

Advancements in Facial Gender Recognition Using Convolutional Neural Networks: A Review

¹Shiv Govind , ²Arun Kumar Rai

¹M.Tech Scholar, ²Assistant Professor,

¹Department of Computer Science and Engineering, Veda Institute of Technology, Bhopal (M.P)

²Department of Computer Science and Engineering, Veda Institute of Technology, Bhopal (M.P)

Email:- shivgovindvns@gmail.com, raiaruniitr@gmail.com

* Corresponding Author: Shiv Govind

Abstract: Predicting age and gender from facial images is important for various uses, like social interactions, security, and customizing user experiences. Despite progress in facial recognition technology, getting accurate and reliable predictions is still a challenge. This paper reviews recent advancements in Convolutional Neural Networks (CNNs) for age and gender classification, comparing their effectiveness to traditional methods. CNNs excel in extracting features from images and have shown notable improvements in accuracy because they can learn relevant features automatically. The review covers several CNN-based methods, such as the Evolutionary-Fuzzy-Integral CNN (EFI-CNN), hybrid CNN-Extreme Learning Machine (ELM) models, and other sophisticated architectures. These approaches tackle issues like image quality, misalignment, and occlusion, resulting in better classification performance. Additionally, the paper explores how CNNs are used in broader image analysis tasks, highlighting their strengths in pattern recognition and efficiency. Overall, the integration of CNNs has significantly advanced age and gender classification, demonstrating their real-world potential.

Keywords: Convolutional Neural Networks (CNNs), age classification, gender classification, feature extraction, deep learning, transfer learning, hybrid models.

I. INTRODUCTION

Age and gender play a fundamental role in social interactions. Many languages use specific greetings and grammatical rules based on gender, and different forms of address are often used depending on a person's age. Despite the importance of these attributes in daily life, accurately estimating age and gender from facial images remains a challenge for commercial systems, even with the rapid advancement of face recognition technologies [1]. Automatic gender detection from facial images has gained significant attention in recent years, driven by the increasing availability of facial data from sources like social media, webcams, CCTV, and mobile devices. Gender identification has a wide range of applications, including visual surveillance, personalized advertising, user interface customization, and as a component of more complex systems like facial recognition and emotion analysis. However, it remains a difficult task, particularly when dealing with issues such as poor image quality, facial misalignment, and occlusion[2].

Historically, researchers approached age and gender recognition using handcrafted feature vectors and traditional machine learning techniques. While these methods were effective to a certain extent, they often struggled with unconstrained and varied image conditions. To address this, recent developments have focused on deep learning—particularly Convolutional Neural Networks (CNNs)—which can automatically learn and extract relevant features from facial images, significantly improving performance[3].

Image recognition has traditionally posed many challenges in machine learning, requiring careful selection and design of features based on color, texture, and spatial patterns. This manual process demanded expertise and often lacked generalization. CNNs, inspired by the human visual system, offer a more adaptive solution by learning features directly from the data. With advances in computing power, deeper and more complex CNN architectures have become feasible, leading to breakthroughs in visual recognition tasks. These networks have been applied successfully across domains such as character recognition and surveillance. In gender classification, earlier systems used methods like Principal Component Analysis (PCA) combined with simple classifiers, but they often failed in nuanced situations where humans could easily detect the correct gender. The shift to CNNs has enabled more accurate and robust gender classification, with various architectures tailored to this specific task and demonstrating high accuracy across different datasets and conditions[4].

II. CONVOLUTIONAL NEURAL NETWORKS (CNNs)

Convolutional Neural Networks (CNNs) have become increasingly prominent across a wide range of domains, including image and pattern recognition, speech processing, natural language understanding, and video analysis. A major factor contributing to their widespread adoption is their ability to automatically learn feature representations during training, as opposed to traditional models that rely on manually designed feature extractors. This capability allows CNNs to achieve

higher performance and efficiency, particularly in applications characterized by strong local correlations, such as those involving visual or auditory data. CNNs are also advantageous in terms of computational and memory efficiency. Their architectural design reduces the complexity and resource demands often associated with traditional models. Nonetheless, the training and evaluation of CNNs can be computationally intensive, frequently requiring high-performance hardware such as graphics processing units (GPUs), digital signal processors (DSPs), or specialized silicon optimized for throughput and energy efficiency[1], [5]–[8]. These platforms are essential to handle the extensive computations involved in deep learning tasks effectively. Structurally, CNNs are composed of multiple layers, typically including convolutional layers, pooling layers, and fully connected layers for classification. One of the earliest examples of a practical CNN architecture is LeNet-5, which laid the groundwork for modern deep learning models. The primary differences among CNN variants stem from how these core layers are structured and configured, including variations in the number, size, and arrangement of convolutional and pooling layers. Additionally, the training strategies used to optimize these networks play a significant role in their overall performance and generalization capabilities.

