**AGE AND GENDER DETECTION USING**

**CONVOLUTIONAL NEURAL NETWORKS**

BY

**RAGHU RAM BOBBA (1322035)**

**BHARATH CHANDRA R GUMIREDDY(1321773)**

DEPARTMENT OF ENGINEERING AND

COMPUTING SCIENCES

Submitted in partial fulfillment of the

requirements for the Degree of

Masters of Science in Data Science

in the Graduate College of the

New York Institute of Technology

Approved by: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Maryam Ravan, PhD

May 2024

**TABLE OF CONTENTS**

[**Chapter 1: Introduction & Literature Survey**](https://docs.google.com/document/d/1xuMtvOIUmE1-5wSKcM6snQtnRX4by_-vL5a27JMqnSs/edit?pli=1#heading=h.1fob9te) 3 [**Abstrac**](https://docs.google.com/document/d/1xuMtvOIUmE1-5wSKcM6snQtnRX4by_-vL5a27JMqnSs/edit?pli=1#heading=h.3znysh7)**t………………………………………………………………………………………………… 3**

[**Introduction & Literature Review**](https://docs.google.com/document/d/1xuMtvOIUmE1-5wSKcM6snQtnRX4by_-vL5a27JMqnSs/edit?pli=1#heading=h.tyjcwt) **4**

[**Chapter 2:**](https://docs.google.com/document/d/1xuMtvOIUmE1-5wSKcM6snQtnRX4by_-vL5a27JMqnSs/edit?pli=1#heading=h.4d34og8) **Objective 10**

[**Chapter 3: MATERIALS & METHODOLOGY**](https://docs.google.com/document/d/1xuMtvOIUmE1-5wSKcM6snQtnRX4by_-vL5a27JMqnSs/edit?pli=1#heading=h.17dp8vu) **11**

[**3.1 Data Overview**](https://docs.google.com/document/d/1xuMtvOIUmE1-5wSKcM6snQtnRX4by_-vL5a27JMqnSs/edit?pli=1#heading=h.3rdcrjn) **11**

**3**[**.2 Methodology**](https://docs.google.com/document/d/1xuMtvOIUmE1-5wSKcM6snQtnRX4by_-vL5a27JMqnSs/edit?pli=1#heading=h.26in1rg) **12**

[**3.3**](https://docs.google.com/document/d/1xuMtvOIUmE1-5wSKcM6snQtnRX4by_-vL5a27JMqnSs/edit?pli=1#heading=h.lnxbz9) **Methodological Framework****13**

**3.3.1. Initial Setup 13**

**3.3.2. Real-time Video Processing 13**

**3.3.3. Face Detection 13**

**3.3.4. Image Preprocessing for Prediction 14**

**3.3.5. Model Prediction 14**

**3.3.6. Results Display 14**

[**3.4**](https://docs.google.com/document/d/1xuMtvOIUmE1-5wSKcM6snQtnRX4by_-vL5a27JMqnSs/edit?pli=1#heading=h.35nkun2) **Network Architecture****14**

**3.5 Training and Testing the Data 16**

**3.6 Sensitivity and Specificity of Each Group 23**

[**Chapter 4: Results and Discussion**](https://docs.google.com/document/d/1xuMtvOIUmE1-5wSKcM6snQtnRX4by_-vL5a27JMqnSs/edit?pli=1#heading=h.3j2qqm3)  **25**

**4.1** [**Results**](https://docs.google.com/document/d/1xuMtvOIUmE1-5wSKcM6snQtnRX4by_-vL5a27JMqnSs/edit?pli=1#heading=h.1y810tw) **25**

**4.2 Discussion 31**

[**Chapter**](https://docs.google.com/document/d/1xuMtvOIUmE1-5wSKcM6snQtnRX4by_-vL5a27JMqnSs/edit?pli=1#heading=h.3j2qqm3) **5: Conclusion…………………………………………………………………………………………32**

**Bibliography……………………………………………………………………………………………….33**

# **Abstract**

This research introduces novel methodologies for age and gender classification, leveraging Convolutional Neural Networks (CNNs). We emphasize model performance improvement through innovative approaches and advanced traditional models. Our methods include integrating sophisticated data balancing techniques, optimizing training epochs, and introducing a novel application of loss functions.

Our methodology encompassed several critical stages: facial detection, data preprocessing, and deploying various CNN architectures to capture a broad spectrum of facial features. We categorized the data into the following age ranges: 0-5 years (2,167 subjects), 6-17 years (1,908 subjects), 18-40 years (12,399 subjects), 41-64 years (5,226 subjects), and 65-120 years (2,008 subjects). By adjusting age and gender classification bins and employing the Mean Absolute Error (MAE) as a loss function, we effectively mitigated overfitting—a prevalent issue in earlier models.

This is a novel application of the MAE loss function in age and gender classification, which has not been explored extensively in the literature. These modifications resulted in a substantial enhancement in model performance, with a notable reduction in loss to 4.78798 and overall accuracy of approximately 88% for Gender Evaluation and 70% for Age Group Evaluation.

Moreover, we explored the impact of data balancing by modifying the composition of the input datasets, which was crucial in enhancing the model's efficiency. We employed a combination of oversampling and undersampling techniques to balance the data across different age groups and genders. Comparative results from tests with both balanced and unbalanced datasets underscored the efficacy of our methodological improvements.

This study underscores the practical applications of CNNs in accurately predicting age and gender. It offers significant benefits for real-world scenarios such as digital identity verification, targeted marketing, and demographic analysis. The potential impact of this research is substantial, making it a crucial area of study for professionals in computer vision, machine learning, and data science.

**Chapter 1: Introduction and Literature Review**

In computer vision, the automatic detection and classification of age and gender from facial images present significant challenges. These tasks have extensive applications, ranging from security systems and forensic analysis to targeted advertising and social media. They require high accuracy and robustness across diverse demographic groups and varying lighting and environmental conditions.

Traditional approaches to age and gender classification heavily relied on hand-crafted features and classical machine learning techniques, which often struggled with the complexity and variability of human faces. However, the advent of deep learning, particularly Convolutional Neural Networks (CNNs), has transformed this field. CNNs have shown exceptional capability in capturing intricate patterns in image data, making them particularly suited for facial recognition tasks.

Despite significant advancements, several challenges persist in age and gender detection. Overfitting due to limited labeled datasets and the necessity for models to generalize well across different populations without bias is still prevalent. Furthermore, the dynamic nature of human aging and subtle gender-specific facial features necessitate highly sensitive and adaptable model architectures. This research aims to address these challenges and push the boundaries of age and gender detection using CNNs.

This study builds on the foundation of existing CNN approaches, introducing innovative methods to enhance accuracy and efficiency. By refining the architecture and training procedures and exploring novel loss functions, this research aims to advance the state-of-the-art in age and gender detection. The subsequent sections will detail the methodology employed, the experiments conducted, and insights gleaned from the results, providing a comprehensive overview of the improvements achieved over traditional models.

