AL-AASAR Journal
Quarterly Research Journal
www. al-aasar.com

Vol. 2, No. 1 (2025)
Online ISSN: 3006-693X
Print ISSN: 3006-6921

IMAGE CLASSIFICATION AND REGRESSION FOR GENDER AND AGE ESTIMATION

Waina Imamdin[1,*], Jawad Ahmed[1], Iftikhar Hussain[1],Fawad Nasim[1]
[1]Department of Computer Science, Faculty of Computer Science & IT Superior University Lahore, 54000, Pakistan
[*]Corresponding Author: wainaimamdin97@gmail.com

Abstract:
We introduce a deep learning method for automatic age and gender estimates from facial images. Using a dataset fit for both classification and regression tasks, simultaneous prediction of gender (reaching 89% accuracy) and age (achieving 92% classification accuracy) was enabled. This work advances the field by providing a strong deep-learning framework for facial analysis and by pointing out important elements like ideal network settings and data augmentation methods that affect the accuracy of age and gender estimates.

Keywords: Age recognition; Face Identification; Gender recognition; Image Classification; Image Regression

1) Introduction:

In social interactions, two of the most significant aspects of human identity are gender and age. These attributes help shape how people perceive and engage with one another, influencing behavior, communication styles, and even decision-making processes. Age and gender detection have become increasingly relevant, especially in technological fields where understanding human demographics can optimize various processes. Age and gender detection systems have three primary components: facial recognition, gender determination, and age estimation. These three parts work together to make the recognition system. To figure out someone's age, gender, or other demographic information from their face, machine learning methods and deep learning models[1] are used. These techniques have aroused great interest in possible practical applications, including marketing analytics, security systems[2], medical imaging [3], and interfaces between humans and computers. While our work focuses on applying deep learning to age and gender estimation from facial images, deep learning techniques have also proven effective in other image classification domains, such as the classification of malware images, as demonstrated by [4]. However, there are always some problems when the accuracy approaches the peak level, especially in complex live conditions with bright and gloomy illumination, facial movements, and potential occlusions. Currently, age and gender detection systems have benefited from the advancements that deep learning models present[5]. The models employ a large amount of

AL-AASAR Journal
Quarterly Research Journal
www. al-aasar.com
Vol. 2, No. 1 (2025)
Online ISSN: 3006-693X
Print ISSN: 3006-6921

data, composed of faces, to find patterns and the primitives associated with gender and age. Hence, these detection systems have evolved as critical components of marketing analytics devices, security systems, and human-machine interfaces. For instance, by applying these algorithms, companies are likely to identify the gender and age of consumers to develop suitable marketing scenarios. Likewise, security systems may employ facial recognition the same way to identify threats[6] or to improve access control systems. Nonetheless, many issues remain unresolved in this subject, which makes these systems imprecise, even in live imaging where shadowing, facial movements, glasses, hats, etc. are obstacles. The efficiency of detection systems is measured in such conditions, and enhancing detection precision in such conditions is an important area of research. Furthermore, although facial recognition technology has come a long way, it currently has its issues, especially when it comes to minorities, or different age groups, as the training datasets often do not represent the entire population. Lack of adequate data and non-universality are some of the problems associated with single modality in human-computer interactions and systems; multimodality could be useful in solving these problems. It can be achieved by combining data from potentially different sources that may give additional or supplementary information that would be retrievable for some specific uses. The second issue regarding age and gender detection is privacy[7]. Defaulting a person's gender or age without permission was quite ethically questionable. With the availability of surveillance systems with the function of facial recognition, the target audience has become more concerned about the violation of their rights. Concerns over moral issues arising from technical surveillance[8], data gathering by gender identification, human-computer interface GUIs, statistics, and databases for data searches make issues concerning gender identification a moral issue. For instance, a retail company employs cameras to collect customer demographic information. Where there is this type of argumentation, there is always the question of whether customers should be informed of this data collection or if they should have a right to refusal. Many people feel uncomfortable exposing their faces to a facial recognition system for gender and age detection, and the lack of ethics and applicable laws in this regard only proves that more attention needs to be paid to respecting people's privacy while embracing the benefits of, let it be so, technological progress.