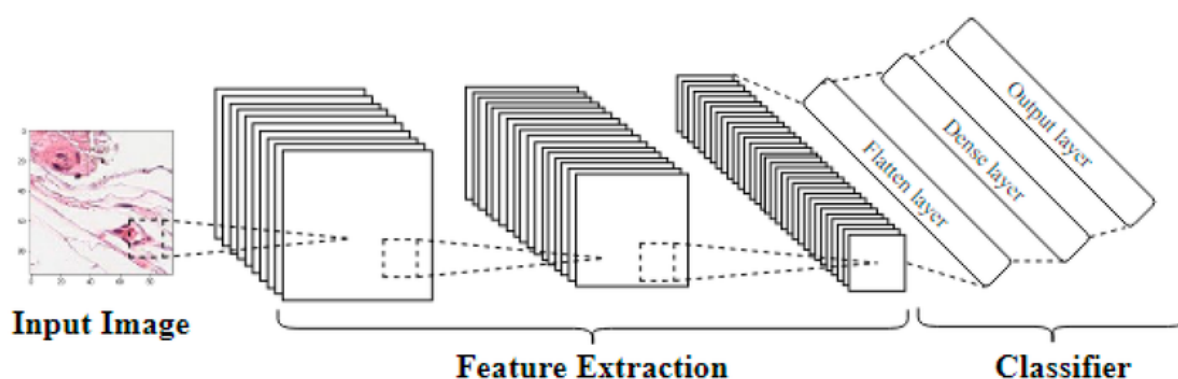


Figure 1 Convolutional neural network (CNN) diagram [8].

The figure 1 illustrates a Convolutional Neural Network (CNN) architecture, where an input image undergoes feature extraction through convolutional layers, followed by classification using flattening and dense layers [8].

In a Convolutional Neural Network (CNN), the convolutional layer plays a critical role in capturing detailed spatial information from the input image. The output from this layer is typically a tensor of dimensions $H \times W \times D$, where H and W denote the height and width of each feature map, and D represents the number of feature maps. These activations can also be interpreted as a set of D -dimensional feature vectors corresponding to each of the $H \times W$ spatial units in the image. Each spatial unit refers to a specific location, and its associated feature vector captures localized information at that point. Pooling layers, another core component of CNNs, help reduce the computational burden by downsampling the feature maps while retaining the most salient features, thereby preserving the essential information. The fully connected layer, unlike the convolutional and pooling layers that operate on localized regions, processes the entire image and integrates global information [2]–[4], [9], [10].

In more recent developments, activations from fully connected layers are increasingly used to describe specific image regions, rather than just summarizing the image as a whole. This approach involves densely sampling regions across the image, which can conceptually resemble applying convolutional operations. To distinguish this, such layers are referred to as "augmented convolutional layers" (AConv layers), whereas the original layers of the pre-trained CNN are called "original convolutional layers" (OConv layers).

III. KEY FINDINGS FROM RECENT STUDIES ON CNN-BASED GENDER CLASSIFICATION

Recent studies on gender and age classification using Convolutional Neural Networks (CNNs) demonstrate a wide range of architectural innovations, dataset selections, optimization strategies, and application-specific adaptations. A comparative analysis reveals the deployment of diverse CNN architectures such as the Evolutionary-Fuzzy-Integral CNN (EFI-CNN), hybrid CNN-Extreme Learning Machine (ELM) models, and simplified CNN frameworks utilizing cross-correlation techniques. These studies utilize benchmark datasets like Adience and MORPH II to address classification tasks including gender, age group, and facial expression recognition [11]–[16]. Optimization approaches such as Particle Swarm Optimization (PSO) and ELM have been applied to enhance model performance and computational efficiency. Notably, Ozbulak et al. (2019) achieved substantial accuracy improvements through the use of transfer learning, while Lin et al. integrated fuzzy logic to refine classification robustness. Haider et al. focused on lightweight models optimized

for smartphone deployment, showcasing the relevance of computational efficiency in real-world applications. Overall, the findings highlight that the success of CNN-based gender classification depends significantly on architectural design, training strategies, and task-specific customization, making them versatile tools for biometric and demographic analysis [17]–[21].

