Age Classification using Different Algorithms of Deep Learning and Computer Vision

Amal Alshahrani[1], Rital Shafiei[2], Dana Alghamdi[3], Rahaf Algethami[4], Shahd Althobaiti [5], and Wafaa Almazmomi[6]

University of Umm Al-Qura, College of Computing, Computer Science and Artificial Intelligent Department, Mecca, KSA

Abstract
The creation of an AI model that reliably classifies age groups based on people's pictures is suggested in this research study. The recommended approach makes use of the YOLO object detection system in combination with the VGG16 Convolutional Neural Network (CNN) architecture. The AI model is trained using a sizable collection of facial photos with matching age categories, allowing it to discover unique patterns and characteristics unique to various age groups. VGG16 and YOLO are integrated into the training process, utilizing their own strengths in object detection and image categorization. The research's conclusions are important for fields including targeted marketing, age-based demographics, and facial recognition. Using VGG16 and YOLO in the AI model to accurately classify users' ages can improve decision-making, targeted advertising, and user experience in general. The objective of the comparative examination of the outcomes produced by various methods is to offer insightful information about their individual capacities.
Keywords:
Age Classification, Convolutional neural network (CNN), YOLOv8, YOLOv5, VGG16.

1. Introduction

This research paper proposes the development of an AI model that accurately categorizes age groups based on individuals' photos. The suggested method involves utilizing the VGG16 Convolutional Neural Network (CNN) architecture in conjunction with the YOLO (You Only Look Once) object detection system. A large dataset comprising images of faces and their corresponding age categories is employed as the training set for the AI model. This enables the model to learn distinctive patterns and features specific to different age groups. The training process encompasses the integration of VGG16 and YOLO, leveraging their respective capabilities in image classification and object detection. The research paper presents a comprehensive outline of the training steps for the AI model. The findings of this research hold significance in areas such as facial recognition, age-based demographics, and targeted marketing strategies. Accurate age classification achieved through the AI model's utilization of VGG16 and YOLO can greatly enhance decision-making, targeted advertising, and the overall user experience. In conclusion, this research suggests the development of a novel AI algorithm that combines the strengths of VGG16 and YOLO to accurately determine the age of individuals based on their photos. The comparative analysis of the results obtained from these approaches aims to provide valuable insights into their respective capabilities and effectiveness in age classification, allowing for informed conclusions regarding their suitability for the task at hand.

2. Literature Review

González-Briones et al. [4] share a study about making a smart picture system to sort out people's gender and age from pictures. The team suggests a fancy system that uses different techniques for taking, changing, and handling pictures. They tried the system in an office and found that using Fisherfaces worked best for guessing age, beating other methods like MLP. Using filters made it easier to work and better at guessing quickly. The paper adds to the field by taking on the problems of age and gender guessing in real places. The authors recommend more research to make the system better and do more things.

Jagzap et al. [5] This article presented a new method for the age classification of people using the Artificial Neural Network (ANN) method. Age classification based on facial images is important in many fields, including computer vision and cognitive sciences. The proposed method combines the Triple Pattern (PTP) method and principal component analysis (PCA) to extract features and then divides the ANN into four categories: child, young, old, and senior. The system strives for originality, demonstrating its capabilities in the fields of biometrics and security applications. Future research will focus on improving the performance of the system and integrating it with additional age classification datasets.

Manuscript received June 5, 2025
Manuscript revised June 20, 2025
https://doi.org/10.22937/IJCSNS.2025.25.6.5

Inik et al. [6] This study aimed to separate human images into different groups using VggNet, one of the deep learning (DL) models. In recent years, we have witnessed the development of artificial intelligence, machine learning, and computer vision, largely thanks to the contributions of DL. DL became popular after the ImageNet 2012 competition, where it significantly reduced the error rate in object recognition. Convolutional Neural Networks (CNNs) form the basis of the DL model, which typically consists of Convolution, reLu, pooling, and full regression. In this study, people were divided into 12 age groups, including gender- and age-specific categories. For this purpose, a new data set consisting of 1800 images was created, 90% of which was used for training and 10% for testing. VggNet was trained using this data and achieved a 78.5% success rate in separating people into different classes. Although large datasets are often required to successfully train the DL model, this study shows that satisfactory training results can be achieved with relatively small datasets. The study highlights the importance of statistical extensions to further improve model performance.