In the context of leveraging AI for healthcare applications, recent advances such as those discussed in [9] demonstrate the broader potential of AI-driven methodologies, including frameworks like image classification and regression for demographic analysis, to inform personalized interventions in domains such as mental health. Recent applications of artificial intelligence in important fields, including developments in cybersecurity techniques [10], highlight the adaptability of AI frameworks, including image classification and regression

Vol. 2, No. 1 (2025)
Online ISSN: 3006-693X
Print ISSN: 3006-6921

models, in addressing varied real-world challenges through strong pattern recognition and predictive analytics. Through adaptive AI-driven frameworks, the growing role of artificial intelligence in predictive modeling[11] shows how approaches like image classification and regression—core to demographic analysis in gender and age estimation—may be extended to address more general issues, including threat anticipation and mitigating. Moreover, in using gender detection, several considerations make the task equally difficult because gender is a diverse concept. Gender identification systems commonly extend their categories into two separate divisions in which an individual falls – either male or female. However, this approach does not capture the gender spectrum fully, especially when persons consider themselves as belonging to other gender identities other than being male or female. This introduces further ethical questions about the suitability of such detection systems in a society that is slowly coming to terms with degrees of flexibility of gender. As the study shows, misclassification is potentially harmful, and, therefore, developers and researchers are to resolve these issues in subsequent systems.

As it will be seen, one way of addressing some of the difficulties associated with gender and age detection systems is through multimodal approaches. Systems applied previously, which employ facial recognition only, what is in this case known as the single-modality approach, may prove inefficient in terms of producing accurate results, especially in cases when working with low-quality or acquired in certain conditions database. In contrast, multimodal interfaces use data from multiple modes to acquire supplementary or orthogonal information which could further enhance accuracy. For example, a multimodal system may include voice, gesture, writing, text documents, and face identification. While most current approaches utilize single-modality data to perform demography analysis, these systems can provide more reliable data since they are designed to analyze multiple information streams. For instance, if a face is partly hidden, the potential for voice to suggest whether or not the person is male or female, young or old, could go on to help where face recognition falls short.

The most current advancements, strategies, and challenges in gender and age detection research are thoroughly examined in this paper. We study the fundamentals of deep learning architectures used in cutting-edge detection systems, such as recurrent neural networks (RNNs), convolutional neural networks (CNNs), and others. We also go over the significance of evaluation metrics, preprocessing methods, and dataset quality when comparing the effectiveness of gender and age detection algorithms.

The facial feature extraction method can be broadly categorized into two groups according to the features employed: appearance-based methods and linear feature-based techniques. The distance between different facial features, like the lips, nose, chin, and eyes, is often called a

AL-AASAR Journal
Quarterly Research Journal
www. al-aasar.com
Vol. 2, No. 1 (2025)
Online ISSN: 3006-693X
Print ISSN: 3006-6921

geometric feature. By returning the location for various characteristics, the Viola Junes method may be used to obtain facial characteristics from a facial image. After that, the detected features' Euclidean distance and the French Ar's triangulation technique can be determined. Because it is essential to ensure that facial features are detected accurately, each image in the database was carefully chosen and mined according to appropriate detection techniques. As a result, a billion variables are utilized for training models on numerous samples.