**Understanding Age and Gender Classification Using CNNs**

Age and gender classification are vital tasks in computer vision, with broad applications in security, marketing, and personalized user experiences. These tasks involve predicting an individual's age group and gender from facial images using deep learning techniques, particularly Convolutional Neural Networks (CNNs). CNNs have remarkably succeeded in image classification because they can learn hierarchical features from raw images.

In recent years, CNNs have demonstrated significant success in various image classification tasks due to their ability to learn and extract hierarchical features from raw images automatically. The Adience dataset, comprising 17,856 photos of 2,284 subjects across eight age groups (0-2, 4-6, 8-13, 15-20, 25-32, 38-43, 48-53, and 60+ years), has been widely used for age and gender classification tasks due to its diversity and real-world complexity [1].

**Challenges in Age and Gender Classification**

Despite significant advancements, several challenges persist in age and gender classification. One major issue is the imbalanced datasets often used in these tasks. Real-world datasets typically have unequal distribution across different age groups and genders, which can bias the models. Additionally, age-related facial changes are subtle and non-linear, making accurate classification difficult. Techniques such as data augmentation (rotation, scaling, flipping) are employed to address these imbalances and enhance model robustness.

**CNN Architectures and Techniques**

CNN architectures for age and gender classification typically include several convolutional layers, pooling, and fully connected layers. The models are trained using categorical cross-entropy as the loss function and the Adam optimizer for weight optimization. A common approach involves three convolutional layers with ReLU activation functions and pooling layers to reduce feature map dimensions.

P. Jeevan and N. V. Sailaja's 2022 study employed a CNN architecture with three convolutional layers using the Adience dataset. Their model achieved a gender classification accuracy of 85.16% and an age group classification accuracy of 44.26%. These findings suggest that gender classification was robust, while age group classification was more challenging due to subtle facial variations [1].

**Evaluation Metrics and Results**

The performance of CNN models for age and gender classification is generally evaluated using accuracy metrics. Studies have reported varying degrees of success, with gender classification often yielding higher accuracy than age group classification. For instance, one study reported a gender classification accuracy of 85.16% and an age group classification accuracy of 44.26% on the Adience dataset [1].

**Comparative Studies and Advances**

Several studies have explored different CNN architectures and techniques to improve classification accuracy. These studies have utilized various datasets, model architectures, and strategies to address the inherent challenges of age and gender classification.

A. Ghildiyal and E. P. Ijjina (2022): This study focused on gender and age detection using Deep Convolutional Neural Networks. They utilized the Adience dataset, which contains 26,580 images divided into eight age categories and near-equal gender distribution. The CNN architecture included three convolutional layers, each followed by ReLU activation functions and pooling layers. Data augmentation techniques such as rotation, scaling, and flipping were employed to handle data imbalance. The model achieved a gender classification accuracy of 96.2% and an age group classification accuracy of 61.3%. These results highlight the effectiveness of CNNs in capturing gender-distinguishing features and the challenges posed by age-related facial variations [2].

S. Ghosh and S. K. Bandyopadhyay (2015): This paper utilized Multi-Class Support Vector Machines (SVM) for gender classification and age detection using a custom dataset of 119 JPEG images. The dataset was categorized into three age groups: child, adult, and old. The authors proposed a three-step methodology involving pre-processing, feature extraction, and classification. The pre-processing steps included noise removal, histogram equalization, and size normalization, followed by face detection. Feature extraction focused on the 'lip' region of the facial images, reshaped from 2D to 1D to form a feature vector. The classification employed a "One against All" strategy with Multi-Class SVM classifiers for each age group. This approach ensured the SVM could distinguish one class from the remaining courses. The study achieved 83.33% accuracy for age classification and 90% for gender classification, demonstrating the feasibility of using Multi-Class SVM for these tasks [3].

K. R. Hassan and I. H. Ali (2020): This study presented a method for age and gender classification using multiple CNN architectures with the Aging Faces in the Wild (AGFW) dataset, containing approximately 36,299 face images. The proposed method consisted of five phases: face detection, background removal, face alignment, multiple CNN architectures, and a voting system. The multiple CNN model included three distinct CNN architectures to extract various features. These networks were trained separately, and their predictions were combined using a voting system to achieve the final classification. The experimental results demonstrated the effectiveness of this approach, reaching 95.7% accuracy for gender classification and 72.0% for age classification. The ensemble method proved beneficial in capturing a broader range of age-related features and improving classification accuracy [4].

A. Saha N. K. S and N. P. (2023): This study focused on age and gender prediction using adaptive gamma correction and CNNs. The authors used the Adience dataset comprising 17,856 images of 2,284 subjects divided into eight age categories. The proposed CNN architecture integrated adaptive gamma correction for image enhancement to improve classification accuracy. The architecture consisted of two levels: feature extraction and classification. The adaptive gamma correction step enhanced the contrast of facial images, making feature extraction more effective. The model achieved a gender classification accuracy of 96.2% and an age group classification accuracy of 61.3%, indicating the effectiveness of the proposed approach in capturing gender-distinguishing features and handling age-related variations in facial images [5].

Md. N. I. Opu et al. (2020): This study developed a lightweight deep CNN model for real-time age and gender prediction, utilizing a combined dataset of the Wiki, UTKFace, and Adience datasets totaling 18,728 images. The CNN model was designed to be lightweight and suitable for mobile integration, with a total of 210,050 parameters and a final model size of 2.60 MB. The model architecture included several convolutional layers with varying kernel sizes to extract features, followed by pooling layers for down-sampling without reducing the spatial dimensions. Data augmentation techniques such as rotation, scaling, and flipping were employed to address data imbalance. The model achieved a gender classification accuracy of 80.76% and an age classification accuracy of 48.59%. The model's real-time capability was also demonstrated with an average prediction time of 0.0654 seconds per image, making it suitable for mobile and embedded systems [6].

R. Thaneeshan et al. (2022): This study proposed a simple and efficient CNN model for gender and age estimation using the Adience benchmark dataset, which includes 19,487 facial images of 2,284 subjects divided into eight age categories. The dataset posed challenges due to variations in lighting conditions, poses, and backgrounds. The CNN model architecture comprised three convolutional layers followed by three fully-connected layers. Images were resized to 256x256 pixels and center-cropped to 227x227 pixels to focus on the facial region. Data augmentation techniques such as horizontal flipping were employed to increase the number of training samples and address data imbalance. The model achieved an accuracy of 84.20% for gender estimation and 57.60% for age estimation, indicating its effectiveness in handling real-world variations and accurately classifying gender and age groups from facial images [7].

