CHAOS
Theory and Applications
in Applied Sciences and Engineering

e-ISSN: 2687-4539
RESEARCH ARTICLE
Vol.5 / No.2 / 2023 / pp.105-110
https://doi.org/10.51537/chaos.1257597

Prediction of Gender and Age Period from Periorbital Region with VGG16

Omer Faruk Akmese [*,1], **Huseyin Cizmeci** [β,2], **Selim Ozdem** [β,3], **Fikri Ozdemir** [α,4], **Emre Deniz** [*,5], **Rabia Mazman** [§,6], **Murat Erdogan** [§,7] and **Esma Erdogan** [§,8]

[*]Department of Computer Engineering, Hitit University, Corum, Türkiye, [β]Department of Computer Technology, Hitit University, Corum, Türkiye, [α]Department of Anatomy, Hitit University, Corum, Türkiye, [§]Department of Forensic Sciences, Hitit University, Corum, Türkiye.

ABSTRACT Using deep learning methods, age and gender estimation from people's facial area has become popular. Recently, with the increase in the use of masks due to Covid-19, only the eye area of people is seen. The periorbital region can give an idea about the person's characteristics, such as age and gender. This study it is aimed to predict gender and age from images obtained by cutting the eye area from facial photographs of people using Visual Geometry Group-16 (VGG16). With the transfer learning method for age group (male, female) and gender group (child, youth, adults, and old) classification, 5714 images in the data set were used for the age group, and 3280 images were used for the gender group. As a result of this study, 99.41% success in age estimation and 95.73% in gender estimation was achieved.

KEYWORDS
Deep learning
Age and gender prediction
VGG16
Periorbital area

INTRODUCTION

Humans are social beings that interact with the environment they live in. Gender plays a fundamental role in social life. With the gender difference situation, people's speech, form of address, and behavior also differ. These differences are just a few of the gender commitments in social interactions (Gündüz and Cedimoğlu 2019).

The use of artificial intelligence and deep learning applications is becoming more and more common (Hinton and Salakhutdinov 2006; Solmaz *et al.* 2020). Artificial intelligence applications that make predictions about age and appearance on social media are frequently preferred by users. It is seen that many mobile applications make applications such as age estimation, aging, and rejuvenation on mobile platforms. These applications are generally used for entertainment purposes. Another important example is that the Xiaomi Mi 6 has an 8-megapixel front camera. In addition to the automatic facial beautification filter, this front camera also offers the user a system that can predict gender and age with its artificial intelligence.

Artificial intelligence applications are becoming widespread in every field, and their applications have been seen in forensic cases recently (Zha *et al.* 2022; Aslan *et al.* 2022). The amount of data stored or transferred has increased with the intensive use of the Internet and information devices. As a result, there has been an increase in the crime rate. The amount of evidence obtained regarding the crimes committed has increased as the amount of data. Increasing evidence has made it difficult for experts in the field of forensic informatics to analyze the data with the available facilities. The disruptions experienced in the forensic informatics evidence and data analysis processes have caused negativities in the forensic trial processes (Dilber and Çetin 2021). In some forensic cases, the evidence is very insufficient. In such cases, artificial intelligence applications can come into play. This study developed models that predict gender and age group from photographs using deep learning methods. The study aimed to estimate the gender and age group using only eye photographs as evidence in a forensic case using deep learning algorithms. It is aimed to speed up the process of gender determination in the evidence related to this method and to facilitate data analysis.

Numerous studies have been conducted to improve and develop methods for assessing age and gender. In order to identify the gender and age of a single individual from their photo, Abu Nada *et al.* (2020) developed a double-check layer validator that makes use of deep learning methods, specifically a Convolu-

Manuscript received: 28 February 2023,
Revised: 31 May 2023,
Accepted: 7 June 2023.