2) Related work:

Here we take a look back at the best methods for estimating age and gender that have been around recently. Review articles on gender classification and studies on face age estimation are available. Furthermore, additional methods are examined since their models provide a graphical representation of CNN's picture categorization decisions. For gender and age estimation, earlier methods used hand-crafted soft-bio metric characteristics. To estimate a person's age and gender, Hayashi et al. collected wrinkle line segments from the corners of their eyes and used them. Using manual methods, Karimi and Tashk were able to acquire a variety of facial traits, including the position of the chin, nose, lips, and eyes. To estimate age and gender, Fukai et al. integrated numerous retrieved facial characteristics, including skin tone, wrinkles, and pigmented patches. By combining SVM classifiers with local picture characteristics, Eidinger et al. created a model. Facial picture appearance variations due to view changes, rotations, and occultations are beyond the capabilities of current methods, which depend on manually created features and their placements. The most up-to-date methods for estimating a person's gender and age use model architectures based on Convolutional Neural Networks (CNNs) trained on raw facial photos. The authors Liao et al. [12] used a Neural Network (NN) architecture to extract classification scores from facial images that had been fine-tuned by dividing them into small patches. Next, they determined a person's age and gender by averaging the categorization scores of these nine patches. The authors Zhang and Xu [13] also suggest a different method, which involves training an NN architecture with landmark patches extracted from face photos. In order to estimate gender and age, Levi and Hassner [14] created a CNN model architecture that is both simple and powerful. Their design can be trained using less data. Abu Nada et al. [15] built a convolutional neural network (CNN) model that can automatically identify the gender and age of a person after it has detected their face. Separate models for estimating age and gender categorization have been proposed by a small number of researchers. A model for age estimate utilizing SVM classifier and Local Binary Patterns (LBP) features has been presented by Shan [16]. In their study, Chen et al. [17] estimated the ages of facial photos using a cascade CNN model and an error-correcting technique. Using a focused loss function, Liu et al. [18] created a CNN model for age classification. By suggesting a CNN model, Mustapha et al. [19] have approached age estimation as a multi-class classification task. The estimation level of CNN-based techniques is far lower than human prediction performances, even though they shown significant performance gains. Also, they don't specify which discriminatory facial features or organs have helped with gender and age estimation; they just use CNN as a black-box classifier. A few methods for graphically explaining the judgments made by different CNN models have been put out in the past few years. By backpropagating the gradient of class activation maps, Selvaraju et al. [20] created a method known as Grad-CAM, which graphically displayed the discriminated features of a class. In order to create class activation maps, they revised their visualization model in [21] to take the second order of gradients into account. A model called LayerCAM was recently presented by Jiang et al. [22] for visualization purposes. This model uses backpropagating positive gradients. If you want to know how CNNs work and find out which traits discriminate between classes, you need to see how they make decisions.

3) Methodology

3.1) Dataset:

In this study, we run our tests on the UTKFace dataset, which contains 20,000 face photos annotated with age, gender, and ethnicity. Being a large-scale face dataset, the data covers a wide age range, from zero to twenty-six years old. This dataset contains over 20,000 pictures of faces annotated with age, gender, and ethnicity. Images' occlusion, lighting, quality, face expressions, and positions can vary greatly. This dataset could be used for a wide variety of applications, including face detection, age estimates, landmark localization, age progression or regression, and many more. But the face detection and age estimation in that dataset are our main concerns. See more examples of UTKFace dataset photos in Fig. 1. Every picture has a three-element tuple that says "age" (in years) and "gender" (Male-0, Female-1).

AL-AASAR Journal
Quarterly *Research Journal*
www. al-aasar.com

Vol. 2, No. 1 (2025)
Online ISSN: 3006-693X
Print ISSN: 3006-6921

Fig 1: Sample Images

To achieve consistent results, We trained, tested, and validated both using the same set of Images. The data sets were divided into 80:20 ratios into the train, test, and validation sets to achieve this.