Table 1: Summary of CNN-Based Approaches for Gender and Age Classification

Architecture	Datasets Used	Classification Task	Optimization Techniques	Performance / Accuracy	Notable Features
EFI-CNN (Evolutionary-Fuzzy-Integral CNN)	CIA, MORPH, CACD2000	Age, Gender	Particle Swarm Optimization (PSO)	5.95% (Age), 3.1% (Gender) improvement over traditional methods	Combines multiple CNN outputs using Sugeno and Choquet fuzzy integrals
Hybrid CNN + ELM	MORPH II, Adience Benchmark	Age, Gender	Extreme Learning Machine (ELM)	Outperforms other models in accuracy and efficiency	CNN used for feature extraction; ELM for final classification
Simplified CNN with Cross-Correlation	SUMS, AT&T	Gender	2nd-order Backpropagation, Annealed Learning Rates	98.75% (SUMS), 99.38% (AT&T)	Efficient architecture; fusion of convolutional and subsampling layers
Hybrid CNN + SVM	Fingerprint datasets	Gender	CNN for features, SVM for classification	99.25%	Unique use of fingerprint data; hybrid deep learning-classical ML approach
Pareto Frontier CNNs (GoogleNet, SqueezeNet, ResNet50)	Unconstrained internet images	Gender	Transfer Learning	High accuracy in unconstrained settings	Leverages transfer learning with compact, high-performance CNNs
5-layer U-Net + AlexNet	NUAA, CASIA, Adience, IOG, CK+, JAFFE	Age, Gender, Spoofing, Expressions	Advanced deep architecture	83.26% (Age), 95.31% (Gender), 94.17% (Spoofing), 96.9% (Expressions)	Multi-task learning using U-Net and AlexNet for robust biometric classification
Simple CNN	Adience Benchmark	Age, Gender	Standard CNN	Outperforms existing methods on Adience	Minimalist design; ideal for limited-data environments
Deep Gender CNN	CAS-PEAL, FEI	Gender	Modified multilayer deep neural network	98%	Optimized for smartphones; low memory and power consumption
AlexNet-like, VGG-Face, GilNet	Adience Benchmark	Age, Gender	Transfer Learning	7% (Age), 4.5% (Gender) improvement over GilNet	Demonstrates effectiveness of domain-specific pre-trained models via transfer learning

Table 1 highlights that CNN-based approaches are highly effective for age and gender classification, especially when tailored with optimization methods and hybrid techniques. The architecture choice, dataset, and intended application domain all influence performance. Modern trends focus on scalability, real-world deployment, and multi-task learning, making CNNs central to intelligent biometric systems [22]-[24].

IV. CONCLUSION

The application of Convolutional Neural Networks (CNNs) to age and gender classification marks a significant advancement in the field of facial image analysis. CNNs are particularly effective due to their ability to automatically extract hierarchical features and adapt across various datasets and classification tasks. This review summarizes a range of innovative CNN-based approaches developed to address the complexities of age and gender prediction. Techniques such as the Evolutionary-Fuzzy-Integral CNN (EFI-CNN) leverage fuzzy logic to improve decision-making, while hybrid models like CNN-Extreme Learning Machine (ELM) and CNN-Support Vector Machine (SVM) demonstrate enhanced performance through the integration of deep learning with traditional classifiers. Furthermore, transfer learning models based on architectures like GoogleNet and ResNet50 have achieved high accuracy, particularly in unconstrained environments with variable image quality. However, challenges such as image noise, occlusions, and facial misalignment continue to affect model robustness. Future research should focus on improving model generalization through enhanced architectures and optimization techniques, enabling more accurate and resilient performance in real-world scenarios. Continued innovation in CNN design and training strategies will be key to advancing intelligent, adaptive systems for applications ranging from biometric security to personalized user experiences.

