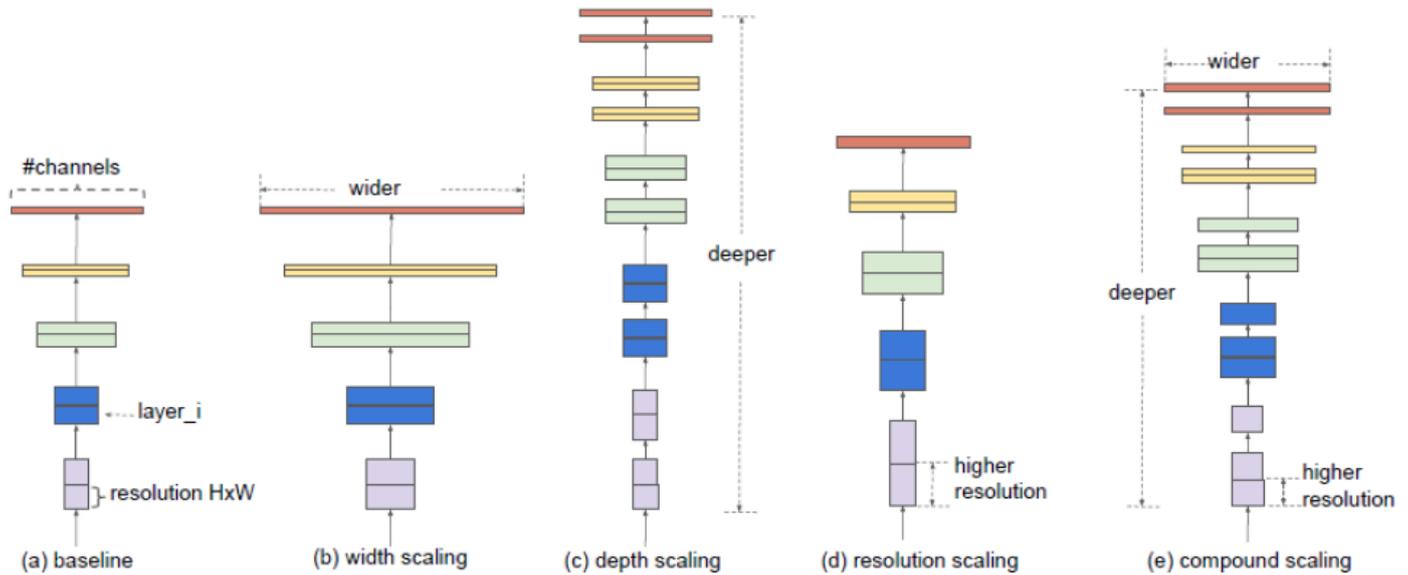


Figure 6. Comparison of Model Scaling [21]

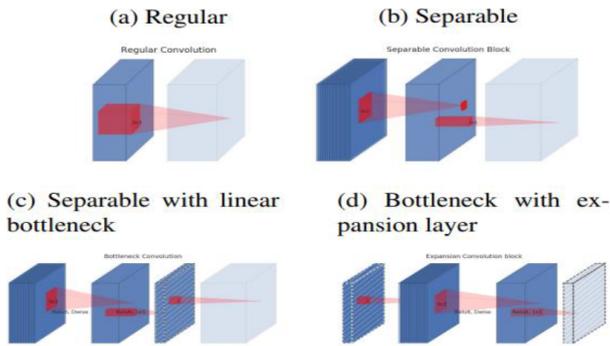


Figure 7. Evolution of separable convolution blocks [20]

image recognition challenge. The authors demonstrated experimentally that the ResNet with 50/101/152 layers has fewer errors in image classification problems and likewise gain an improvement of 28% on the well-known image recognition dataset name COCO.