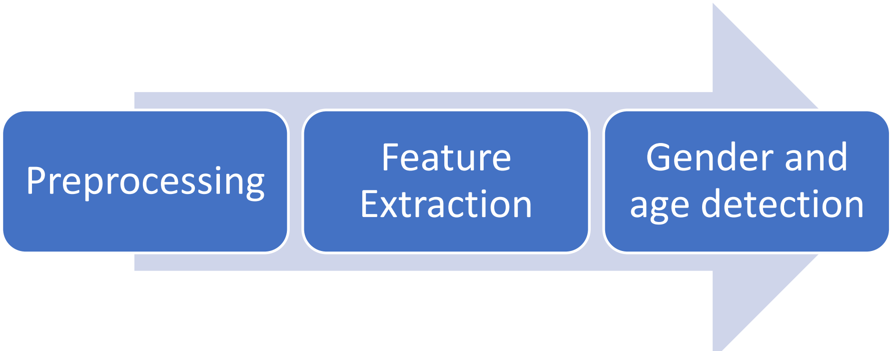

Fig 2: Steps involved in gender classification

.

AL-AASAR Journal
Quarterly Research Journal
www. al-aasar.com

Vol. 2, No. 1 (2025)
Online ISSN: 3006-693X
Print ISSN: 3006-6921

3.2) PlottIng the Age Distribution:

In this study, we plot the age distribution. As shown in Fig. 3, the distribution is roughly normal, although imperfect. It is executed on the right side, but there are outliers at the higher end of the curve. The median we can get after plotting the age distribution is 27.

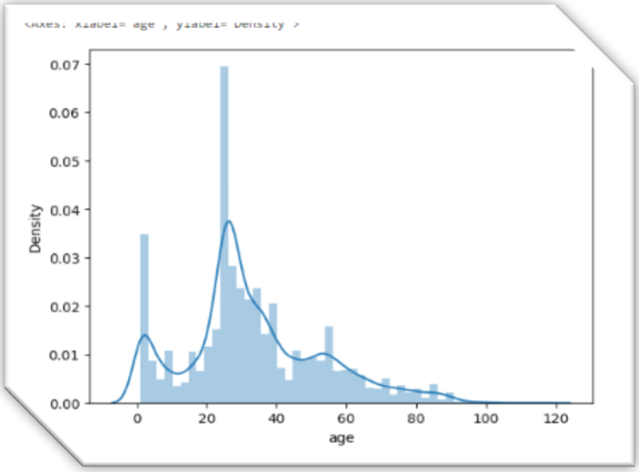

Fig 3: Age distribution

3.3) Methodology:

In addition to describing the system we have suggested, this section offers information on replacing the procedure for further study. Our method includes two primary elements. An estimator of gender, which classifies input photos based on the gender of their faces, makes up the first part. Two VGG16 models comprise the age estimation module, the second part. Model A, the first model, has only been trained on images of female subjects. Model B is the second model, which has only been trained on male subjects. The labels A and B only refer to the models. Our gender classifier is trained using the Kaggle gender dataset , and our age estimation models are tested and trained using the UTKFace. To reduce biases that could occur if we trained our gender estimation model on the age estimation dataset, we opted to train it on a completely different dataset.

