International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 02 | Feb - 2025 SJIF Rating: 8.448 ISSN: 2582-3930

Emotion Based Age and Gender Detection Using ML

Sahil Tarle[1], Mayur Vanve[2], Ritesh Rayate[3] Shreyas Shete[4]

[1]Sahil Tarle Department Information Technology From Matoshri Aasarabai Polytechnic.
[2] Mayur Vanve Department Information Technology From Matoshri Aasarabai Polytechnic.
[3] Ritesh Rayate Department Information Technology From Matoshri Aasarabai Polytechnic
[4]Shereyas Shete Department Information Technology From Matoshri Aasarabai Polytechnic.
[5] Ms.Ashwini Gaikwad Lecturer Of Information Technology From Matoshri Aasarabai Polytechnic.
[6] Mr.Mahesh Bhandakkar HOD of Information Technology From Matoshri Aasarabai Polytechnic.

Abstract -In recent years, emotion-based age and gender detection using machine learning has gained significant attention due to its potential applications in various domains, including marketing, security, and human-computer interaction. This study explores the effectiveness of convolutional neural networks (CNNs) in accurately classifying age and gender based on facial expressions. Specifically, we leverage advanced architectures such as VGG16 and ResNet, which are renowned for their deep learning capabilities. By utilizing transfer learning techniques, we enhance the feature extraction process from facial images, thereby improving the model's performance on age and gender classification tasks.Our experiments demonstrate that integrating emotional analysis with demographic detection can yield more personalized and context-aware interactions.we present a comparative analysis of the performance of VGG16 and ResNet models in this domain. The results indicate that while both models achieve commendable accuracy, ResNet's residual connections significantly mitigate the vanishing gradient problem, leading to superior performance in recognizing nuanced emotional expressions across different age groups and genders. Additionally, we discuss the implications of these findings in real-world applications, emphasizing the importance of ethical considerations when deploying emotion detection systems. This study contributes to the growing body of research in affective computing and provides a foundation for future advancements in emotion-based demographic detection technologies.

Keywords: Emotion detection, Age classification, Gender classification, Machine learning, Convolutional neural networks, VGG16, ResNet, Transfer learning, Affective computing

1.INTRODUCTION

Emotion-based age and gender detection is an emerging field that leverages advancements in machine learning (ML) and computer vision to analyze human emotions through facial expressions. With the proliferation of digital technologies and the increasing amount of visual data generated every day, understanding emotions has become crucial for various applications, such as targeted marketing, user experience enhancement, and security systems. Accurately detecting age and gender based on emotional cues can lead to more personalized interactions in digital environments, enabling businesses and service providers to tailor their offerings effectively.Consequently, this research aims to develop a robust framework for emotion-based age and gender detection leveraging deep learning techniques. Convolutional Neural Networks (CNNs) have emerged as a fundamental tool in image analysis, renowned for their capability to automatically learn spatial feature hierarchies from input images.By employing multiple layers of convolutions, CNNs can capture intricate details and patterns that are crucial for emotion recognition. Among the various architectures available, VGG16 and ResNet stand out for their performance in classification tasks. VGG16 is known for its simplicity and depth, utilizing small convolutional filters to extract fine-grained features, while ResNet introduces residual learning, allowing for the training of deeper networks without suffering from degradation in performance. These characteristics make both architectures suitable for analyzing facial images to detect emotional states, age, and gender effectively.The integration of emotion detection with age and gender classification enhances the understanding of user behavior and preferences. For instance, in marketing applications, brands can tailor advertisements based on the emotional responses of different demographic groups, ensuring that content resonates with the intended audience. In security and surveillance systems, accurately identifying individuals' emotional states can help assess potential threats or disturbances in public spaces. Thus, this research not only explores the technical aspects of emotion-based classification but also highlights its potential impact across various industries.In this study, we implement a comprehensive evaluation of VGG16 and ResNet architectures for emotion-based age and gender detection. We employ machine learning techniques to leverage pre-trained models on large datasets, enhancing their performance on specific tasks while minimizing the need for extensive labeled data. Our approach involves collecting a diverse dataset of facial images annotated with emotional labels, age groups, and gender information. By systematically comparing the results obtained from both architectures, we aim to identify the most effective model for this dual-task classification problem, contributing to the field of affective computing and its applications in real-world scenarios.