Priyadarshni et al. [7] suggested a way to determine human age using faces. They used K-Nearest Neighbor (K-NN) to classify ages based on facial traits. Their process had three parts: preparing, finding aging features, and classifying. First, they changed colorful pictures to black and white. Then, they found aging traits using skin texture and wrinkles, using a method called the Local Gabor Binary Pattern Histogram (LGBPH). Finally, they used K-NN to quickly classify the data. With the LGBPH method, they got 90% accuracy, showing their method works well for age classification.

Othmani et al. [8] This paper presented a comparative analysis of deep learning frameworks for automatic age estimation from facial images. The study evaluates the performance of various Convolutional Neural Network (CNN) architectures using multiple public datasets, including MORPH, FG-NET, FACES, PubFig, and CASIA-web Face. The results demonstrate that the CNN-based frameworks outperform existing methods in age estimation. The paper also explores the robustness of the best architecture under different conditions, such as noise, expression changes, crossing ethnicity, and crossing gender. Additionally, the study investigates layer-wise transfer learning and knowledge transfer from face recognition tasks to age estimation.

Jiu-Cheng et al. [9] Age prediction is a challenging task due to the discrete nature of age labels. To tackle this issue, researchers have proposed the use of Convolutional Neural Networks (CNNs). These networks utilize two- and three-group classification algorithms and incorporate deep learning and ordinal ensemble techniques. By focusing on specific age groups and utilizing aggregation methods, these approaches aim to improve the accuracy of age predictions. In conclusion, the integration of CNNs with ordinal regression and ensemble learning holds the potential for enhancing age prediction performance in a variety of applications.

Benkaddour [10] demonstrated enhanced performance in age and gender classification using deep learning. where the dataset was divided into training and testing sets, with each person having 10 images for age and gender estimation. The method used involved employing Convolutional Neural Networks (CNNs) to extract features from face images, utilizing activation functions such as ReLU and pooling layers for dimensionality reduction. The results showed a significant improvement in accuracy, with a rate of 91.75%.

Avishek Garain et al. [11] proposed GRA_Net (Gated Residual Attention Network), a deep learning model for accurate age and gender classification from facial images. While there have been previous research efforts in this area, the authors highlight that age and gender identification have received less attention compared to other face recognition tasks. The proposed GRA_Net model is a modified version of the Residual Attention Network, incorporating the concept of gates similar to Gated Residual Units. The model combines classification and regression approaches to tackle the binary gender classification problem and the age prediction regression problem. The authors conducted experiments using five publicly available standard datasets: FG-Net, Wikipedia, AFAD, UTKFace, and AdienceDB. The results obtained from the experiments demonstrate the effectiveness of the GRA_Net model for age and gender classification.

Olatunbosun Agbo-Ajala et al. [12] Advancements in Facial Age Classification: A Comprehensive Survey of Deep Learning Approaches. Traditional handcrafted methods for age estimation have limitations in accurately predicting age. However, the availability of large datasets and computational power has made deep learning with convolutional neural networks (CNNs) a more effective approach. The paper focuses on the utilization of CNNs for age estimation from facial images. The authors analyze various facial aging databases including MORPH-II, IMDb-WIKI, OIU-Adience, CACD, AFAD, WIT-BD, HOIP, Gallagher's web-collected, Ni's web-collected, AgeDB, and UTKface databases, which have been shown to be the most suitable for age estimation using CNN techniques due to their large data size. This paper also discusses age estimation algorithms such as multi-class classification, hybrid algorithms, the ranking algorithm, and deep label distribution learning. The hybrid algorithm combines two or more algorithms, resulting in a better and more robust model. The ranking algorithm addresses the specific problems of classification algorithms, while deep label distribution learning leverages adjacent ages to generate label distributions for each age, leading to improved model performance.

ELKarazle et al. [13] provided an overview of facial age estimation using machine learning techniques. It highlights the process of building automatic age estimation models, including acquiring suitable datasets, preprocessing steps, feature extraction, and model training. The paper also discusses common challenges in age estimation, such as factors influencing aging and data inconsistency.