[1]ofarukakmese@hitit.edu.tr
[2]huseyincizmeci@hitit.edu.tr (**Corresponding author**)
[3]selimozdem@hitit.edu.tr
[4]fikriozdemir@hitit.edu.tr
[5]emredeniz@hitit.edu.tr
[6]rabiayasaroglu.1969@gmail.com
[7]muraterdogan41@gmail.com
[8]esmardgn19@gmail.com

tional Neural Network (CNN). In their investigation, they discovered that estimating age was 57% correct while estimating gender was 82% accurate. Duan *et al.* (2018) introduced a hybrid structure which includes CNN and Extreme Learning Machine to perform age and gender classification. They stated that the accuracy rate in the test results was 52.3% for age prediction and 88.2% for gender prediction. Oladipo *et al.* (2022) developed an age estimations system using genetic algorithm and backpropagation trained artificial neural network. Kumar *et al.* (2022) proposed a study, which is based on Seg-Net based architecture and machine learning algorithms to classify person's gender anda ge from diverse facial photos. Although it is seen that various age and gender estimation studies have been carried out for different application areas, it has been observed that these studies have not been studied sufficiently in the field of forensics.

MATERIAL AND METHOD

Artificial intelligence refers to systems or machines in which the human learning process is mathematically modeled. It imitates human intelligence to perform tasks and can gradually improve itself with the information it collects. Different techniques have emerged with the increase in artificial intelligence studies. Deep learning is one of the machine learning methods used in artificial intelligence studies that allow computers to learn from experience (Kim 2016). When the literature is scanned, it is seen that there are many areas where deep learning is used. Deep learning applications are being developed in various subjects, such as image and video processing, biomedical signal processing, object recognition, robotics, chemistry, finance, search engines, and autonomous vehicle systems (Şeker *et al.* 2017). VGG16 including deep learning, is an artificial neural network that is effective in prediction and classification (Zhu *et al.* 2023). Therefore, the VGG16 method was preferred in this study.

Deep Learning

Machine learning is the scientific study of statistical models for computer systems to perform a specific predefined task without specifying an explicit command or instruction by the user (Bingol *et al.* 2020). A sub-branch of artificial intelligence studies is machine learning. One of the most popular applications of machine learning is image and image recognition. In image recognition applications, machine-like images must be introduced for the machine to learn the image in question. As a result of this learning, the machine can easily distinguish different pictures from each other or detect the common points of similar pictures. VGG16 architecture is effective in computer vision tasks and is used in detail in the modeling of this study.

Many libraries are available in the Python programming language, each suitable for a different purpose. Selecting the appropriate libraries according to the data to be studied increases the accuracy rate. This study used Keras and TensorFlow libraries containing machine learning algorithms. In the study, images belonging to two classes (female and male) were used for gender estimation with the VGG16 model. Images belonging to four classes (child, young, adult, and old) were used for age estimation. In the study, age and gender classification was made with the VGG16 transfer learning model.

Dataset

The photos used in the dataset were taken from (Generated Photos 2022). These photos are artificial faces created by artificial intelligence methods, which are not in reality but are very realistic. As seen in Figure 1, the eye parts of the photographs in the data set were cut to the same dimensions. Each of the images has a resolution of 120x280 pixels. The photographs used in the data set were divided into two classes, 1630 female and 1650 male. This dataset is also divided into four more classes, 1486 of which are children, 1354 are young, 1472 are adults, and 1402 are old. The application is written using the python programming language on the TUBITAK ULAKBIM, High Performance and Grid Computing Center (TRUBA resources).

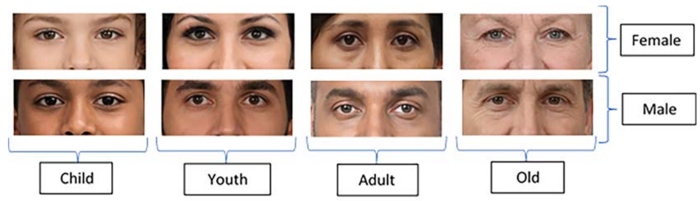

Figure 1 Examples of datasets compiled for the application.

Figure 2 shows the distribution of train, test, and train-test classes by age group. The distribution for train is 26% child, 23.9% youth, 25.5% adult, and 24.7% old. The distribution of each group for train-test is 80% as train and 20% as test.