2. Problem Statement

The problem addressed by this project is the accurate detection of a person's age, gender, and emotional state based on facial images using machine learning and deep learning algorithms, specifically CNN, VGG16, and ResNet. Existing systems often struggle with the complexity of predicting multiple attributes, especially when considering the variability in human facial expressions, diverse age groups, and genders. Furthermore, real-time emotion recognition poses additional challenges due to dynamic facial expressions

and environmental conditions, such as lighting and occlusion. This project aims to build a robust solution that can predict age and gender while simultaneously identifying emotions using advanced neural network architectures. The key challenge lies in training the models to generalize well across diverse datasets and ensuring the system can perform accurately in real-world scenarios. By integrating CNN, VGG16, and ResNet, the model is expected to achieve high precision in extracting relevant facial features for accurate predictions across these three dimensions.

3. Literature Review

1) In their study, "Face gender recognition in the wild: an extensive performance comparison of deep-learned hand-crafted and fused features with deep and traditional models," Althnian et al. (2021) explore various approaches to gender recognition from facial images in real-world conditions. The authors compare the effectiveness of deep learning methods against traditional models and hand-crafted features, highlighting the merits of deep learning, such as superior accuracy and robustness to variations in lighting and occlusion. Conversely, they point out the demerits of deep learning, including the requirement for large datasets and significant computational resources. The findings underscore the importance of selecting appropriate feature extraction techniques and model architectures, suggesting that hybrid approaches that fuse deep-learned and hand-crafted features can yield optimal performance in gender recognition tasks.[1]
2) In their paper, "A Sequential Iterative Detection Framework for Gender and Age Classification," Altun and Aksoy (2021) present a novel approach for classifying gender and age using a sequential iterative detection framework. The study demonstrates the merits of this framework, which improves classification accuracy by leveraging multiple iterations of detection to refine results progressively. The authors also highlight the robustness of their approach against variations in image quality and demographics, making it adaptable to diverse datasets. However, the paper notes some demerits, including potential increased computational complexity and time consumption due to the iterative process, which may hinder real-time applications. Overall, the framework shows promise for enhancing gender and age classification tasks while emphasizing the need for efficient implementation in practical scenarios.[2]
3) In the paper "Prediction of the age and gender based on human face images based on deep learning algorithm," Haseena et al. (2022) explore the application of deep learning techniques for predicting age and gender from facial images. The authors highlight the merits of utilizing advanced convolutional neural networks (CNNs), which significantly enhance prediction accuracy and efficiency compared to traditional methods. Their research demonstrates how deep learning can effectively extract intricate features from images, resulting in robust performance across diverse datasets. However, the paper also addresses certain demerits, such as the need for extensive labeled training data and the potential for overfitting, particularly when models are not properly regularized. Overall, the findings suggest that while deep learning algorithms offer substantial advantages in age and gender prediction, careful consideration of data quality and model complexity is crucial for practical implementation.[3]
4) In the paper "An Approach Based on Deep Learning for Recognizing Emotion, Gender and Age," Juel Sikder (2022) investigates a deep learning framework designed to simultaneously classify emotion, gender, and age from facial expressions. The study highlights the merits of this integrated approach, demonstrating that deep learning models can achieve high accuracy and reliability in recognizing complex human attributes by effectively capturing nuanced facial features. Additionally, the research emphasizes the framework's potential for real-time applications, enhancing its utility in various fields such as marketing and security. However, the paper also discusses certain demerits, including the significant computational resources required for training deep learning models and the challenges posed by diverse lighting conditions and facial occlusions that may affect recognition performance. Overall, the findings suggest that while deep learning offers promising advancements in multi-faceted facial analysis, addressing its limitations is essential for broader applicability.[4]
5) In their paper "Age, Gender Prediction and Emotion Recognition Using Convolutional Neural Network," Singh et al. (2021) present a convolutional neural network (CNN)-based approach to simultaneously predict age, gender, and recognize emotions from facial images. The authors highlight the merits of using CNNs, which allow for effective feature extraction and significantly improve classification accuracy compared to traditional machine learning methods. Their research demonstrates strong performance across multiple datasets, indicating the robustness of their model in diverse scenarios. However, the study also notes some demerits, such as the requirement for large and well-annotated training datasets, which can be challenging to obtain. Additionally, the authors mention the model's vulnerability to variations in image quality and environmental conditions that could affect recognition accuracy. Overall, the findings suggest that while CNNs are powerful tools for multi-task facial analysis, addressing data-related challenges and enhancing model generalization are critical for practical applications.[5]