3. Gap

A significant gap in age classification from AI images is the lack of diverse and comprehensive datasets specifically designed for age estimation. Bridging this gap requires the use of large-scale datasets that encompass a wide range of ages and ethnicities. Such datasets would also enable the training of more accurate and unbiased age estimation models applicable to diverse populations, using a more capable AI model to ensure a higher accuracy percentage. Also, in this research paper, we will be comparing the three tools or software that we are using to build this model, which are YOLOv8, YOLOv5, and VGG16, to see which of them improves the accuracy of the model.

4. Methodology

We implemented the age classification model and adjusted it according to the algorithm we used. As for the dataset we used, the Age Prediction from Images dataset, obtained from Kaggle, forms the basis for training and evaluating age prediction models. It is divided into three age groups (20 to 29, 30 to 39, and 40 to 49), each containing both training and testing samples. Within each age group, there are 200 test samples and 1000 training samples, totaling 600 test samples and 3000 training samples across all age groups. This allocation ensures that 20% of the samples are reserved for testing, while the remaining 80% are utilized for training.

YOLO (You Only Look Once) is an object detection algorithm introduced in 2015. YOLO performs detection in a single pass through a neural network, enabling real-time object detection, not like traditional methods. It's used a lot in applications like autonomous driving and video surveillance. YOLO has evolved with versions like YOLOv2, YOLOv3, and YOLOv4; till now, YOLOv9 is the latest, each improving accuracy and speed. In our research, and taking into consideration that our project is image classification, not object detection, we were able to use YOLOv5 and YOLOv8 to implement the model.

VGG16 is a convolutional neural network used for image classification, image recognition, and object detection tasks. It is characterized by its simplicity and uniform architecture, making it easy to understand and implement. We implemented the age classification model and adjusted it according to the algorithm we used.

As for the dataset we used, the Age Prediction from Images dataset, obtained from Kaggle, forms the basis for training and evaluating age prediction models. It is divided into three age groups (20 to 29, 30 to 39, and 40 to 49), each containing both training and testing samples. Within each age group, there are 200 test samples and 1000 training samples, totaling 600 test samples and 3000 training samples across all age groups. This allocation ensures that 20% of the samples are reserved for testing, while the remaining 80% are utilized for training.

YOLO (You Only Look Once) is an object detection algorithm introduced in 2015. YOLO performs detection in a single pass through a neural network, enabling real-time object detection, not like traditional methods. It's used a lot in applications like autonomous driving and video surveillance. YOLO has evolved with versions like YOLOv2, YOLOv3, and YOLOv4; till now, YOLOv9 is the latest, each improving accuracy and speed.

In our research, taking into consideration that our project is image classification, not object detection, we were able to use YOLOv5 and YOLOv8 to implement the model.

A. YOLOv5

The computer vision model YOLOv5 is a major shift from the Darknet framework to PyTorch. It makes some significant adjustments while maintaining the improvements from YOLOv4 [14]. A strided convolution layer is incorporated into the architecture to lower memory and computational expenses [14]. With five variations (YOLOv5n, YOLOv5s, YOLOv5m, YOLOv51, and YOLOv5x), YOLOv5 can be customized to meet various hardware needs and applications. The breadth and depth of the convolution modules differ in each variation [14].YOLOv5 has a large community of contributors, is regularly maintained by Ultralytics, and is open source. It is renowned for being user-friendly and being simple to implement and train. In addition to several connectors for labeling, training, and deployment, Ultralytics offers a mobile version for the iOS and Android operating systems [14].

B. YOLOv8

To meet the needs of different applications, YOLOv8 is an object detection model that comes in five versions: YOLOv8n (nano), YOLOv8s (small), YOLOv8m (medium), YOLOv81 (big), and YOLOv8x (extra large) [14]. The following are YOLOv8's salient characteristics:

a) Similar Backbone Architecture: With notable changes to the CSPLayer, YOLOv8 maintains a backbone architecture like that of YOLOv5. With this change, detection accuracy is increased by efficiently combining contextual information with high-level features [14].

b) Semantic Segmentation Model: YOLOv8 presents the YOLOv8-Seg semantic segmentation model. Using a C2F module after a CSPDarknet53 feature extractor, it departs from the traditional YOLO neck architecture.Two segmentation heads in YOLOv8-Seg are in charge of predicting semantic segmentation masks for incoming images [14].