### 3.6. EfficientNet

The EfficientNet model was developed by [21] to scale models uniformly in all dimensions of depth/width/resolution using a simple yet extremely effective compound scaling that can lead to greater performance. This method may scale up MobileNets and ResNets by employing neural architecture search to construct a new baseline network and scale it up to obtain a family of models known as Efficient-Nets, which achieve significantly higher accuracy and efficiency than earlier ConvNets.

As demonstrated in Figure 6, EfficientNet employs a compound scaling method that uniformly scales all three dimensions with a constant ratio. The EfficientNet family ranges from B0 to B7. Starting with the baseline EfficientNet-B0, the compound scaling method is used to scale it up in two steps: The compound scaling approach begins with a grid search to de-

termine the relationship between multiple scaling dimensions of the baseline network under a set resource limitation (e.g., 2x more FLOPs). This finds the appropriate scaling coefficient for each dimension before scaling up the baseline network to the specified target model size or computational budget. When scaling up existing models, this compound scaling strategy reliably enhances model accuracy and efficiency such as MobileNet (+1.4% ImageNet accuracy), and ResNet (+0.7%), compared to conventional scaling methods.

### 3.7. MobileNetV2

MobileNetV2 network was developed by [20] based on depthwise separable convolution. MobileNetV2 is a tensorflow-based family of mobile computer vision models designed specifically to meet the requirements of low-resource efficient systems, maximize accuracy, and improve the state-of-the-art performance of mobile models on multiple tasks and benchmarks, as well as across a spectrum of model sizes. It is built on an inverted residual structure with quick connections between narrow bottleneck layers as shown in Figure 7. As a source of non-linearity, the intermediate expansion layer filters features using lightweight depthwise convolutions. Non-linearities were also reduced in the narrow layers to maintain representational power and increase performance.

Figure 7's diagonally hatched texture denotes non-linear layers. The final (lightly colored) layer marks the start of the next block. when stacked, c and d are equivalent blocks. The network performance was measured on ImageNet classification, COCO object detection, VOC image segmentation, the architecture improved the state of the art for wide range of performance points [20].

### 3.8. Performance Metrics

The practice of analysing how well a model works against real data is known as performance evaluation. Standard performance metrics like as precision, recall, and the F1-score, which