6) In the paper "Age, Gender, and Emotion Recognition Based Deep Learning Models," M. N. J. et al. (2021) explore various deep learning models tailored for the simultaneous recognition of age, gender, and emotions from facial images. The authors emphasize the merits of utilizing advanced architectures, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), which enhance prediction accuracy by effectively capturing complex patterns in facial features. Their findings indicate that the proposed models perform well across different datasets, demonstrating robustness in varied conditions. However, the study also highlights several demerits, including the necessity for extensive labeled data to train the models adequately and the potential for overfitting, especially with smaller datasets. Additionally, the models' performance can be adversely affected by variations in facial expressions, lighting, and occlusions. Overall, the research suggests that while deep learning models show significant promise in multi-faceted facial analysis, addressing data limitations and environmental influences is essential for achieving optimal performance.[6]
7) In the paper "Machine Learning Based Approaches for Age and Gender Prediction from Tweets," Katna et al. (2022) investigate the use of machine learning techniques to predict users' age and gender based on the content of their tweets. The authors highlight the merits of their approach, demonstrating that linguistic patterns and stylistic features can serve as effective indicators for demographic predictions, achieving reasonable accuracy even in the absence of explicit demographic information. The study showcases the versatility of machine learning algorithms in text analysis, emphasizing their potential in social media analytics and targeted marketing. However, the paper also discusses several demerits, such as the inherent challenges in dealing with noisy and unstructured data typical of social media platforms, which may affect model performance. Additionally, the models may face limitations in generalizability across different demographic groups and cultural contexts. Overall, the findings suggest that while machine learning offers valuable insights for age and gender prediction from tweets, addressing data quality and contextual variations is crucial for enhancing reliability and applicability.[7]
8) In the paper "Age and Gender Prediction from Face Images Using Attentional Convolutional Network," Abdolrashidi et al. (2020) propose an attentional convolutional network (ACN) designed to enhance age and gender prediction from facial images. The authors highlight the merits of integrating attention mechanisms, which allow the model to focus on relevant features in the face, thereby improving prediction accuracy and robustness against variations in lighting and facial expressions. Their approach demonstrates superior performance compared to traditional CNNs, indicating the effectiveness of attention in feature extraction. However, the study also points out several demerits, including the increased computational complexity and the requirement for extensive labeled datasets for training, which may limit its applicability in real-world scenarios. Additionally, the model's reliance on facial images raises concerns about privacy and ethical implications in its deployment. Overall, the findings suggest that while attentional convolutional networks offer promising advancements in age and gender prediction tasks, careful consideration of data requirements and ethical issues is essential for practical implementation.[8]
9) In the paper "A Convolutional Neural Network for Real-Time Face Detection and Emotion & Gender Classification," Uddin et al. (2020) present a convolutional neural network (CNN) framework aimed at real-time face detection alongside emotion and gender classification. The authors highlight several merits of their approach, including the ability to process images rapidly, making it suitable for real-time applications in various fields such as security and human-computer interaction. Their model achieves commendable accuracy in detecting faces and classifying emotions and gender, showcasing the effectiveness of CNNs in handling complex visual tasks. However, the study also discusses some demerits, such as the dependency on high-quality datasets for training, which can be challenging to curate. Furthermore, the model's performance may be affected by variations in lighting conditions, facial occlusions, and diverse ethnic backgrounds, which can introduce bias in classification results. Overall, the research indicates that while the proposed CNN framework offers significant advancements in real-time face detection and classification, addressing data quality and environmental challenges is crucial for its successful deployment.[9]
10) In the paper "Face-Based Age and Gender Estimation Using Improved Convolutional Neural Network Approach," Sharma, Sharma, and Jindal (2022) introduce an enhanced convolutional neural network (CNN) model designed specifically for age and gender estimation from facial images. The authors highlight the merits of their improved architecture, which incorporates advanced techniques such as data augmentation and transfer learning, leading to significant improvements in prediction accuracy and model robustness across diverse datasets. The study demonstrates the effectiveness of their approach in mitigating common challenges associated with age and gender classification, such as variations in facial expression and lighting conditions. However, the paper also notes certain demerits, including the requirement for substantial computational resources and the potential for overfitting if the model is not properly regularized. Additionally, the reliance on high-quality labeled datasets may pose challenges in real-