Vol. 2, No. 1 (2025)
Online ISSN: 3006-693X
Print ISSN: 3006-6921

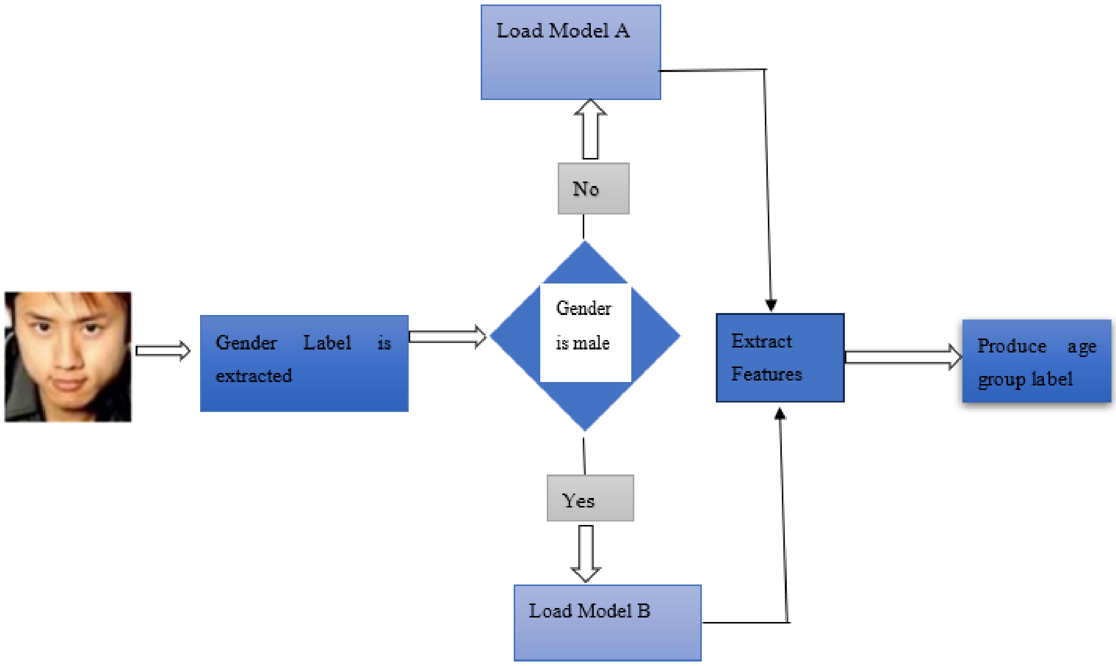

Figure 4: Overview of proposed system

3.4) Face recognition:

The CNN model was selected because it performed better than the other methods (Haar Cascade, Eigenface, and Fisherface) regarding speed, accuracy, CPU-based real-time operation, and detecting faces of different sizes and alignments. First, the image was removed from the dataset and converted to grayscale. The pixels were then fed through the opencv pre-trained model ("haar cascade frontal face alt2.xml"), which produced the corner points for the rectangle where the face was detected using the OpenCV cascade classifier. The face was then placed into a different folder for further processing. After the images were processed through a size filter function that removed outlier images (those smaller than 5 kb in size), the resulting image was used as the dataset for training the model.

3.5) Gender Forecasting:

Vol. 2, No. 1 (2025)
Online ISSN: 3006-693X
Print ISSN: 3006-6921

The output layer of the architecture uses a 'Sigmoid' function, with one node representing Male or Female (0: Male, 1: Female). Gender prediction is thought to be a 'classification issue.' The model used is a four-layer convolution architecture with a max-pooling layer and a convolution layer in each layer. The output of these layers is subsequently subjected to a "real" activation function, succeeded by a dense layer. Thick layers with sigmoid functions are used to create gender-specific nodes. This model was 88.09 percent accurate in terms of precision. The 'Sigmoid' function is used in the output layer of the following model architecture, where two nodes (index 0: Male, index 1: Female) represent the two probabilities for the Male and Female classes, respectively. The model used is a three-layer convolution architecture consisting of a max pooling layer, a batch-normalization layer, and a convolution layer. These layers send their output to a flattening layer, which uses the "real" activation function to send the output to a dense layer. The node with gender probability for both classes is finally generated using a thick layer and a sigmoid function, following the removal of 50% of random nodes.

3.6) Preprocessing the model:

A pre-trained face identification model was used to analyze every picture in the dataset. The OpenCV cascade classifier 'haar cascade frontal face alt2.xml' was employed. After being placed in a new folder, the images were filtered to remove any outliers, such as faces that were misidentified. This was the final picture folder where processing was done. Finally, the photos were hard coded to OpenCV for scanning. The pixels were then parsed into an array with each picture being independent of the other while “Age, Gender, and Ethnicity” were parsed from the filename into a face detection model which was used in analyzing all pictures in that set. 642 p. employed the OpenCV cascade classifier haar cascade frontal face alt2.xml. The images that were downloaded and then copied into a new folder were sorted to eliminate outliers, i.e. typical benchmark faces. The last folder contains the imagery with elements subjected to preprocessing. Ultimately, images were scanned using OpenCV, and after the pixels were extracted into an array, the filename's "age, gender, and ethnicity" were extracted and added to a data frame. This procedure was continued in Step 2. Every age was linked to an age class, and the age array was mapped to various bins. Twenty-four age classes were formed, and the information was forwarded for additional processing.