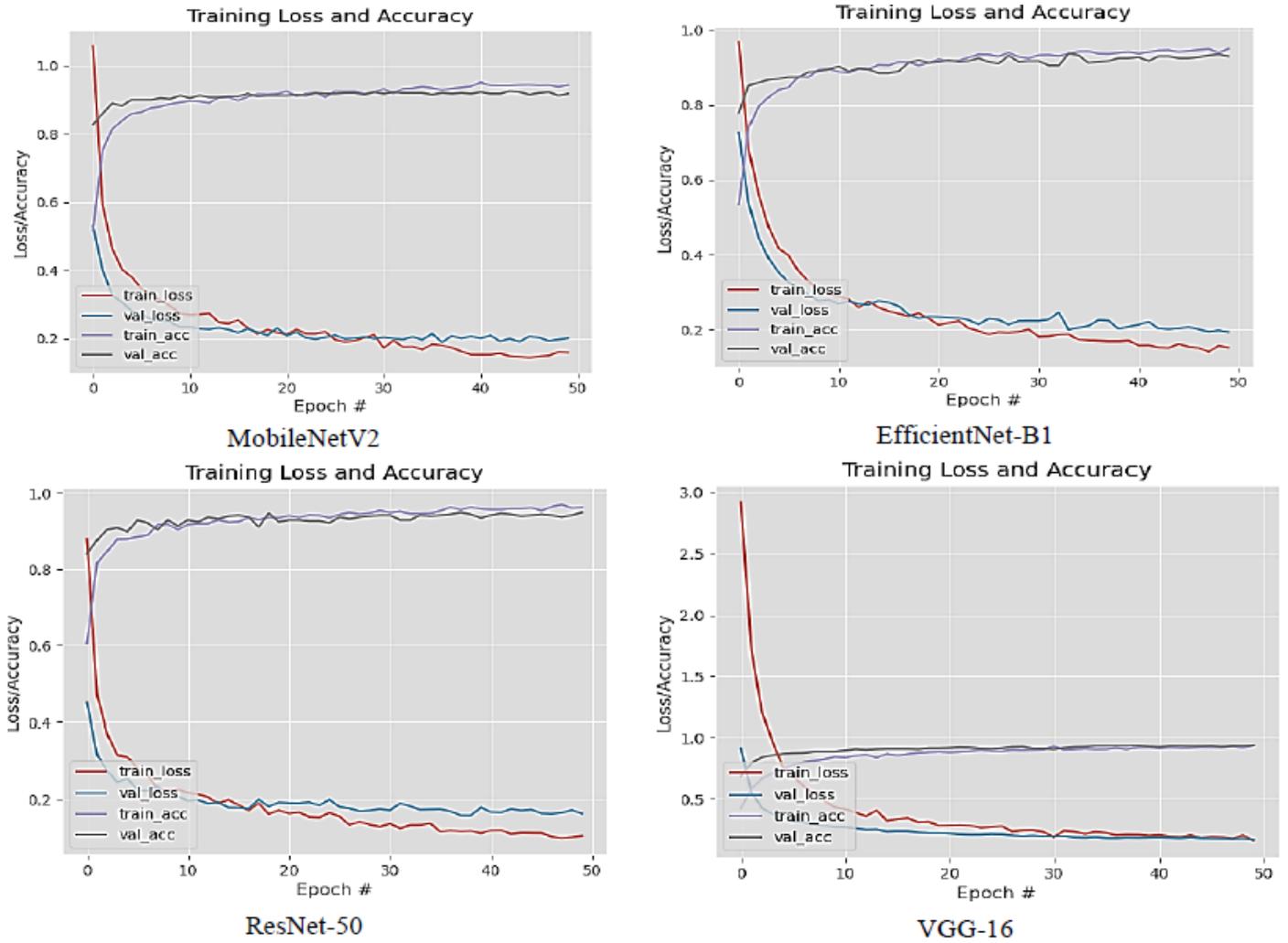


Figure 8. Graphs of Accuracy and Loss Score for the Four Pre-trained Models

are defined as follows [27], are used to assess the classification models' performance:

*Accuracy*

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (1)$$

This is one of the popular metrics for assessing a classification model. It is the percentage of correct predictions in a given test data.

*Precision*

It is about how precise or how often the prediction is correct. It is the ratio of true positives to the sum of true positives and false positives. It is calculated as

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

*Recall*

How often is the prediction correct when the actual value is positive? It is calculated mathematically as the ratio of true

positives to the sum of true positives and false negatives as in equation 3.

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

*F1-measure*

The F1-measure is also known as the F1 Score. It is the harmonic mean of precision and recall. A harmonic mean is appropriate for situations where the average of rates (a ratio between two related quantities) is desired. It is calculated as

$$F1\ Score = \frac{2 \times Precision \times Recall}{(Precision \times Recall)} \quad (4)$$

where, TP denotes the number of true positive pixels, FP stands for the number of false positive pixels, FN represents the number of false negative pixels and TN represents the number of true negative pixels.