REFERENCES

- [1] S. Kumar, A. Kumar, C. Gupta, A. Chaturvedi, and A. P. Tripathi, "Synergy of AI and PMBLDC Motors: Enhancing Efficiency in Electric Vehicles," *IEEE Int. Conf. "Computational, Commun. Inf. Technol. ICCIT 2025*, pp. 68–73, 2025, doi: 10.1109/ICCIT62592.2025.10927757.
- [2] A. Kumar and S. Jain, "Predictive Switching Control for Multilevel Inverter using CNN-LSTM for Voltage Regulation," vol. 11, pp. 1–9, 2022.
- [3] C. Gupta and V. K. Aharwal, "Design and simulation of Multi-Input Converter for Renewable energy sources," *J. Integr. Sci. Technol.*, vol. 11, no. 3, pp. 1–7, 2023.
- [4] C. B. Singh, A. Kumar, C. Gupta, S. Cience, T. Echnology, and D. C. Dc, "Comparative performance evaluation of multi level inverter for power quality improvement," vol. 12, no. 2, pp. 1–7, 2024.
- [5] A. Kumar and S. Jain, "Multilevel Inverter with Predictive Control for Renewable Energy Smart Grid Applications," *Int. J. Electr. Electron. Res.*, vol. 10, no. 3, pp. 501–507, 2022, doi: 10.37391/IJEER.100317.
- [6] A. Kumar and S. Jain, "Enhancement of Power Quality with Increased Levels of Multi-level Inverters in Smart Grid Applications," vol. 14, no. 4, pp. 1–5, 2022, doi: 10.18090/samriddhi.v14i04.07.
- [7] S. Kumar, A. Kumar, C. Gupta, and A. Chaturvedi, "Future Trends in Fault Detection Strategies for DC Microgrid," *Proc. - 2024 IEEE 16th Int. Conf. Commun. Syst. Netw. Technol. CICON 2024*, pp. 727–731, 2024, doi: 10.1109/CICON63059.2024.10847358.
- [8] C. Gupta and V. K. Aharwal, "Design of Multi Input Converter Topology for Distinct Energy Sources," *SAMRIDDHI*, vol. 14, no. 4, pp. 1–5, 2022, doi: 10.18090/samriddhi.v14i04.09.
- [9] C. Gupta and V. K. Aharwal, "Optimizing the performance of Triple Input DC-DC converter in an Integrated System," *J. Integr. Sci. Technol.*, vol. 10, no. 3, pp. 215–220, 2022.
- [10] A. Kumar and S. Jain, "Critical Analysis on Multilevel Inverter Designs for," vol. 14, no. 3, 2022, doi: 10.18090/samriddhi.v14i03.22.
- [11] S. Kumar and A. Kumar, "A Review on PWM Based Multicarrier Multilevel Inverter with Reduced Number of Switches," *Smart Moves J. Ijoscience*, vol. 6, no. 7, pp. 24–31, 2020, doi: 10.24113/ijoscience.v6i7.309.
- [12] V. Meena and C. Gupta, "A Review of Design , Development , Control and Applications of DC – DC

Converters,” no. 2581, pp. 28–33, 2018.

- [13] A. K. Singh and C. Gupta, “Controlling of Variable Structure Power Electronics for Self-Contained Photovoltaic Power Technologies,” vol. 05, no. 02, pp. 70–77, 2022.
- [14] A. Hridaya and C. Gupta, “Hybrid Optimization Technique Used for Economic Operation of Microgrid System,” *Academia.Edu*, vol. 5, no. 5, pp. 5–10, 2015, [Online]. Available: http://www.academia.edu/download/43298136/Aditya_pape_1.pdf.
- [15] P. Verma and C. Gupta, “A Survey on Grid Connected Solar Photovoltaic System,” *Int. Conf. Contemp. Technol. Solut. Towar. fulfilment Soc. Needs*, pp. 106–110, 2018, [Online]. Available: https://www.academia.edu/37819420/A_Survey_on_Grid_Connected_Solar_Photovoltaic_System.
- [16] P. Verma and M. T. Student, “Three Phase Grid Connected Solar Photovoltaic System with Power Quality Analysis,” pp. 111–119, 2018.
- [17] C. G. Aditya Hridaya, “International Journal of Current Trends in Engineering & Technology ISSN : 2395-3152 AN OPTIMIZATION TECHNIQUE USED FOR ANALYSIS OF A HYBRID International Journal of Current Trends in Engineering & Technology ISSN : 2395-3152,” *Int. J. Curr. Trends Eng. Technol.*, vol. 06, no. October, pp. 136–143, 2015.
- [18] P. Ahirwar and C. Gupta, “Simulation of Continuous Mode Hybrid Power Station with Hybrid Controller,” vol. 03, no. 02, pp. 58–62, 2020.
- [19] S. Khan, C. Gupta, and A. Kumar, “An Analysis of Electric Vehicles Charging Technology and Optimal Size Estimation,” vol. 04, no. 04, pp. 125–131, 2021.
- [20] P. Mahapatra and C. Gupta, “Study of Optimization in Economical Parameters for Hybrid Renewable Energy System,” *Res. J. Eng. Technol. ...*, vol. 03, no. 02, pp. 63–65, 2020, [Online]. Available: http://www.rjetm.in/RJETM/Vol03_Issue02/Study of Optimization in Economical Parameters for Hybrid Renewable Energy System.pdf.
- [21] B. B. Khatua, C. Gupta, and A. Kumar, “Harmonic Investigation Analysis of Cascade H Bridge Multilevel Inverter with Conventional Inverter using PSIM,” vol. 04, no. 03, pp. 9–14, 2021.
- [22] Serin, J., Vidhya, K. T., Deepa, I. M. I., Ebenezer, V., & Jenefa, A. (2024). Gender Classification from Fingerprint Using Hybrid CNN-SVM. *Journal of Artificial Intelligence and Technology*, 4(1), 82-87.
- [23] Kumar, S., Rani, S., Jain, A., Verma, C., Raboaca, M. S., Illés, Z., & Neagu, B. C. (2022). Face spoofing, age, gender and facial expression recognition using advance neural network architecture-based biometric system. *Sensors*, 22(14), 5160.