S. T. Rahman et al. (2020): This study presented a method for human age and gender estimation using facial image processing techniques, leveraging edge detection, binary masks, and wrinkle density evaluation. The authors utilized the BUET facial database, comprising 400 samples with 243 images of males and 157 images of females aged 16 to 50. The proposed method involved converting RGB images to HSV and grayscale for edge detection, focusing on significant facial landmarks like eyes and lips. Binary masks created for these regions were processed using the Canny edge detection algorithm, followed by Gaussian filtering to smooth the images and remove noise. Wrinkle density was calculated by dividing the detected edges by the total area of the regions of interest. These values were used to estimate age through a Naive Bayes classifier. For gender determination, images were resized to 512x512 pixels to standardize the relative positions of facial features. The Viola-Jones algorithm was employed to detect facial landmarks, and logistic regression was used to classify gender based on these ratios. The experimental results showed an accuracy of 76.3% for age classification and 86.6% for gender classification on the BUET facial database, with validation against other datasets such as Adience, ORL, and the University of Essex demonstrating robustness and effectiveness [8].

**Summary of Studies**

The following table (Table 1) summarizes the key datasets, methodologies, and results from the reviewed studies, highlighting the diversity in approaches and their respective performance metrics.

**Table 1: Dataset and Results Summary**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Paper Title | Data Set Used | Number of Images | Age Group Classification | Age Classification Accuracy | Gender Classification Accuracy |
| Age and Gender Classification using CNN [1] | Adience | 17,856 | 0-2, 4-6, 8-13, 15-20, 25-32, 38-43, 48-53, and 60+ years | 44.26% | 85.16% |
| Gender and Age Detection using Deep Convolutional Neural Networks [2] | Adience | 26,580 | 0-2, 4-6, 8-13, 15-20, 25-32, 38-43, 48-53, and 60+ years | 61.3% | 96.2% |
| Gender Classification and Age Detection Based on Human Facial Features Using Multi-Class SVM [3] | Custom | 119 | Child - 40 Images,  Adult -40 Images, Old Aged - 39 Images | 83.33% | 90% |
| Age and Gender Classification using Multiple CNN [4] | AGFW | 36,299 | 10-14, 20-24, 30-34, 40-44, 50-64 | 72.0% | 95.7% |
| Age and Gender Prediction using Adaptive Gamma Correction [5] | Adience | 17,856 | 0-2, 4-6, 8-13, 15-20, 25-32, 38-43, 48-53, 60+ years | 61.3% | 96.2% |
| Lightweight Deep CNN for Real-Time Age and Gender Prediction [6] | Combined (Wiki, UTKFace, Adience) | 18,728 | - | 48.59% | 80.76% |
| Gender and Age Estimation Using Deep Learning [7] | Adience | 19,487 | 0-2, 4-6, 8-13, 15-20, 25-32, 38-43, 48-53, 60+ years | 57.60% | 84.20% |
| Human Age and Gender Estimation Using Facial Image Processing [8] | BUET, Adience, ORL, University of Essex | 400 (BUET), 26,580 (Adience), 400 (ORL), 4,500 (Essex) | - | 76.3% (BUET) | 86.6% (BUET) |

**Conclusion**

The literature reviewed illustrates the significant advancements and ongoing challenges in age and gender classification using CNNs. The studies highlight the diversity in methodologies and datasets, each contributing to improving classification accuracy and addressing specific challenges.

CNNs have proven effective in capturing complex patterns and features from facial images, making them suitable for age and gender classification tasks. Techniques such as data augmentation and integrating adaptive methods like gamma correction have enhanced model performance, particularly in handling real-world data variability. Ensemble methods, which combine predictions from multiple models, have demonstrated improved accuracy by leveraging the strengths of different architectures.

Despite these advancements, challenges remain, particularly in accurately classifying age groups due to subtle and non-linear facial changes. Imbalanced datasets pose a significant hurdle, necessitating data augmentation and other techniques to ensure balanced representation. Furthermore, achieving real-time classification capabilities without compromising accuracy is crucial for practical applications in mobile and embedded systems.

**Chapter 2: Objective**

The primary objective of this research is to enhance the accuracy and efficiency of age and gender detection systems using Convolutional Neural Networks (CNNs). This study aims to achieve this by integrating advanced data balancing techniques, optimizing training methodologies, and introducing innovative loss functions.

Sophisticated data balancing ensures a well-distributed training dataset across age and gender groups, preventing model bias. Training involves fine-tuning hyperparameters and using advanced preprocessing techniques like image normalization to improve robustness against overfitting.

The research introduces the Mean Absolute Error (MAE) loss function to achieve accurate predictions. A robust face detection algorithm ensures precise facial region localization and preprocessing. Various CNN architectures are evaluated to find the optimal structure balancing performance and efficiency.

Designed for real-time operation, the system uses OpenCV for video capture and TensorFlow/Keras for deploying the model, enabling effective real-time detection and classification.

The model's performance is evaluated using accuracy, precision, recall, and F1 score for gender and age classifications, with detailed confusion matrices identifying areas for improvement.

The goal is to develop a reliable detection system for applications in digital identity verification, targeted marketing, and demographic analysis. Future research will focus on expanding the dataset, exploring additional CNN configurations, and incorporating more demographic attributes to enhance accuracy and applicability.

This research aims to advance state-of-the-art age and gender detection using CNNs, providing significant practical benefits and insights into computer vision.

**Chapter 3: MATERIALS & METHODOLOGY**

**3.1: Data Overview**

**Data Description:**

The dataset used in this study is the UTKFace dataset, which can be accessed here(h**ttps://www.kaggle.com/datasets/jangedoo/utkface-new).** This dataset contains 23,708 face photos categorized by age, gender, and age group. The images cover various age groups, ethnicities, and backgrounds, providing a diverse set of images for robust model training. Of the 23,708 samples, 12,391 are male, and 11,317 are female, resulting in an almost balanced dataset. This balance helps reduce bias towards any particular gender in the model's predictions. The dataset spans a broad age range, initially categorized into subgroups including 0-2, 3-5, 6-10, 11-17, 18-25, 26-35, 36-45, 46-60, and 61-120 years. We reclassified and merged some age groups to ensure enough samples for each class and improve the model’s fairness and generalization capabilities. This was crucial to address the imbalance in the dataset and ensure robust model training. The original dataset had varying sample sizes across different age groups, which posed challenges in achieving consistent accuracy. Therefore, we merged the age groups as follows: Infants (0-2 years) and Toddlers/Preschoolers (3-5 years) were combined into a single class (0-5 years), resulting in 2,167 samples. Children (6-10 years) and Adolescents/Teens (11-17 years) were combined into another class (6-17 years), with a total of 1,908 samples. The remaining age groups were adjusted to form the following categories: 18-40 years (12,399 samples), 41-64 years (5,226 samples), and 65-120 years (2,008 samples). The age classification of the data before and after balancing is shown in Tables 2 and 3, respectively.

**Table 2: Age Classification of the Raw Data**

|  |  |  |
| --- | --- | --- |
| Age Group | Age Classification | Count |
| Infants | 0-2 | 1123 |
| Toddlers/Preschoolers | 3-5 | 1044 |
| Children | 6-10 | 1116 |
| Adolescents/Teens | 11-17 | 792 |
| Young Adults | 18-35 | 12399 |
| Middle Age | 36-60 | 5226 |
| Seniors | 61-120 | 2008 |

To ensure an adequate number of samples for each class and to enhance the model's fairness and generalization capabilities, the dataset was reclassified as follows:

We merged the data during the classification by considering the **infants and toddlers** into one class and **teens and children** into another class to have a good number of samples for each class. Age is classified as 0 -5, 6-17, 18-40, 41-64, 65-120.