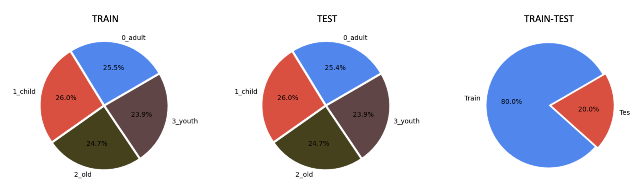

Figure 2 Train data set, test data set, train and test data set graphs for the age group.

In the study, data duplication was applied using the DataImageGenerator Function. The parameters used are shown in Table 1.

Regarding gender, Figure 3 shows the distribution of train, test, and train-test classes, respectively. For train, it is 49.7% female and 50.3%, male. It was determined as 50% female and 50% male for the test. For the train test, 80% was reserved as train and 20% as test.

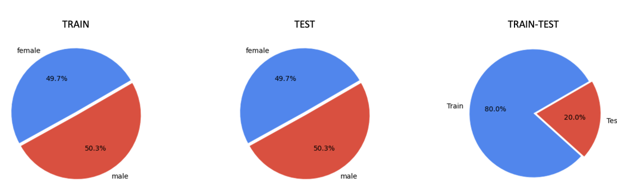

Figure 3 Graphs of train dataset, test dataset, train and test dataset for Gender.

Convolutional Neural Network

CNN, a deep learning algorithm, uses images as input data. It performs the classification process of images. CNN architecture consists of three layers. Convolutional Layer, Pooling Layer, and Fully Connected Layer (Bulut 2017). The features such as edge

Table 1 Parameters of ImageDataGenerator Function.

Parameter	Value
Rotation_range	10
Zoom_range	0.1
Width_shift_range	0.1
Height_shift_range	0.1

Figure 4 VGG16 model.

Table 2 Parameters of Model.fit_generator Function.

Parameter	Value
Batch_size	8
Epochs	15
Steps_per_epoch	8
Verbose	True

Table 3 Gender Confusion Matrix.

Accuracy : 95.73%	True Female	True Male	Total	Class Precision
Predicted Felame	230 (TP)	6 (FP)	236	94%
Predicted Male	15 (FN)	241 (TN)	256	98%
Total	245	247	492	96%
Class Recall	97%	94%	96%	

Table 4 Age Group Confusion Matrix.

Accuracy : 99.41%	True Adult	True Child	True Old	True Youth	Total	Class Precision
Predicted Adult	216	0	0	1	217	100%
Predicted Child	0	221	0	4	225	100%
Predicted Old	0	0	211	0	211	100%
Predicted Youth	0	0	0	198	198	98%
Total	216	221	211	203	851	99%
Class Recall	100%	98%	100%	100%	99%	

Table 5 Gender and Age Group Confusion Matrix.

	Accuracy	F1-Score	Precision	Recall
Gender	96%	96%	96%	96%
Age Group	99%	99%	99%	99%