world applications where such data is scarce. Overall, the findings suggest that while the proposed CNN approach offers notable advancements in face-based age and gender estimation, careful attention to model training and data quality is essential for practical implementation.[10]

4. Proposed System

The proposed methodology for emotion-based age and gender detection utilizes a hybrid deep learning approach, integrating Convolutional Neural Networks (CNN), VGG16, and ResNet to create a robust and efficient system capable of accurately predicting age, gender, and emotional states from facial images. The process begins with data collection, where a diverse dataset of labeled facial images is acquired, encompassing various ages, genders, and emotions. Preprocessing techniques such as resizing, normalization, and data augmentation are applied to enhance image quality and model robustness. The CNN architecture is employed for initial feature extraction, where a series of convolutional and pooling layers learn spatial hierarchies of features from the input images, capturing both low-level and high-level characteristics essential for distinguishing between different age groups, genders, and emotional expressions.Following feature extraction, the system leverages pretrained models VGG16 and ResNet to enhance performance further. VGG16 is used for its ability to capture intricate details in facial features, while ResNet, with its residual connections, enables training of very deep networks without the vanishing gradient problem. This multi-task learning approach allows the system to simultaneously predict age, gender, and emotion, utilizing shared features for better accuracy and efficiency. The model is trained using a combination of loss functions tailored for each task, optimizing performance through methods such as Adam or SGD. Finally, thorough evaluation and testing are conducted on a separate dataset to ensure the model's generalizability and robustness in real-world scenarios, aiming for high accuracy and real-time performance in applications such as personalized marketing, healthcare, and interactive systems.

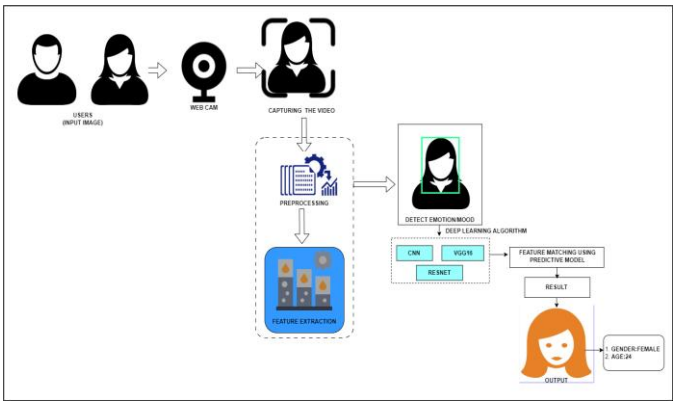

Fig -1: System Architecture

The system architecture for emotion-based age and gender detection begins with the integration of a web camera that captures real-time facial images of individuals. The captured images are then preprocessed to ensure they are suitable for analysis. This preprocessing stage includes resizing the images to a standard resolution, normalizing pixel values for consistent input, and applying data augmentation techniques to introduce variability and enhance the model's robustness. Once the images are prepared, they are fed into a Convolutional Neural Network (CNN) model, which acts as the initial feature extractor. The CNN layers process the images through a series of convolutional and pooling operations, effectively capturing essential facial features such as textures, edges, and other distinguishing characteristics relevant to age and gender classification.Following feature extraction, the processed data is passed to more advanced architectures, specifically VGG16 and ResNet, which further refine the features and enhance predictive accuracy. VGG16's depth allows for the capture of intricate details related to age and gender, while ResNet's residual connections help maintain performance even in deeper architectures. The outputs from these models are directed to separate fully connected layers that produce predictions for age, gender, and emotional state. The results are then displayed in real-time on a user interface, providing immediate feedback on the detected age and gender of the individual captured by the web camera. This architecture not only facilitates accurate age and gender detection but also supports real-time applications in various fields, such as personalized marketing, healthcare monitoring, and interactive systems.