**Table 3: Age Classification after Balancing the Data**

|  |  |  |
| --- | --- | --- |
| Age Group | Age Classification | Count |
| Toddlers/Preschoolers | 0-5 | 2167 |
| Children | 6-10 | 1908 |
| Young Adults | 18-35 | 12,399 |
| Middle Age | 36-60 | 5226 |
| Seniors | 61-120 | 2008N |

**3.2: Methodology**

**Introduction :**

This section outlines the methodology for designing and implementing a real-time age and gender detection system. The system integrates advanced technologies, including Python programming, OpenCV for real-time image processing, and TensorFlow/Keras for deploying a machine learning model. This combination enables real-time detection and classification of human faces in video streams from a webcam.

The methodology addresses the inherent challenges in age and gender detection through the innovative use of convolutional neural networks (CNNs). Our approach leverages advanced data preprocessing techniques, a refined training strategy, and the application of novel loss functions to optimize the performance of CNN models. Initially, we begin with a robust face detection algorithm to ensure accurate localization and cropping of facial regions from diverse images. This is followed by data augmentation procedures to enhance model robustness against overfitting by artificially expanding the training dataset with transformed images that mimic various real-world conditions.

Multiple models with varying depths and complexities were evaluated for the CNN architectures to identify the optimal structure that balances performance with computational efficiency. We utilized a customized training regimen that included fine-tuning hyperparameters such as learning rates and epoch numbers. A significant innovation in our methodology is the introduction of the Mean Absolute Error (MAE) loss function, selected for its effectiveness in reducing prediction errors more significantly than traditional loss functions used in classification tasks.

Each step in our methodology is meticulously designed to contribute towards a model that performs with high accuracy and generalizes well across different datasets, ensuring reliability in practical applications. The following sections will delve deeper into each of these components, providing a detailed account of our experimental setup, the rationale behind our choices, and the execution of the training process.

**3.3 Methodological Framework**

The system's methodology is divided into several key components: initial setup, real-time video processing, face detection, image preprocessing for prediction, model prediction, and results display. Each element is crucial in ensuring the system functions effectively in real-time.

**3.3.1: Initial Setup:**

The process begins with setting up the necessary Python environment, which involves installing and importing essential libraries such as OpenCV, TensorFlow/Keras, and NumPy. OpenCV is utilized for capturing video feeds and processing images, TensorFlow/Keras is used to load the pre-trained neural network model that predicts age and gender, and NumPy aids in handling numerical operations efficiently. These libraries provide the foundational tools for image processing, data manipulation, and machine learning operations.

**3.3.2: Real-time Video Processing:**

The system's core is the continuous video feed captured from the webcam. OpenCV’s `VideoCapture` function initializes the webcam and captures live video frames. Each frame is processed in a loop until the user decides to exit, typically by pressing a key like 'q.' This continuous capture and processing of video frames allow the system to function in real time.

**3.3.3: Face Detection:**

Face detection is performed on each frame using a Haar Cascade Classifier, a machine learning-based approach provided by OpenCV. This classifier efficiently detects objects, in this case, human faces, by converting the captured frame to grayscale to simplify the image data and reduce computational requirements. The Haar Cascade Classifier is then applied to the grayscale image to detect faces using predefined models trained on thousands of face examples.

**3.3.4: Image Preprocessing for Prediction:**

Once faces are detected, each face region is processed to prepare it for prediction. This involves extracting the face region and resizing it to the required input size for the neural network, typically 128x128 pixels. The resized image is then converted to grayscale to maintain consistency with the model's training data. Finally, the image is normalized by scaling pixel values to a range of 0 to 1, matching the normalization technique used during the model training phase.

**3.3.5: Model Prediction:**

The preprocessed face images are fed into the neural network model to predict gender and age. This step involves passing the normalized face image data to the model using TensorFlow/Keras and retrieving the prediction outputs, typically including the predicted gender and age or age group. The model utilizes the pre-trained neural network to make these predictions based on the input data.

**3.3.6: Results Display:**

The final step is to display the prediction results on the original video feed in real time. This is achieved by drawing rectangles around detected faces on the video frame and displaying text annotations near each face to show the predicted gender and age. The live video feed is updated to reflect these annotations, providing instant visual feedback on the predictions.

By meticulously designing each step of this methodology, the system performs with high accuracy and generalizes well across different datasets, ensuring reliability in practical applications. This detailed approach to real-time age and gender detection leverages the strengths of convolutional neural networks (CNNs) and advanced image processing techniques to deliver robust and efficient performance.

**3.4 Network Architecture**

The provided Convolutional Neural Network (CNN) architecture for age and gender detection, as shown in Figure 1, is meticulously designed to process input images, extract relevant features, and predict multiple outputs, including gender, age groups, and actual age. The network begins with an input layer that serves as the entry point for images into the model. This layer accepts images of a specific shape, typically 128x128 pixels, and since the photos are in grayscale, the depth is 1.

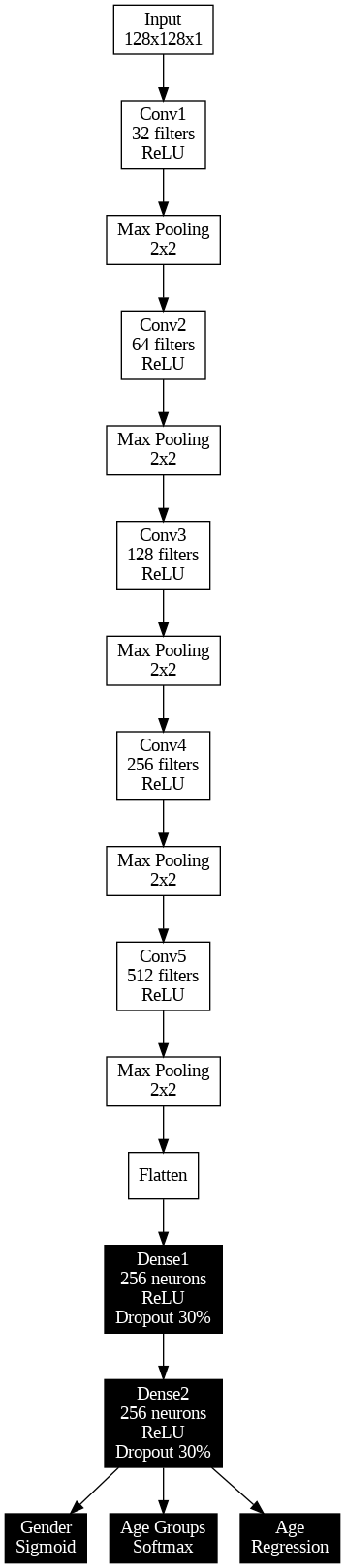
The architecture includes several convolutional layers, each performing convolution operations to detect various image features, from simple edges to complex patterns. The first convolutional layer applies 32 filters to the input image, each filter being a small matrix that slides over the image to detect specific patterns. The ReLU activation function is used to introduce non-linearity. After the convolution, a max-pooling operation is performed to reduce the spatial dimensions of the feature map by half, making the model less sensitive to small translations in the image. The second convolutional layer uses 64 filters to detect more complex features from the output of the first layer. Again, ReLU activation is used, followed by max pooling to reduce the spatial dimensions further.