AL-AASAR Journal
Quarterly Research Journal
www. al-aasar.com

Vol. 2, No. 1 (2025)
Online ISSN: 3006-693X
Print ISSN: 3006-6921

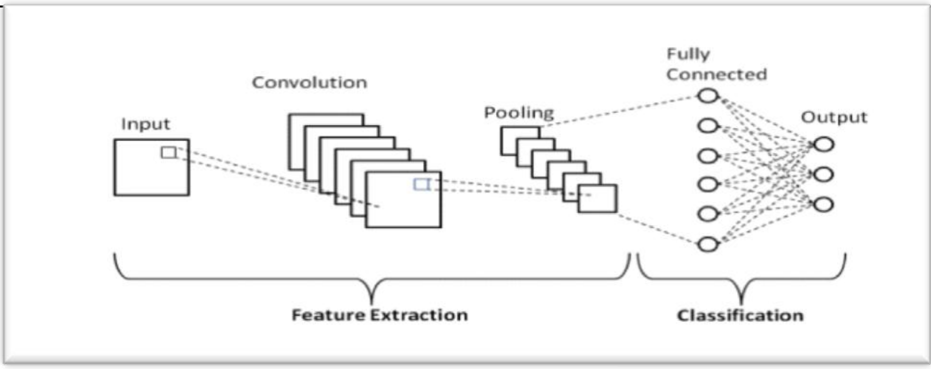

Figure 5: Overview of CNN

3.7) Feature Extraction:

A convolution technique called feature extraction helps to isolate and identify each unique feature of the image so that it can be analyzed. The convolution output is fed into a fully connected layer, which forecasts the class of the image by utilizing the data collected in previous rounds. "Convolutional, pooling, and fully-connected (FC) layers" are the three layers that make up the CNN. A CNN is created when these layers are combined. Two matrices representing photos are multiplied to generate an output from which features can be extracted.

4) Network Architecture:

Age estimation, as well as gender and age classification, are addressed by the deep CNN method. The core components of all three models are a collection of convolutional units after a sequence of Fully connected layers for regression and classification. The model is given an RGB image, which resizes it to 180 x 180 x 3. Convolutional units, or stacks of convolutional layers with a 3x3 filter size, make up each architecture. The next steps are batch normalization, max pooling (2x2), non-linear activation (ReLU), and shift in co-variates. To encourage independence between them, the more profound levels in this instance also have spatial dropout (drop value of 0.15–0.2), which removes entire feature maps. The output is fed into the FC layers after being flattened after the convolutional blocks. Batch normalization, dropout (value between 0.2 & 0.4), and ReLU activation function are features of these FC layers. Fig 5 shows the architecture that was used to calculate age. The output layer of the architectures for age and gender classifications has a softmax activation function, and the convolution layer has 3 & 2 blocks with 256 filter.

Figure 6: Loss

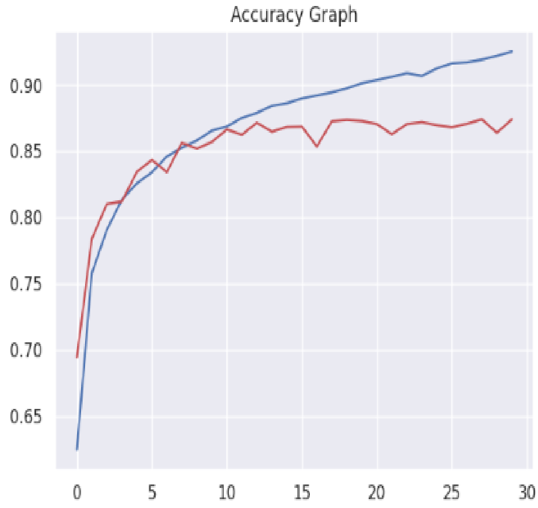

Fig 7: Accuracy

Table 1: Result Summary

Parameter	Value
age_out_loss	**5.7**
Gender_Out_Accuracy	**92.5%**
val_age_out_loss	**6.7**
val_gender_out_accuracy	**89 %**