5. Algorithm Used for Proposed System

1) CNN:

The Convolutional Neural Network (CNN) algorithm is a crucial component for emotion-based age and gender detection using machine learning due to its powerful ability to automatically extract and learn hierarchical features from images. CNNs are specifically designed to handle image data, making them ideal for analyzing facial features that are key to predicting age, gender, and emotional expressions. In this context, the CNN algorithm works by applying multiple convolutional layers that learn spatial hierarchies of features such as edges, textures, and shapes in a person's face, which are critical for distinguishing between different age groups, genders, and emotional states.CNNs are composed of convolutional layers, pooling layers, and fully connected layers, which allow them to process raw pixel data and generate high-level feature maps. By using activation functions like ReLU and applying pooling operations, the CNN progressively reduces the dimensionality of the input while retaining essential information. This process helps the network focus on the most relevant features for classification.In emotion-based age and gender detection, CNNs can learn to differentiate subtle variations in facial features, such as wrinkles for age estimation, or distinct facial structures for gender classification. Additionally, CNNs are adept at capturing the nuances of facial expressions, enabling the detection of emotions like happiness, sadness, or anger. The use of CNNs in this task improves the accuracy of predictions by effectively handling challenges like lighting variations, facial pose

changes, and occlusions, making them an essential algorithm for real-time and robust performance in multi-task facial analysis.

2) VGG16:
VGG16 is a deep learning model that has proven to be highly effective for emotion-based age and gender detection using machine learning. Known for its simplicity and powerful performance, VGG16 consists of 16 layers, mainly made up of convolutional and fully connected layers, designed to capture rich and hierarchical feature representations from images. The model's strength lies in its deep architecture, which allows it to extract increasingly complex features from facial images as data passes through successive layers. This makes it highly suitable for tasks such as age and gender prediction, as well as emotion detection.In the context of emotion-based age and gender detection, VGG16's convolutional layers are adept at identifying important facial features like skin texture, facial structure, and expression changes. For age prediction, VGG16 can capture subtle patterns such as wrinkles, skin elasticity, and facial contours that vary with age. For gender classification, it recognizes distinctive facial structures, while for emotion detection, the model learns to detect facial expressions like smiles, frowns, or raised eyebrows.What makes VGG16 particularly useful in this application is its ability to generalize well across diverse datasets, despite its relatively simple structure compared to more complex models like ResNet. By using small receptive fields in its convolutional layers, VGG16 can capture fine-grained details from facial images, which are essential for accurate detection of emotions, age, and gender. Its performance is further enhanced when pretrained on large image datasets like ImageNet, allowing for faster training and better feature extraction in emotion-based age and gender detection models.VGG16 contributes to a highly accurate and robust solution for multi-task facial analysis, ensuring reliable predictions even in complex real-world scenarios.

3) ResNet:
The ResNet (Residual Network) algorithm is highly effective for emotion-based age and gender detection using machine learning due to its ability to train very deep neural networks without suffering from the problem of vanishing gradients, which often occurs in deep learning models. ResNet achieves this by introducing "residual learning" through shortcut connections, allowing the model to skip one or more layers. This innovation enables the training of much deeper networks, which are crucial for capturing complex and subtle patterns in facial images needed for accurate age, gender, and emotion detection.In the context of this task, ResNet's deep architecture allows it to learn more detailed and refined features from facial images, which are essential for distinguishing between different age groups, genders, and emotional expressions. For example, age prediction requires the detection of fine-grained features such as skin texture, wrinkles, and facial contours, while gender classification focuses on structural facial differences. Emotion detection, on the other hand, involves recognizing subtle shifts in facial expressions, such as frowns or smiles.ResNet's skip connections make it particularly well-suited for this multi-task problem, as it ensures that the model continues to learn effectively even as it becomes deeper, enhancing both accuracy and performance. Additionally, ResNet is robust in handling challenges like image noise, occlusions, and variations in lighting, which are common issues in real-world facial data. By using ResNet, the model can achieve high accuracy in emotion-based age and gender detection, making it an ideal choice for this application.