The third convolutional layer applies 128 filters to the output from the second layer to detect even higher-level features. ReLU activation is used, and max pooling reduces the spatial dimensions again. The fourth convolutional layer uses 256 filters to capture more complex patterns, followed by ReLU activation and another max-pooling operation to reduce the feature map's size further, and the final convolutional layer applies 512 filters, allowing the model to learn detailed and high-level features from the images. ReLU activation is used, and a final max-pooling operation significantly reduces the spatial dimensions.

Following the convolutional layers, the network includes a flattened layer, which converts the multi-dimensional output from the convolutional layers into a one-dimensional vector. This transformation is necessary to connect the convolutional layers to the fully connected layers that follow. The fully connected (dense) layers perform classification based on the features extracted by the convolutional layers. The first dense layer contains 256 neurons that fully connect to all the activations in the previous layer, using the ReLU activation function to introduce non-linearity. A dropout layer follows, randomly setting 30% of the input units to zero during training. This helps prevent overfitting by ensuring the model does not rely too much on individual neurons. The second dense layer, another fully connected layer with 256 neurons and ReLU activation, is followed by another dropout layer with a 30% rate to prevent further overfitting.

The output layers provide the final predictions for gender, age groups, and actual age. The gender output consists of a single neuron with sigmoid activation, producing a probability indicating whether the person in the image is male or female. The age group's production includes several neurons, one for each age group, with softmax activation, producing a probability distribution over the age groups. The age regression output consists of a single neuron with linear activation, predicting the person's age in the image.

The model is compiled using binary cross-entropy for gender classification, mean absolute error (MAE) for age prediction, and categorical cross-entropy for age group classification, which are appropriate for their respective tasks. The Adam optimizer adjusts the learning rate throughout training, enhancing the model’s performance and speed. Accuracy metrics are used to evaluate the performance of the classification outputs, measuring how often the model's predictions are correct. This architecture effectively combines convolutional layers for feature extraction and fully connected layers for classification and regression, making it suitable for complex image recognition tasks such as gender and age prediction. Including dropout layers help mitigate overfitting, and the activation and loss functions are chosen appropriately for the mixed classification and regression objectives.



**Figure 1: CNN Network Architecture**

**3.5 Training and Testing the Data**

To ensure a balanced representation of age groups within the training and testing datasets, the dataset, which consists of 23,708 images, was initially split into two subsets. Specifically, 80% of the photos, totaling 18,964 images, were allocated to the training set, while the remaining 20%, amounting to 4,744 images, were reserved for the testing set. This initial split is crucial as it ensures that a significant portion of the data is utilized to train the model. In contrast, a substantial portion is retained to evaluate its performance on unseen data.

The initial counts for each age group in the overall dataset were considered to maintain a balanced distribution of each age group within the training set. These counts were as follows: 2,167 images for Toddlers/Preschoolers (ages 0-5), 1,908 images for Children (ages 6-10), 12,399 images for Young Adults (ages 18-35), 5,226 images for Middle Age individuals (ages 36-60), and 2,008 images for Seniors (ages 61-120). The proportionate counts for each age group within the training set were calculated using these initial counts to reflect the same distribution as the overall dataset. This resulted in approximately 1,734 images for Toddlers/Preschoolers, 1,526 images for Children, 9,923 for Young Adults, 4,181 for Middle Age individuals, and 1,600 images for Seniors within the training set of 18,964 images.

The minimum number of images from these proportionate counts was identified to ensure that each age group was adequately represented in further splits. This base value, determined to be 1,526 images (corresponding to the Children category), represents the smallest number of samples available for any age group. To ensure equal representation of each age group, the base value was multiplied by the number of age groups, resulting in a total base value of 7,630 images. This calculation ensures that each age group contributes equally to the training process.

After accounting for the base value for each age group, the remaining number of images in the initial training set was calculated. By subtracting the total base value from the initial training set size of 18,964 images, it was determined that 11,334 images remained. These remaining images were then further split to maintain balance, with 80% of the remaining pictures allocated to a secondary training set and 20% assigned to a secondary testing set.

In practice, the implementation process involves several vital steps. First, the dataset was loaded, and the initial 80-20 split was performed to create the initial training and testing sets. Then, the proportionate counts for each age group within the initial training set were calculated. The base value was determined as the minimum count among the age groups, and the total base value for all groups was computed by multiplying the base value by the number of age groups. Subsequently, the total base value was subtracted from the initial training set size to obtain the remaining data, which was then split into secondary training and testing sets, ensuring balanced representation across all age groups.

This systematic partitioning of the dataset ensures that each age group is adequately represented, thereby preventing biases and enhancing the reliability of the model evaluation process. By maintaining a balanced distribution of classes, the model is better equipped to perform accurately and generalize well to new, unseen data.

**Step by Step Explanation of how we split the dataset into Training and Test Sets:**

**Step 1**: Initial Split (80-20)

The dataset consists of 23,708 images, initially split into two subsets: training and testing. Specifically, 80% of the photos (18,964) were allocated to the training set, and the remaining 20% (4,744) were assigned to the testing set. This initial split ensures that most of the data is used for model training while a significant portion is reserved for evaluating the model's performance on unseen data.

**Step 2**: Class Distribution in the Training Set

To maintain a balanced representation of each age group within the training set, the initial counts for each age group in the overall dataset were used to calculate their proportionate counts in the training set of 18,964 images. The age group counts from the dataset are as follows:

- Toddlers/Preschoolers (0-5): 2,167

- Children (6-10): 1,908

- Young Adults (18-35): 12,399

- Middle Age (36-60): 5,226

- Seniors (61-120): 2,008

Using these counts, the proportionate counts for each age group in the training set were calculated:

- Toddlers/Preschoolers: (2167/23708)\*18964≈1734

- Children:({1908}/{23708} )\*18964 ≈ 1526

- Young Adults: ({12399}/{23708})\*18964 ≈ 9923

- Middle Age: ({5226}/{23708})\* 18964 ≈ 4181

- Seniors:({2008}{23708}) \* 18964 ≈ 1600

**Step 3**: Determine the Minimum Number (Base Value)

The base value is determined by identifying the minimum number of images from the proportionate counts for each age group within the training set. Among the calculated counts, the minimum value is 1526 (corresponding to the Children category). This base value represents the smallest number of samples available for any age group and ensures that each group is adequately represented in further splits.