Conclusion:

This research paper examined very carefully an exhaustive list of methodologies used in gender determination. The fact that accurate identification of gender is required by multiple complex uses made our research take into consideration the varied array of applications. Such uses cross a variety of fields, among them, but not limited to, tracking of individuals for one reason or another, high-end computer forensics investigation, targeted electronic campaigns, thorough statistical analysis on varying populations, and the vital acquisition of dependable demographic information to support sound decision-making.

In order to assess the performance of our suggested gender determination approaches, we created a strong dataset. The dataset was strategically divided into separate subsets to enable the training as well as strict assessment of our models. In particular, 80% of data was reserved for the training phase so that the models could identify the underlying patterns and characteristics specific to various genders. The remaining 20% of the dataset was held back for the test phase, which gave an objective evaluation of how well the models generalize to new data.

The experimental findings derived from our research show a strong level of accuracy in gender categorization. We obtained a remarkable gender output precision of 92% on the held-out test set, proving the model to be well capable of accurately determining gender in unseen cases.

Additionally, the validation age output precision was 89%, implying an equally high level of performance on a distinct validation set applied throughout training for hyperparameter optimization and overfitting avoidance.

Although the performance of the gender classification was good, we also tracked the age prediction loss within our multi-task learning setup. The loss of the output for age in the validation dataset was at 5.7%, which was the rate of error in predicting age. The total validation gender and age output loss was at 6.7%, representing the total error when both gender prediction and age prediction were considered together. These loss metrics provide valuable insights into the model's performance across different prediction tasks and highlight areas for potential future improvement.

Reference:

[1] Khan, Muhammad Ismaeel, Aftab Arif, and Ali Raza A. Khan. "The Most Recent Advances and Uses of AI in Cybersecurity." BULLET: Jurnal Multidisiplin Ilmu 3, no. 4 (2024): 566-578.

[2]Khan, Muhammad Ismaeel, Aftab Arif, and Ali Raza A. Khan. "AI's Revolutionary Role in Cyber Defense and Social Engineering." International Journal of Multidisciplinary Sciences and Arts 3, no. 4 (2024): 57-66.

[3]Zainab, Hira, Ali Raza A. Khan, Muhammad Ismaeel Khan, and Aftab Arif. "Ethical Considerations and Data Privacy Challenges in AI-Powered Healthcare Solutions for Cancer and Cardiovascular Diseases." Global Trends in Science and Technology 1, no. 1 (2025): 63-74.

[4]Tariq, Muhammad Arham, Muhammad Ismaeel Khan, Aftab Arif, Muhammad Aksam Iftikhar, and Ali Raza A. Khan. "Malware Images Visualization and Classification With Parameter Tunned Deep Learning Model." Metallurgical and Materials Engineering 31, no. 2 (2025): 68-73. https://doi.org/10.63278/1336.

[5]Zainab, Hira, Muhammad Ismaeel Khan, Aftab Arif, and Ali Raza A. Khan. "Development of Hybrid AI Models for Real-Time Cancer Diagnostics Using Multi-Modality Imaging (CT, MRI, PET)." Global Journal of Machine Learning and Computing 1, no. 1 (2025): 66-75.

[6]Arif, Aftab, Muhammad Ismaeel Khan, and Ali Raza A. Khan. "An overview of cyber threats generated by AI." International Journal of Multidisciplinary Sciences and Arts 3, no. 4 (2024): 67-76.