YOLOv8 is an effective tool with many uses because of its improvements in architecture, loss functions, and segmentation capabilities [14]. The Darknet framework is connected to the YOLO variants, and as time goes on, more advanced versions such as CSPDarknet are reached [14]. The COCO dataset offers a stable foundation for testing and comparing various architectural ideas, and is frequently used as a benchmark for assessing the performance of YOLO versions [14].

C. VGG16

The typical Convolutional Neural Network (CNN) architecture known for its depth is called VGG, short for Visual Geometry Group. Notable variations of this architecture include VGG-16 and VGG-19, which have 16 and 19 convolutional layers, respectively. It forms the basis of state-of-the-art object recognition models, outperforming baselines on a range of tasks and datasets outside of ImageNet. Designed as a deep neural network, VGGNet is still very well-liked in the image recognition industry [15].

5. Results and discussions

A. YOLO algorithm.

The YOLO algorithm is an object detection and classification algorithm. It uses a deep convolutional neural network (CNN) to classify objects based on their visual features. YOLO divides the input image into a grid and predicts the probability distribution of different classes for each grid cell. The highest probability class is assigned as the label for each detected object [16].

1. YOLOv5

YOLOv5 is a popular object detection algorithm that uses a deep neural network to recognize and locate objects in images. By training the model for more epochs, it had more opportunities to learn and refine its understanding of the objects and their distinguishing features.

The comparison between 10 epochs and 20 epochs using YOLOv5 for object detection and age classification showed that increasing the number of training epochs resulted in higher accuracy. The model demonstrated improved performance in detecting objects and accurately classifying ages, specifically in the age groups of 20, 30, and 40 years old.

Table.1. YOLOv5 10 epochs results

Epoch	accuracy
1	0.334
2	0.334
3	0.348
4	0.334
5	0.321
6	0.321
7	0.314
8	0.291
9	0.331
10	0.304

Table.2. YOLOv5 20 epochs results

Epoch	accuracy
1	0.334
2	0.334
3	0.331
4	0.334
5	0.331
6	0.331
7	0.321
8	0.344
9	0.331
10	0.304
11	0.341
12	0.338
13	0.324

14	0.355
15	0.338
16	0.328
17	0.341
18	0.355
19	0.308
20	0.301

The achieved accuracy of 36% is a quantitative measure that reflects the model's performance. The improvement in accuracy observed in the 20 epochs compared to the 10 epochs suggests that increasing the number of training iterations resulted in better object detection and classification capabilities for the model.

Fig.1. YOLOv5 testing1.

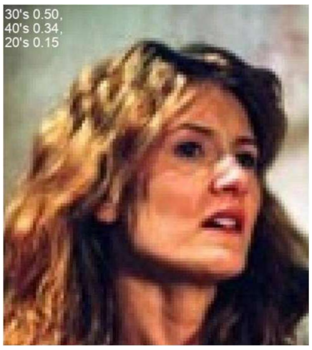

Fig.2. YOLOv5 testing2.

In the figures above, it shows that the model achieved an accuracy of 0.50 on the testing set in YOLOv5.

2. YOLOv8

YOLOv8 is a highly efficient algorithm that excels in fast classification tasks.

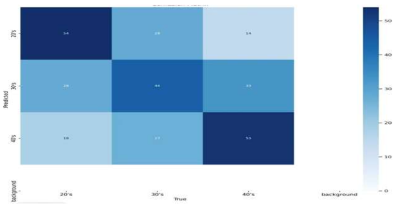
Fig.3. Confusion matrix Normalized.

In YOLOv8, the normalized confusion matrix is a compact representation of the model's performance in classification tasks. It provides a summary of how well the model predicts different classes by measuring the proportion of true positive, true negative, false positive, and false negative predictions. The normalized confusion matrix allows for a quick assessment of the model's accuracy, precision, recall, and F1-score for each class, enabling a comprehensive evaluation of its classification performance. The model used for analysis had 10 epochs, and the accuracy achieved was 51%.