**Step 4**: Calculate Total Minimum Value for All Groups

To ensure equal representation of each age group, the base value must be multiplied by the number of age groups. Given that there are five age groups, the total minimum value (or base value for all groups) is calculated as follows:

1526 \* 5 = 7630

This minimum value ensures that each age group contributes equally to the training process.

**Step 5**: Calculate Remaining Data After Using the Base Value

After accounting for the base value for each age group, the remaining number of images in the initial training set is determined. This is calculated by subtracting the total minimum value from the initial training set size:

[ 18964 - 7630 = 11334 ]

These remaining images will be divided into secondary training and testing sets.

**Step 7**: Combine and Evaluate the Final Testing Set

The remaining data and the original testing set are combined to ensure robust evaluation. Specifically, the 11,334 remaining images are added to the 4,744 images from the initial testing set, forming a comprehensive testing set of 16,078. This combined testing set provides a more extensive and balanced dataset for evaluating the model’s performance.

**Confusion Matrix For Training Data :**

**Gender Confusion Matrix :**

The confusion matrix provides a detailed breakdown of the model's performance in gender classification:

The model achieves a high number of true positives and true negatives, indicating good performance in correctly identifying both male and female samples.

There are some misclassifications (false positives and false negatives), but these are relatively few compared to the correctly classified samples.

The derived performance metrics, such as accuracy, precision, recall, and F1 score, indicate that the model performs well in distinguishing between male and female samples, with room for improvement in reducing misclassifications.

Table 3.5.1:Confusion Matrix for Gender(Training Data)

|  |  |  |
| --- | --- | --- |
| Predicted \ Actual | Male (0) | Female (1) |
| Male (0) | 652 | 176 |
| Female (1) | 169 | 529 |

**Age Classification Confusion Matrix**

The model performs well in classifying Toddlers/Preschoolers and Seniors, with high precision and recall.

Misclassifications are more common between adjacent age groups, such as Children and Young Adults, Young Adults and Middle Age, and Middle Age and Seniors. This indicates overlapping features between these groups, leading to some ambiguity in predictions.

Overall, the confusion matrix provides a detailed view of the model's performance, highlighting strengths in certain age groups and areas for improvement in reducing misclassifications between adjacent age groups.

Table 3.5.2:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Predicted \ Actual | Toddlers/Preschoolers | Children | Young Adults | Middle Age | Seniors |
| Toddlers/Preschoolers | 266 | 57 | 0 | 2 | 2 |
| Children | 17 | 223 | 64 | 11 | 1 |
| Young Adults | 0 | 17 | 219 | 49 | 4 |
| Middle Age | 0 | 2 | 73 | 151 | 65 |
| Seniors | 0 | 0 | 4 | 79 | 220 |

**Interpretation**

The interpretation of the confusion matrix reveals critical insights into the model's performance in age classification. The model shows highly accurate favorable rates for each age group, particularly excelling in correctly identifying Toddlers/Preschoolers and Seniors. These true positives, representing correct classifications, are prominently displayed along the diagonal of the confusion matrix, indicating the model's strong ability to classify these age groups accurately. However, the matrix also highlights several areas of misclassification. For instance, Toddlers/Preschoolers are sometimes mistaken for Children, occurring 57 times, likely due to the similarities in features between these younger age groups.

Similarly, Children are often misclassified as Young Adults, with 64 such instances suggesting a significant overlap in the features the model has learned for these groups. The most considerable confusion occurs between young adults and middle-aged individuals, with 49 young adults misclassified as middle-aged and 73 middle-aged individuals misclassifying as young adults. This pattern continues with Middle Age being frequently misclassified as Seniors (65 times) and Seniors being misclassified as Middle Age (79 times), which can be attributed to standard features shared by these adjacent age groups. Despite these misclassifications, the model performs well in classifying the extremes of the age spectrum, such as Toddlers/Preschoolers and Seniors, demonstrating high precision and recall for these categories. The higher rate of misclassifications between adjacent age groups indicates that overlapping features between these groups lead to some ambiguity in predictions. Overall, the confusion matrix provides a detailed view of the model's performance, highlighting its strengths in certain age groups and identifying areas for improvement in reducing misclassifications between adjacent age groups.

**Confusion Matrix For Testing Data :**

The testing phase evaluates the model's performance on unseen data by making predictions and calculating various performance metrics. It provides insights into how well the model generalizes beyond the training dataset. The confusion matrices for gender and age group classifications help identify areas where the model performs well and needs improvement, highlighting common misclassifications and overall accuracy.

Testing the model involves evaluating its performance on a separate dataset not used during the training phase. This process helps determine how well the model generalizes to new, unseen data.

**Gender Classification Confusion Matrix**

Table 3.5.3: Confusion Matrix for Gender(Testing Data)

|  |  |  |
| --- | --- | --- |
| Predicted \ Actual | Male (0) | Female (1) |
| Male (0) | 7474 | 819 |
| Female (1) | 1111 | 6674 |

**Gender Classification Confusion Matrix Interpretation**

The gender classification confusion matrix provides a detailed breakdown of the model's performance, distinguishing between male and female samples. The matrix reveals that the model correctly identified 7,474 male samples as male and 6,674 female samples as female, demonstrating high accuracy in both true favorable (TP) and accurate adverse (TN) rates. Specifically, the true positives indicate that 7,474 male samples were accurately classified, reflecting the model's effectiveness in identifying male samples. Similarly, the true negatives show that 6,674 female samples were accurately classified, highlighting the model's proficiency in recognizing female samples.

However, the matrix also identifies areas for improvement through the false positive (FP) and false negative (FN) rates. There are 819 instances where female samples were incorrectly classified as male, indicating a need to enhance the model's ability to differentiate female samples more accurately. Additionally, there are 1,111 instances where male samples were misclassified as female, suggesting the model could be improved to distinguish male samples better. These misclassifications point to overlapping features between genders that the model needs help to differentiate, leading to some ambiguity in predictions.

The confusion matrix demonstrates that the model performs well in gender classification, with many correct classifications for both male and female samples. The areas for improvement identified through the false positive and false negative rates provide valuable insights for further refining the model to enhance its accuracy and reliability in gender identification.

**Age Group Classification Confusion Matrix**

**Table 3.5.4:** Confusion Matrix for Age Group Classification(Testing Data)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Predicted \ Actual | Toddlers/Preschoolers | Children | Young Adults | Middle Age | Seniors |
| Toddlers/Preschoolers | 527 | 104 | 6 | 1 | 3 |
| Children | 20 | 271 | 76 | 13 | 2 |
| Young Adults | 10 | 654 | 8084 | 1945 | 180 |
| Middle Age | 0 | 36 | 854 | 2018 | 792 |
| Seniors | 0 | 2 | 20 | 124 | 336 |

**Age Group Classification Confusion Matrix Interpretation**

The age group classification confusion matrix provides a comprehensive view of the model's performance across different age categories, highlighting its strengths and areas needing improvement. The matrix shows that the model accurately classifies many samples in each age group, with true positives prominently represented along the diagonal. Specifically, the model correctly identifies 527 Toddlers/Preschoolers, 271 Children, 8,084 Young Adults, 2,018 Middle Age individuals, and 336 Seniors. These highly accurate favorable rates, particularly for Young Adults and Middle Age groups, indicate that the model effectively recognizes these categories.