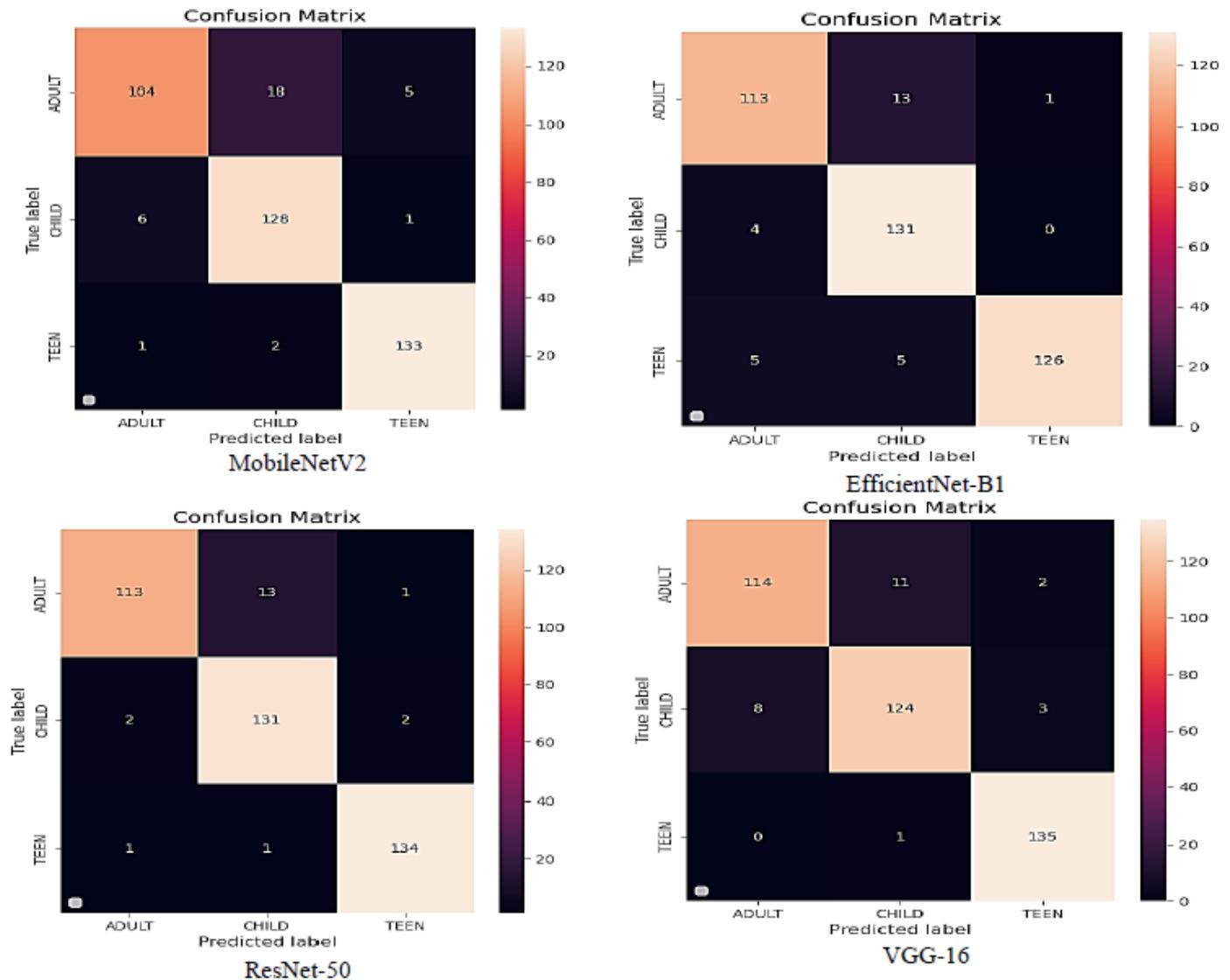


Figure 9. Confusion Matrix of Different Pre-train Networks used

#### 4. Results and Discussion

The goal of the study was to see how accurate a pre-trained deep learning model could be in classifying a person's age group from the sclera of the human eye. Previous studies have demonstrated that the sclera ages and changes over time, and that it can be utilized to estimate age [11]. The dataset was trained and tested on four distinct pre-trained convolutional neural networks, including MobileNetV2, VGG-16, ResNet-50, and EfficientNet-B1, as shown in the results. Table 3 compares the four different neural networks with respect to accuracy, precision, recall, and F1-Score while Figure 8 shows their accuracy and loss score across 50 epochs.