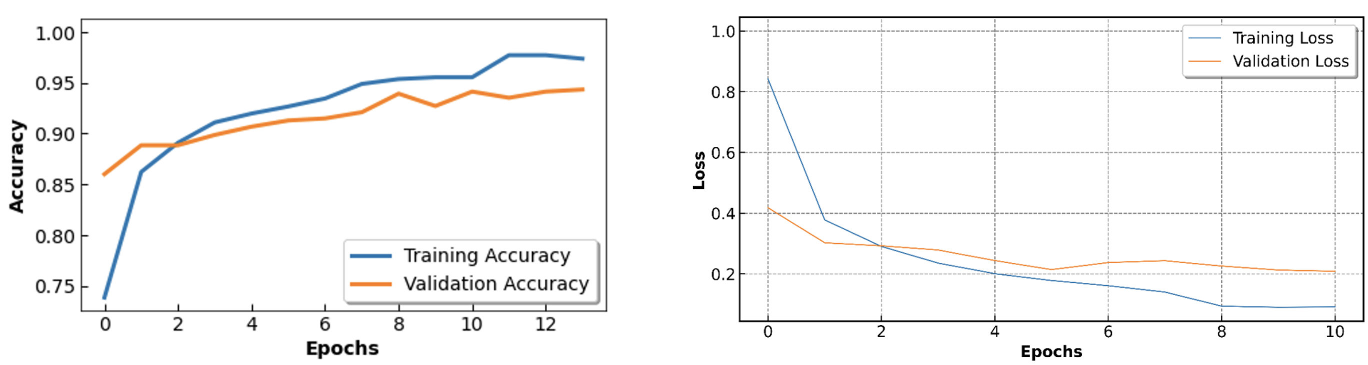

Figure 5 Gender classification success and loss graph.

and texture belonging to the features obtained from the image are found and transmitted to the other sublayers, respectively, and the values in the result layer are obtained (Metlek and Kayaalp 2020).

Convolutional Layer The Convolutional layer is the main layer of the CNN model. CNN extract features automatically in the convolution layer (Gündüz and Cedimoğlu 2019). Attributes are extracted using matrices whose input sizes are determined in the convolution layer (such as 11x11 in AlexNet, 5x5, 3x3, and 2x2 in VggNet). In our study, features were extracted using 2x2 matrices for the age group and 3x3 matrices for gender. With these extracted features, a new matrix was created, and data smaller than the input data were obtained. The matrix to be circulated on the image impacts the network's training and success (Metlek and Kayaalp 2020).

Pooling Layer The pooling layer usually comes after the activation process. In this layer, the data is reduced to smaller sizes. This process makes the network work faster and can lead to data loss. The preferred matrices for data reduction are the maximum (max pooling) value, the smallest (min pooling) value, and the average (average pooling) value matrices. A new matrix is obtained by circulating these matrices over the matrix obtained from the activation process. In our study, the softmax activation formula was used for the age group, and the sigmoid activation formula for gender. The equations of the functions are as follows:

$$softmax : y = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}} \tag{1}$$

$$sigmoid : y = \frac{1}{1+e^{-x}} \tag{2}$$

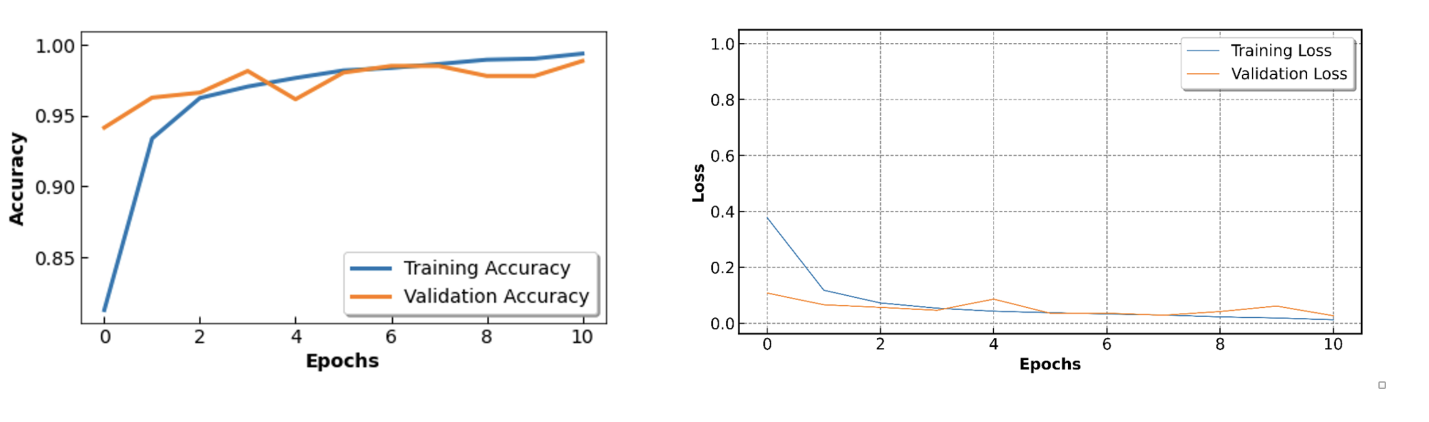

Figure 6 Age group classification success and loss graph.