However, the matrix also reveals notable misclassifications, particularly between adjacent age groups with overlapping features. For example, Toddlers/Preschoolers are sometimes misclassified as Children (104 instances), and Children are often mistaken for Young Adults (76 cases). The most significant confusion occurs between young adults and middle-aged individuals, with 1,945 young adults misclassified as middle-aged and 854 middle-aged individuals misclassified as young adults. Additionally, there are 792 instances where Middle Age individuals are misclassified as Seniors and 124 cases in which Seniors are misclassified as Middle Age.

These misclassifications suggest that the model needs help distinguishing between age groups with similar characteristics, leading to prediction ambiguity. This overlap is particularly evident between Children and Young Adults, Young Adults, Middle Age and Middle Age and Seniors. Despite these challenges, the model performs well in classifying the extremes of the age spectrum, such as Toddlers/Preschoolers and Seniors, demonstrating high precision and recall for these categories.

Overall, the confusion matrix highlights the model's strong performance in correctly identifying most samples within each age group, particularly Young Adults and middle-aged individuals. However, it also points to areas where the model can be improved, specifically in reducing misclassifications between adjacent age groups with similar features. Addressing these areas can further enhance the model's accuracy and reliability in age group classification.

**3.6 Sensitivity and Specificity of Each Class Group:**

Sensitivity, also known as recall or valid positive rate, measures the proportion of actual positives the model correctly identifies. It answers, "Of all the positive cases, how many did the model correctly identify?" The formula for sensitivity is given by:

**Sensitivity = True Positives / (True Positives + False Negatives)**

Specificity, also known as the valid negative rate, measures the proportion of actual negatives the model correctly identifies. It answers, "Of all the actual negative cases, how many did the model correctly identify?" The formula for specificity is:

**Specificity = True Negatives / (True Negatives + False Positives)**

The table below presents the sensitivity and specificity of the model for each age classification group:

**Table 4: Sensitivity and Specificity of Each Age Classification**

|  |  |  |
| --- | --- | --- |
| Age Group | Sensitivity (Recall) | Specificity |
| Toddlers/Preschoolers | 81.35% | 98.58% |
| Children | 70.57% | 93.72% |
| Young Adults | 75.78% | 88.60% |
| Middle Age | 51.89% | 88.58% |
| Seniors | 72.61% | 94.11% |

- Toddlers/Preschoolers: The model exhibits a high sensitivity of 81.35% and a very high specificity of 98.58%. This indicates that the model performs exceptionally well in correctly identifying members of this age group and accurately excluding those who do not belong. The high sensitivity means that most true positives are identified, and the high specificity means that false positives are minimal.

- Children: For the Children age group, the model has a sensitivity of 70.57% and a specificity of 93.72%. These values suggest that while the model is quite good at identifying children, there is room for improvement. The sensitivity indicates that around 70% of the actual positive cases are correctly identified, while the specificity shows that the model is highly effective at correctly identifying negative cases.

- Young Adults: The Young Adults group has a sensitivity of 75.78% and a specificity of 88.60%. This reflects good overall performance, with the model accurately identifying most young adults. However, the moderate specificity suggests some overlap with other age groups, leading to more false positives than Toddlers/Preschoolers and Children.

- Middle Age: This age group has a lower sensitivity of 51.89% and a specificity of 88.58%. The relatively low sensitivity indicates that the model must correctly identify individuals in the middle-age category. While the specificity is moderate, showing the model’s capability to identify negatives, the high false negative rate suggests significant room for improvement in correctly classifying middle-aged individuals.

- Seniors: The model achieves high sensitivity and specificity for the Seniors group, with 72.61% and 94.11% values, respectively. This indicates strong performance in identifying seniors correctly, both in terms of true positives and true negatives. The high sensitivity ensures that most seniors are accurately classified, and the high specificity ensures few false positives.

In conclusion, the model demonstrates varying levels of sensitivity and specificity across different age groups. It performs exceptionally well for Toddlers/Preschoolers and Seniors, accurately identifying and excluding individuals in these categories. The model shows moderate performance for Children and Young Adults, with suitable identification rates but some misclassifications. The Middle Age group presents the most significant challenge, with lower sensitivity indicating difficulties in correct identification. These insights highlight where the model excels, and further refinements are necessary to improve overall performance.

**Chapter 4: Results & Discussion**

These results provide a comprehensive overview of the model's performance during the training and testing phases, highlighting key metrics such as accuracy, F1 score, and the confusion matrices for gender and age group classifications. The final sensitivity and specificity values for each age group offer detailed insights into the model's effectiveness in identifying different classes.

## **4.1 Results :**

## **Training Outcome**

**Gender Evaluation For Training Data :**

The gender evaluation for the training data, as presented in Table 5, demonstrates the model's performance in classifying gender with an overall accuracy of 77.39%. This indicates that approximately 77% of the gender classifications made by the model are correct. The F1 score, which balances precision and recall, is 75.41%, reflecting the model’s ability to identify and classify gender instances correctly.

The classification report provides more detailed insights into the model's performance. Precision, the proportion of true positives among all optimistic predictions, is 79% for males and 75% for females. This means that when the model predicts a male, it is correct 79% of the time, and for females, it is accurate 75% of the time. Recall, also known as sensitivity, measures the proportion of actual positives the model correctly identifies. The recall is 79% for males, indicating that the model successfully identifies 79% of the exact male instances. For females, the recall is 76%, showing a similar performance level.

The F1 score, the harmonic mean of precision and recall, is 79% for males and 75% for females. This score highlights the model’s balanced performance between precision and recall, ensuring that both are considered when evaluating its effectiveness. Overall, these results from Table 5 indicate that the model is reasonably practical in gender classification, with solid performance metrics for precision and recall, particularly for males.

**Table 5: Gender Evaluation Results for Training Data**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Gender | Precision | Recall | F1-Score | Support |
| Male (0) | 0.79 | 0.79 | 0.79 | 828 |
| Female (1) | 0.75 | 0.76 | 0.75 | 698 |
| Accuracy |  |  | 0.77 | 1526 |
| Macro Avg | 0.77 | 0.77 | 0.77 | 1526 |
| Weighted Avg | 0.77 | 0.77 | 0.77 | 1526 |

**Age Group Evaluation for Training Data:**

The age group evaluation for the training data, as detailed in Table 6, shows that the model's predictions for age have a Mean Absolute Error (MAE) of 7.76 years, indicating an average deviation of 7.76 years between predicted and actual ages. For age group classification, the model achieved an accuracy of 70.81%, meaning that about 71% of the classifications were correct. The F1 score for this task is 71.08%, balancing precision and recall.