6. Applications

1) Airports
2) Hospitals
3) Offices
4) Metro Stations
5) Schools
6) Banks
7) Residential Areas
8) Marketing
9) Healthcare
10) Security
11) Human-Computer Interaction.

7. CONCLUSIONS

In conclusion, emotion-based age and gender detection utilizing machine learning algorithms such as CNN, VGG16, and ResNet presents a significant advancement in understanding human emotions and demographics through facial recognition technology. The integration of these deep learning models allows for high accuracy and robustness in predicting age, gender, and emotional states from facial images, making the system valuable for various applications, including personalized marketing, healthcare, and human-computer interaction. Despite the challenges related to data quality, privacy concerns, and potential biases, ongoing advancements in technology and methodologies promise to enhance the effectiveness and ethical deployment of these systems. As the field continues to evolve, the ability to analyze emotional expressions and demographic information will not only improve user experiences but also foster greater insights into human behavior, paving the way for innovative solutions across multiple industries.

REFERENCES

1. Althnian, N. Aloboud, N. Alkharashi, F. Alduwaish, M. Alrshoud and H. Kurdi, "Face gender recognition in the wild: an extensive performance comparison of deep-learned hand-crafted and fused features with deep and traditional models", Appl. Sci., vol. 11, no. 1, pp. 89, 2021.

2. H. Altun and H. Aksoy, "A Sequential Iterative Detection Framework for Gender and Age Classification", 2021 Innovations in Intelligent Systems and Applications Conference (ASYU), pp. 1-6, 2021.
3. Haseena, S., et al. "Prediction of the age and gender based on human face images based on deep learning algorithm." Computational and Mathematical Methods in Medicine 2022 (2022).
4. Juel, Juel Sikder. "An Approach Based on Deep Learning for Recognizing Emotion, Gender and Age." (2022).
5. Singh, Arjun, et al. "Age, gender prediction and emotion recognition using convolutional neural network." Proceedings of the International Conference on Innovative Computing & Communication (ICICC). 2021.
6. M. N. J. . et. al., "Age, Gender, and Emotion Recognition based Deep learning models", TURCOMAT, vol. 12, no. 13, pp. 2723–2731, Jun. 2021.
7. Katna, Rishabh, et al. "Machine learning based approaches for age and gender prediction from tweets." Multimedia Tools and Applications 81.19 (2022): 27799-27817.
8. Abdolrashidi, Amirali, et al. "Age and gender prediction from face images using attentional convolutional network." arXiv preprint arXiv:2010.03791 (2020).
9. Uddin, Md Jashim, et al. "A convolutional neural network for real-time face detection and emotion & gender classification." IOSR Journal of Electronics and Communication Engineering (IOSR-JECE) e-ISSN (2020): 2278-2834.
10. Sharma, Neha, Reecha Sharma, and Neeru Jindal. "Face-based age and gender estimation using improved convolutional neural network approach." Wireless Personal Communications 124.4 (2022): 3035-3054.
11. Mellouk, Wafa, and Wahida Handouzi. "Facial emotion recognition using deep learning: review and insights." Procedia Computer Science 175 (2020): 689-694.
12. Adenowo, Adetokunbo A., and Adekunle Olu Adewole. "Performance Evaluation of Deep Learning Method for Age, Gender and Emotion Recognition Tasks." African Journal of Computing & ICT 13.4 (2020).
13. Agbo-Ajala, Olatunbosun, and Serestina Viriri. "Deeply learned classifiers for age and gender predictions of unfiltered faces." The Scientific World Journal 2020 (2020).
14. jyoti Sharma, Dibya, Sachin Dutta, and Dibya jyoti Bora. "REGA: Real-Time Emotion, Gender, Age Detection Using CNN—A Review." (2020).