The comparison of the four pre-trained models is shown in Figures 8 and 9. The ResNet-50 achieved the best accuracy of 95%, while VGG-16, EfficientNet-B1, and MobileNetV2 had 94, 93, and 91%, respectively. It also shows that the adult class has a higher rate of misclassification, with that class accounting for the majority of the misclassification. This could be due to

the large quantity of images for children aged 13 and 14, with 13 falling into the children's class and 14 falling into the teen category. These are the boundaries for each class, meaning only months separate the two classes, which could result in misclassification.

#### 5. Conclusion

In this study, we used a classification technique to evaluate an approach to age prediction. Other methods for age prediction were carefully investigated. It was discovered that facial images for age prediction lack generalizability while retina and iris images for age prediction requires the images to be captured using a specialized cameras. Our method demonstrated that sclera images captured through hand-held cameras can also provide a very high accuracy. About 2000 images were taken in an uncontrolled environment from 250 people. To solve the problem of inadequate data, pre-trained networks were used to

train the data. The classification process was evaluated with four different models: MobileNetv2, EfficientNet-B1, ResNet-50, and VGG-16 resulting in an accuracy of 92%, 93%, 95% and 94% respectively. Based on the findings, it can be stated that template aging of the sclera in the eye does exist over time, as evidenced by the findings of this study and also ResNet-50 was most successful when applied for sclera classification. The study also suggests that with more data and better segmentation, greater accuracy can be attained, potentially rivalling other areas such as facial and iris age prediction.

## Acknowledgments

We thank the referees for the positive enlightening comments and suggestions, which have greatly helped us in making improvements to this paper.