Fully Connected Layer The fully connected layer is where all the connections in the previous layer are collected. The data from the fully connected layer is transferred to the result layer in one dimension (Metlek and Kayaalp 2020).

VGG16 Model

VGG16 is a deep learning model developed by the Visual Geometry Group at the University of Oxford. VGG16 represents a network architecture called VGGNet, an important milestone in the field of Convolutional Neural Networks. VGG16 is a convolutional neural network with 16 deep layers. These layers consist of convolutional layers, fully connected layers and activation layers. The VGG16 model differs from previous models in that it is deeper and has more parameters. The VGG16 model is specifically designed to be used in image classification tasks (Theckedath and Sedamkar 2020).

The model (Fig.4) was trained on the ImageNet dataset and achieved successful results in recognizing many different object classes. In addition, VGG16 can be used in different tasks with transfer learning methods, often using it as a pre-trained model. VGG16 is considered a milestone in the field of deep learning and convolutional neural networks and is used today as a basic model in many research and applications (Alkurdy *et al.* 2023).

The parameters used in the training of the model in the study are given in Table 2.

Performance Measurement Metrics

In classification models, the success rate is mostly determined by the relationships between the class values labeled by the practitioner and the actual class value (Aslan 2022). Accordingly, the performance is evaluated based on the TP (True Positive), TN (True Negative), FP (False Positive), and FN (False Negative) values in the complexity matrix. In scientific studies, values such as Accuracy, Precision, Recall, and F1 score are generally used for performance evaluation criteria (Bulut 2017). Accuracy is the overall correctness rate.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{3}$$

Precision is the proportion of correctly detected positive classes to all positives.

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

Recall expresses the proportion of correctly identified Positive classes to true positives.

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

The harmonic mean of sensitivity and precision is the F1-score.

$$F1 - score = \frac{2xPrecisionxRecall}{Precision + Recall} \tag{6}$$

RESULTS

The training-validation accuracy and training-validation loss graphs of the VGG16 model are shown in Figure 5, respectively. As seen during the network training phase, the training process was completed with a data loss that could be considered insignificant, with a loss rate of 0.12.

The training-validation accuracy and training-validation loss graphs of the VGG16 model are shown in Figure 6, respectively. As seen during the network training phase, the training process was completed with a data loss that could be considered insignificant, with a loss rate of 0.015.

According to Table 3, the model correctly predicted gender at 96%. TP: true positives, TN: true negatives, FN: false negatives, and FP: false positives. Precision is the proportion of correctly detected positive classes to all positives. Recall expresses the proportion of correctly identified positive classes to true positives.

According to Table 4, Precision: Ratio of correctly detected Positive classes to all positives. Recall: Ratio of correctly detected Positive classes to true positives. Table 5 shows the model's accuracy, F1, precision, and recall values for gender and age group prediction. Accuracy is the overall correctness rate. F1-score: The harmonic mean of sensitivity and precision is the F1-score. Precision is the proportion of correctly detected positive classes to all positives. Recall expresses the proportion of correctly identified Positive classes to true positives.

CONCLUSION

This article it is aimed to estimate the gender and age group of individuals belonging to different gender and age groups, with photographs obtained by cutting the eye parts from the face area. A success rate of 95.73% was achieved in estimating gender and 99.41% in estimating age group. In the study, eye photographs of individuals wearing masks can be used to estimate their gender and age group information. As age and gender determination can be important in revealing victim or suspect profiles, this and similar studies can support forensic processes. When the results obtained in the study were compared with those in the relevant literature, it was seen that more successful results were obtained for age and gender estimation. In future studies, new methods can be developed to enable us to reach faster, more accurate, and more reliable results using different data sets and models.

Conflicts of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgement

The numerical calculations reported in this paper were fully/partially performed at TUBITAK ULAKBIM, High Performance and Grid Computing Center (TRUBA resources).

Availability of data and material

Python code used in the study remains confidential as potential intellectual property.