Table.3. YOLOv8 10 epochs

Epoch	accuracy
1	0.387
2	0.434
3	0.468
4	0.441
5	0.438
6	0.474
7	0.498
8	0.498
9	0.494
10	0.505

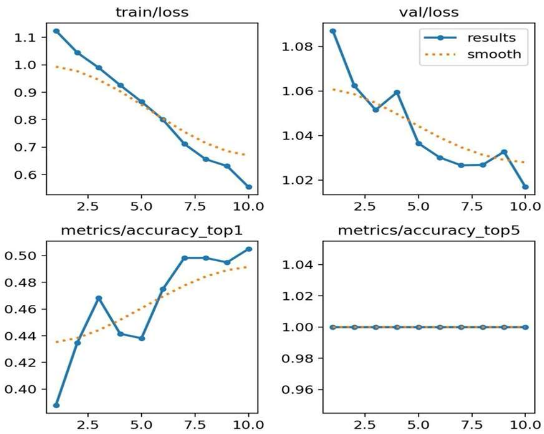

Fig.4. YOLOv8 results

Based on the results analysis, YOLOv8 outperforms YOLOv5 with a higher accuracy of 51%. The primary reason for this improvement in performance is that YOLOv8 incorporates architectural and general model enhancements, resulting in more advanced technical features and improved object detection capabilities.

Fig.5. YOLOv8 testing1.

Fig.6. YOLOv8 testing2.

Figure 6 above shows that the model achieved an accuracy of 0.92 on the testing set in YOLOv8, which is a significant improvement compared to the previous attempts with an accuracy of 0.50.

Here is the hyperparameters for the YOLO versions:

Table.4. YOLO Hyperparameters.

Hyperparameters	Value of YOLOv8	Value of YOLOv5
Input image size	128	128
Epochs	10	20
Optimizer	64	64
Initial learning rate value	AdamW	Adam
Final learning rate value	0.001	0.001
Momentum	0.01	0.01
Weight decay	0.9	0.9
Input image size	0.00005	0.00005

3. VGG16

When accuracy was compared between epochs 10 and 20 in the VGG16 model analysis, it was discovered that epoch 20 had a 35% accuracy increase. In addition, an important finding concerning the training procedure was that, in the first epoch, each step lasted notably longer than in the subsequent epochs.

Table.5. VGG16 10 epochs

Epoch	accuracy
1	0.320
2	0.321
3	0.337
4	0.323
5	0.328
6	0.323
7	0.334
8	0.314
9	0.314
10	0.329

Table.6. VGG16 20 epochs

Epoch	accuracy
1	0.345
2	0.350
3	0.319
4	0.327
5	0.326
6	0.323
7	0.325
8	0.316
9	0.340
10	0.331
11	0.337
12	0.315
13	0.339
14	0.332
15	0.331
16	0.337
17	0.333
18	0.330
19	0.320
20	0.326

One reason for the improved performance of VGG16 at 20 epochs compared to 10 epochs is advanced learning. With the extended training period of 20 epochs, the model could benefit from more training data and deep learning. This allowed the model to understand patterns and information in the data better, enhancing its predictive capabilities. The model could capture more intricate relationships between features and improve its overall performance by training for a longer duration.

Table.7. YOLO and VGG16 results.

Model	Epoch	Accuracy for each class	Best Accuracy
YOLOv5	10	35%	36%
	20	36%	

YOLOv8	10	51%	51%
VGG16	10	34%	35%
	20	35%	

When comparing VGG16 and YOLO, it became evident that YOLO outperformed VGG16 in terms of both performance and accuracy. This could be attributed to several factors, including YOLO's specialized.

Architecture and design for object detection tasks, its unique approach to object detection, potentially superior training data and methodology, as well as optimized implementation choices.

6. Conclusion

The model integrates the VGG16 convolutional neural network architecture and the YOLO object detection system. A large dataset of face images, categorized by age groups, is used for training the model. The combination of VGG16 and YOLO aims to improve the accuracy of age estimation while minimizing training time. The research findings indicate that the YOLO algorithm outperformed the VGG16 algorithm in performance and accuracy, achieving an accuracy rate of 51%. This can be attributed to YOLO's specialized architecture for object detection tasks, unique approach to object detection, and potentially superior training data and methodology. The proposed AI algorithm that combines VGG16 and YOLO shows promise in accurately determining age from photos. Further research and development in this area can advance age estimation technologies and benefit various fields relying on accurate age categorization. We also acknowledge the importance of using a better computer with better resources to achieve a higher accuracy than the one connected in the research paper.