## References

- [1] S. Narejo, E. Pasero & F. Kulsoom, "EEG based eye state classification using deep belief network and stack encoder", *International Journal of Electrical Computer Engineering* **6** (2016) 3131.
- [2] A. Othmani, A. Taleb, H. Abdelkawy & A. Hadid, "Age estimation from faces using deep learning: A comparative analysis", *Computer Vision and Image Processing* **196** (2020) 102961.
- [3] D. O. Oyewola, G. D. Emmanuel, N. N. Juliana, A. U. Terrang & S. A. Akinwunmi, "COVID-19 Risk Factors, Economic Factors, and Epidemiological Factors nexus on Economic Impact: Machine Learning and Structural Equation Modelling Approaches", *Journal of the Nigerian Society of Physical Sciences* **3** (2021) 395.
- [4] V. Umarani, A. Julian & J. Deepa, "Sentiment Analysis using various Machine Learning and Deep Learning Techniques", *Journal of the Nigerian Society of Physical Sciences* **3** (2021) 385.
- [5] A. B. Yusuf, R. M. Dima & S. K. Aina, "Optimized Breast Cancer Classification using Feature Selection and Outliers Detection", *Journal of the Nigerian Society of Physical Sciences* **3** (2021) 298.
- [6] O. Olubi, E. Oniya & T. Owolabi, "Development of predictive model for radon-222 estimation in the atmosphere using stepwise regression and grid search based-random forest regression" *Journal of the Nigerian Society of Physical Sciences* **3** (2021) 132.
- [7] A. Mollahosseini, D. Chan & M. H. Mahoor, "Going deeper in facial expression recognition using deep neural networks", *IEEE Winter Conference on Applications of Computer Vision* (2016) 1.
- [8] R. Angulu, J. R. Tapamo & A. O. Adewumi, "Age estimation via faces: A Survey", *EURASIP Journal of Image and Video Processing* **2018** (2018) 1.
- [9] K. Wang & A. Kumar, "Cross-spectral iris recognition using supervised CNN and supervised discrete hashing", *Pattern Recognition* **86** (2019) 85.
- [10] J. R. Beattie, A. M. Pawlak, J. J. McGarvey & A. W. Stitt, "Sclera as a surrogate marker for determining AGE-modifications in Bruch's membrane using a Raman spectroscopy-based index of aging", *Investigate ophthalmology & visual science* **52** (2011) 1593.
- [11] A. Abbasi & M. Khan, "Iris-pupil thickness based method for determining age group of a person", *International Arab Journal of Information Technology* **13** (2016).
- [12] S. Das, I. D. Ghosh & A. Chattopadhyay, "Deep age estimation using sclera images in multiple environment", *Applied Information Processing Systems* (2022) 93.
- [13] V. Patil & A. Patil, "Human identification method: sclera recognition", *International Journal of Computer Science Networks* **6** (2017) 24.
- [14] R. Russell, J. R. Sweda, A. Porcheron & E. Mauger, "Sclera color changes with age and is a cue for perceiving age, health and beauty", *Psychology and Aging* **29** (2014) 626.
- [15] S. Taheri & O. Toygar, "On the use of DAG-CNN architecture for age estimation using multi-stage features fusion", *Neurocomputing* **329** (2019) 300.
- [16] C. C. Aggarwal, "Neural Networks and Deep Learning", *Springer* **10** (2018) 978.
- [17] S. Tammina, "Transfer learning using vgg-16 with deep convolutional neural network for classifying images." *International Journal of Scientific and Research Publications (IJSRP)* **10** (2019) 143.
- [18] K. Simonyan & A. Zisserman, "Very deep convolutional networks for large-scale image recognition" *arXiv preprint arXiv:1409.1556* (2014).
- [19] K. He, X. Zhang, S. Ren & J. Sun, "Deep residual learning for image recognition" *In Proceedings of the IEEE conference on computer vision and pattern recognition* (2016) 770.
- [20] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov & L.C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks" *In Proceedings of the IEEE conference on computer vision and pattern recognition* (2018) 4510.
- [21] M. Tan & Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks" *International conference on machine learning*. PMLR (2019) 6105.
- [22] A. Sgroi, K. W. Bowyer & P. J. Flynn, "The Prediction of Old and Young Subjects from Iris Texture", *International Conference on Biometrics* (2013) 1.
- [23] M. Erbilek, M. Fairhurst & M. D. C. A. Cristiany, "Age prediction from iris biometrics", *5th International Conference on Imaging for Crime Detection and Prevention (ICDP)* (2013) 1.
- [24] M. R. Rajput & G. S. Sable, "Age Group Estimation from Human Iris", *Advances in Intelligent Systems and Computing: Soft Computing and Signal Processing*, **1118** (2019) 519.
- [25] M. Rajput & G. Sable, "Deep learning based gender and age estimation from human iris", *Proceedings of the International Conference on Advances in Electronics Electrical & Computational Intelligence (ICAEEC)* (2019).
- [26] S. Das, I. D. Ghosh & A. Chattopadhyay, "An efficient deep learning strategy: Its application in sclera segmentation", *IEEE Applied Signal Processing Conference (ASPCON)* (2020) 232.
- [27] P. Rot, M. Vitek, K. Grm, Z. Emersic, P. Peer & V. Struc, "Deep sclera segmentation and recognition", in *Handbook of vascular biometrics Cham* (2020) 395.
- [28] J. T. C. Ming, N. M. Noor, O. M. Rijal, R. M. Kassim & A. Yusuf, "Lung disease classification using different deep learning architectures and principal component analysis", *2nd International Conference on Biosignal Analysis, Processing and Systems (ICBAPS)*, IEEE (2018) 187.
- [29] S.H. Tsang, "Review: MobileNetV2 - lightweight model (image classification)", 19 May 2019. [Online]. Available: <https://towardsdatascience.com/review-mobilenetv2-light-weight-model-image-classification-8febb490e61c>.
- [30] M. N. Musa, N. O. Badmos, I. R. Saidu, & U. Abdulrazaq, "Protective face covering: An application of MobileNetV2 detector", *International Research Journal of Science, Technology, Education, and Management* **2** (2022) 1.
- [31] L. H. Thai, T. S. Hai & N. T. Thuy, "Image classification using support vector machine and artificial neural network", *International Journal of Information Technology and Computer Science* **4** (2012) 32.