[7]Zainab, Hira, Muhammad Ismaeel Khan, Aftab Arif, and Ali Raza A. Khan. "Deep Learning in Precision Nutrition: Tailoring Diet Plans Based on Genetic and Microbiome

Data." Global Journal of Computer Sciences and Artificial Intelligence 1, no. 1 (2025): 31-42.

[8] Khan, Ali Raza A., Muhammad Ismaeel Khan, and Aftab Arif. "AI in Surgical Robotics: Advancing Precision and Minimizing Human Error." Global Journal of Computer Sciences and Artificial Intelligence 1, no. 1 (2025): 17-30.

[9] Zainab, Hira, Ali Raza A. Khan, Muhammad Ismaeel Khan, and Aftab Arif. "Innovative AI Solutions for Mental Health: Bridging Detection and Therapy." Global Journal of Emerging AI and Computing 1, no. 1 (2025): 51-58.

[10] Khan, M. I., A. Arif, and A. R. A. Khan. "AI-Driven Threat Detection: A Brief Overview of AI Techniques in Cybersecurity." BIN: Bulletin of Informatics 2, no. 2 (2024): 248-61.

[11] Arif, A., A. Khan, and M. I. Khan. "Role of AI in Predicting and Mitigating Threats: A Comprehensive Review." JURIHUM: Jurnal Inovasi dan Humaniora 2, no. 3 (2024): 297-311.

[12] Z. Liao, S. Petridis, and M. Pantic, "Local deep neural networks for age and gender classification," arXiv preprint arXiv:08497, 2017.

[13] Y. Zhang and T. Xu, "Landmark-guided local deep neural networks for age and gender classification," Journal of Sensors, vol. 2018, 2018.

[14] G. Levi and T. Hassner, "Age and gender classification using convolutional neural networks," in Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp. 34-42,

[15] A. M. Abu Nada, E. Alajrami, A. A. Al-Saqqa, and S. S. Abu-Naser, "Age and Gender Prediction and Validation Through Single User Images Using CNN," International Journal of Academic Engineering Research, vol. 4, no. 8, pp. 21-24, 2020.

[16] C. Shan, "Learning local features for age estimation on real-life faces," in Proceedings of the 1st ACM international workshop on Multimodal pervasive video analysis, pp. 23-28, 2010.

[17] J.-C. Chen, A. Kumar, R. Ranjan, V. M. Patel, A. Alavi, and R. Chellappa, "A cascaded convolutional neural network for age estimation of unconstrained faces," in IEEE 8th International Conference on Biometrics Theory, Applications, and Systems (BTAS), pp. 1-8, 2016.

[18] W. Liu, L. Chen, and Y. Chen, "Age classification using convolutional neural networks with the multi-class focal loss," in IOP conference series: materials science and engineering, vol. 428, no. 1, p. 012043, 2018.

[19] M. F. Mustapha, N. M. Mohamad, G. Osman, and S. H. Ab Hamid, "Age Group Classification using Convolutional Neural Network (CNN)," in Journal of Physics: Conference Series, vol. 2084, no. 1, pp. 1-20, 2021.

[20] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in Proceedings of the IEEE International Conference on computer vision, pp. 618-626, 2017

[21] A. Chattopadhyay, A. Sarkar, P. Howlader, and V. N. Balasubramanian, "Grad-CAM++: Generalized gradient-based visual explanations for deep convolutional networks," in IEEE Winter conference on applications of computer vision (WACV), pp. 839-847, 2018.

[22] P.-T. Jiang, C.-B. Zhang, Q. Hou, M.-M. Cheng, and Y. Wei, "Layercam: Exploring hierarchical class activation maps for localization," IEEE Transactions on Image Processing, vol. 30, pp. 5875-5888, 2021

[23] V. Raman, K. ELKarazle, and P. Then, "Gender-specific facial age group classification using deep learning," *Intell. Autom. Soft Comput*, vol. 34, pp. 105–118, 2022.