7. Acknowledgment

We would like to thank God one and foremost for finishing this research paper and to all of the people who contributed to writing and working really hard on this project you did a great work, and this would have never happened if it weren't for you. Lastly, we would like to thank our supervisor for always guiding us and helping us through the work.

References

[1] S. A. Rizwan, A. Jalal, M. Gochoo, and K. Kim, "Robust Active Shape Model via Hierarchical Feature Extraction with SFS-Optimized Convolution Neural Network for Invariant Human Age Classification," Electronics, vol. 10, no. 4. MDPI AG, p. 465, Feb. 14, 2021. doi: 10.3390/electronics10040465.

[2] Diwan, T., Anirudh, G., & Tembhurne, J. v. (2023). Object detection using YOLO: challenges, architectural successors, datasets and applications. *Multimedia Tools and Applications*, *82*(6), 9243–9275. https://doi.org/10.1007/s11042-022-13644-y

[3] Girdhar, R., Carreira, J., Doersch, C., & Zisserman, A. (2018). A Better Baseline for AVA (Version 1). arXiv. https://doi.org/10.48550/ARXIV.1807.10066

[4] González-Briones, A., Villarrubia, G., De Paz, J. F., & Corchado, J. M. (2018). A multi-agent system for the classification of gender and age from images. Computer Vision and Image Understanding, 172, 98-106.

[5] Jagzap Shamli, M., & Kendule, J. (n.d.). POSITIONAL TERNARY PATTERN FEATURES BASED HUMAN AGE CLASSIFICATION AND ESTIMATION USING ARTIFICIAL NEURAL NETWORK. Asian Journal of Convergence in Technology, IV. www.asianssr.org

[6] İnik, Ö., & Turan, B. (n.d.). Classification of Different Age Groups of People by Using Deep Learning. JOURNAL OF NEW RESULTS IN SCIENCE. http://dergipark.gov.tr/jnrs

[7] Priyadarshni, V., Nayyar, A., Solanki, A., & Anuragi, A. (2019). Human age classification system using K-NN classifier. In *Advanced Informatics for Computing Research: Third International Conference, ICAICR 2019, Shimla, India, June 15–16, 2019, Revised Selected Papers, Part I 3* (pp. 294-311).

[8] Othmani, A., Taleb, A. R., Abdelkawy, H., & Hadid, A. (2020). Age estimation from faces using deep learning: a comparative analysis. Computer Vision and Image Understanding.

[9] Xie, J. C., & Pun, C. M. (2020). Deep and Ordinal Ensemble Learning for Human Age Estimation from Facial Images. *IEEE Transactions on Information Forensics and Security*, *15*, 2361–2374. https://doi.org/10.1109/TIFS.2020.2965298].

[10] Benkaddour, M. K. (2021). CNN-based features extraction for age estimation and gender classification. *Informatica (Slovenia)*, *45*(5), 697–703. https://doi.org/10.31449/INF.V45I5.3262.

[11] Garain, Avishek, Biswarup Ray, Pawan Kumar Singh, Ali Ahmadian, Norazak Senu, and Ram Sarkar. "GRA_Net: A

deep learning model for classification of age and gender from facial images." IEEE Access 9 (2021): 85672-85689.

[12] Agbo-Ajala, Olatunbosun, and Serestina Viriri. "Deep learning approach for facial age classification: a survey of the state-of-the-art." Artificial Intelligence Review 54, no. 1 (2021): 179-213.

[13] ELKarazle, K., Raman, V., & Then, P. (2022). Facial Age Estimation Using Machine Learning Techniques: An Overview. Big Data Cogn. Comput., 6, 128.

[14] M. Hussain, "YOLOv1 to v8: Unveiling Each Variant–A Comprehensive Review of YOLO," in IEEE Access, vol. 12, pp. 42816-42833, 2024, doi: 10.1109/ACCESS.2024.3378568

[15] Viso AI. (n.d.). VGG Very Deep Convolutional Networks (VGGNet) - What Kou Need to Know. viso.ai. Retrieved from https://viso.ai/deep-learning/vgg-very-deep-convolutional-networks/

[16] Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767.