LITERATURE CITED

Abu Nada, A. M., E. Alajrami, A. A. Al-Saqqa, and S. S. Abu-Naser, 2020 Age and gender prediction and validation through single user images using cnn .

Alkurdy, N. H., H. K. Aljobouri, and Z. K. Wadi, 2023 Ultrasound renal stone diagnosis based on convolutional neural network and vgg16 features. Int J Electr Comput Eng **13**: 3440–3448.

Aslan, M., 2022 Derin öğrenme tabanlı otomatik beyin tümör tespiti. Fırat Üniversitesi Mühendislik Bilimleri Dergisi **34**: 399–407.

Aslan, M. F., K. Sabanci, A. Durdu, and M. F. Unlersen, 2022 Covid-19 diagnosis using state-of-the-art cnn architecture features and bayesian optimization. Computers in Biology and Medicine p. 105244.

Bingol, K., A. E. Akan, H. T. Örmecioğlu, and A. Er, 2020 Artificial intelligence applications in earthquake resistant architectural design: Determination of irregular structural systems with deep learning and imageai method .

Bulut, F., 2017 Örnek tabanlı sınıflandırıcı topluluklarıyla yeni bir klinik karar destek sistemi. Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi **32**.

Dilber, İ. and A. Çetin, 2021 Adli bilişim incelenme süreçlerinde yapay zeka kullanımı: Vgg16 ile görüntü sınıflandırma. Düzce Üniversitesi Bilim ve Teknoloji Dergisi **9**: 1695–1706.

Duan, M., K. Li, C. Yang, and K. Li, 2018 A hybrid deep learning cnn–elm for age and gender classification. Neurocomputing **275**: 448–461.

Generated Photos, 2022 AI Generated Photos. https://generated.photos/faces/datasets.

Gündüz, G. and İ. H. Cedimoğlu, 2019 Derin öğrenme algoritmalarını kullanarak görüntüden cinsiyet tahmini. Sakarya University Journal of Computer and Information Sciences **2**: 9–17.

Hinton, G. E. and R. R. Salakhutdinov, 2006 Reducing the dimensionality of data with neural networks. science **313**: 504–507.

Kim, K. G., 2016 Book review: Deep learning. Healthcare informatics research **22**: 351–354.

Kumar, S., S. Singh, J. Kumar, and K. Prasad, 2022 Age and gender classification using seg-net based architecture and machine learning. Multimedia Tools and Applications **81**: 42285–42308.

Metlek, S. and K. Kayaalp, 2020 Derin öğrenme ve destek vektör makineleri ile görüntüden cinsiyet tahmini. Düzce Üniversitesi Bilim ve Teknoloji Dergisi **8**: 2208–2228.

Oladipo, O., E. O. Omidiora, and V. C. Osamor, 2022 A novel genetic-artificial neural network based age estimation system. Scientific Reports **12**: 19290.

Şeker, A., B. Diri, and H. H. Balık, 2017 Derin öğrenme yöntemleri ve uygulamaları hakkında bir inceleme. Gazi Mühendislik Bilimleri Dergisi **3**: 47–64.

Solmaz, R., A. ALKAN, and M. GÜNAY, 2020 Mobile diagnosis of thyroid based on ensemble classifier. Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi **11**: 915–924.

Theckedath, D. and R. Sedamkar, 2020 Detecting affect states using vgg16, resnet50 and se-resnet50 networks. SN Computer Science **1**: 1–7.

Zha, W., Y. Liu, Y. Wan, R. Luo, D. Li, *et al.*, 2022 Forecasting monthly gas field production based on the cnn-lstm model. Energy p. 124889.

Zhu, F., J. Li, B. Zhu, H. Li, and G. Liu, 2023 Uav remote sensing image stitching via improved vgg16 siamese feature extraction network. Expert Systems with Applications p. 120525.

How to cite this article: Akmese, O.F., Cizmeci, H., Ozdem, S., Ozdemir, F., Deniz, E., Mazman, R., Erdogan, M., and Erdogan, E. Prediction of Gender and Age Period from Periorbital Region with VGG16. *Chaos Theory and Applications*, 5(2), 105-110, 2023.