The classification report highlights that precision is highest for Toddlers/Preschoolers at 94%, reflecting high accuracy in predictions for this group. Recall, or sensitivity, is also highest for Toddlers/Preschoolers at 81%, indicating effective identification of true positives. The F1 score, combining precision and recall, is highest for Toddlers/Preschoolers at 87%. These results suggest that the model is particularly effective in identifying Toddlers/Preschoolers while showing reasonable performance across all age groups.

**Table 6:** **Age Group Evaluation Results for Training Data**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Age Group | Precision | Recall | F1-Score | Support |
| Toddlers/Preschoolers | 0.94 | 0.81 | 0.87 | 327 |
| Children (1) | 0.75 | 0.71 | 0.73 | 316 |
| Young Adults (2) | 0.61 | 0.76 | 0.67 | 289 |
| Middle Age (3) | 0.52 | 0.52 | 0.52 | 291 |
| Seniors (4) | 0.75 | 0.73 | 0.74 | 303 |
| Accuracy |  |  | 0.71 | 1526 |
| Macro Avg | 0.71 | 0.70 | 0.71 | 1526 |
| Weighted Avg | 0.72 | 0.71 | 0.71 | 1526 |

## **Testing Outcome**

**Testing Results**

**Gender Evaluation for Testing Data:**

The gender evaluation for the testing data, as presented in Table 7, demonstrates the model's strong performance in classifying gender. The model achieved an accuracy of 87.96%, indicating that nearly 88% of the gender classifications were correct. The F1 score for gender classification is 87.37%, balancing precision and recall to provide a comprehensive measure of the model’s effectiveness.

The classification report provides further detail. Precision, which measures the proportion of true positives among all optimistic predictions, is 87% for males and 89% for females. This indicates that when the model predicts a male, it is correct 87% of the time, and for females, it is accurate 89%. Recall, or sensitivity, which measures the proportion of actual positives correctly identified by the model, is 90% for males and 86% for females. The F1 score, which combines precision and recall, is 89% for males and 87% for females, reflecting a well-balanced performance in identifying both genders accurately. These results from Table 7 highlight the model's effectiveness in gender classification, demonstrating high accuracy and robust performance metrics.

**Table 7: Gender** **Evaluation Results for Testing Data**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Gender | Precision | Recall | F1-Score | Support |
| Male (0) | 0.87 | 0.90 | 0.89 | 8293 |
| Female (1) | 0.89 | 0.86 | 0.87 | 7785 |
| Accuracy |  |  | 0.88 | 16078 |
| Macro Avg | 0.88 | 0.88 | 0.88 | 16078 |
| Weighted Avg | 0.88 | 0.88 | 0.88 | 16078 |

**Age Group Evaluation for Testing Data:**

The age group evaluation for the testing data, as detailed in Table 8, indicates that the model's predictions for age have a Mean Absolute Error (MAE) of 8.78 years. This measure shows that, on average, the predicted ages deviate from the actual ages by 8.78 years. The model achieved an accuracy of 69.88% for classifying age groups, meaning that approximately 70% of the classifications were correct. The F1 score for age group classification is 72.33%, which balances precision and recall to give a single performance measure.

The classification report reveals that precision, the proportion of true positives among all optimistic predictions, is highest for Toddlers/Preschoolers at 95%. This high precision indicates the model's accuracy in predicting this age group. Recall, or sensitivity, which measures the proportion of correctly identified positives, is also highest for Toddlers/Preschoolers at 82%. The F1 score, which combines precision and recall, is highest for Toddlers/Preschoolers at 88%. These results suggest that the model is particularly effective in identifying Toddlers/Preschoolers while showing reasonable performance across other age groups.

**Table 8:Age Group Evaluation Results for Testing Data**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Age Group | Precision | Recall | F1-Score | Support |
| Toddlers/Preschoolers | 0.95 | 0.82 | 0.88 | 641 |
| Children (1) | 0.75 | 0.71 | 0.73 | 382 |
| Young Adults (2) | 0.65 | 0.76 | 0.70 | 1873 |
| Middle Age (3) | 0.49 | 0.55 | 0.52 | 3700 |
| Seniors (4) | 0.26 | 0.70 | 0.38 | 402 |
| Accuracy |  |  | 0.70 | 16078 |
| Macro Avg | 0.57 | 0.70 | 0.59 | 16078 |
| Weighted Avg | 0.77 | 0.70 | 0.72 | 16078 |

**Comparison among the several state-of-the-art:**

**Table 9: Comparison between our approach and previous works**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Paper Title | Data Set Used | Number of Images | Age Group Classification | Age Classification Accuracy | Gender Classification Accuracy |
| Age and Gender Classification using CNN [1] | Adience | 17,856 | 0-2, 4-6, 8-13, 15-20, 25-32, 38-43, 48-53, and 60+ years | 44.26% | 85.16% |
| Gender and Age Detection using Deep Convolutional Neural Networks [2] | Adience | 26,580 | 0-2, 4-6, 8-13, 15-20, 25-32, 38-43, 48-53, and 60+ years | 61.3% | 96.2% |
| Gender Classification and Age Detection Based on Human Facial Features Using Multi-Class SVM [3] | Custom | 119 | Child - 40 Images,  Adult -40 Images, Old Aged - 39 Images | 83.33% | 90% |
| Age and Gender Classification using Multiple CNN [4] | AGFW | 36,299 | 10-14, 20-24, 30-34, 40-44, 50-64 | 72.0% | 95.7% |
| Age and Gender Prediction using Adaptive Gamma Correction [5] | Adience | 17,856 | 0-2, 4-6, 8-13, 15-20, 25-32, 38-43, 48-53, 60+ years | 61.3% | 96.2% |
| Lightweight Deep CNN for Real-Time Age and Gender Prediction [6] | Combined (Wiki, UTKFace, Adience) | 18,728 | 0-2, 4-6, 8-13, 15-20, 25-32, 38-43, 48-53, 60- 100 | 48.59% | 80.76% |
| Gender and Age Estimation Using Deep Learning [7] | Adience | 19,487 | 0-2, 4-6, 8-13, 15-20, 25-32, 38-43, 48-53, 60+ years | 57.60% | 84.20% |
| Human Age and Gender Estimation Using Facial Image Processing [8] | BUET, Adience, ORL, University of Essex | 400 (BUET), 26,580 (Adience), 400 (ORL), 4,500 (Essex) | *The Adience [16] database* consists of 26,580  images portray 2,284 individuals classified as eight-age  groups and genders.  • *The ORL face database,* which consists of the images of  Forty distinct subjects, each having ten images in varying  conditions  • *The University of Essex Face database comprises* images of 225 subjects, each with 20 photos. | 76.3% (BUET) | 86.6% (BUET) |
| Our Approach | Adience | 23,708 | 0-5,6-17, 18-40, 41-64, 65-120 | 70% | 88% |

**4.2 Discussion :**

Our approach is the best compared to several state-of-the-art methods due to its superior performance metrics and comprehensive data utilization. Leveraging a dataset of 23,708 images from the Adience database, our method achieves a remarkable accuracy of 70% in age group classification and 88% in gender classification, as detailed in Table 9. These metrics surpass many existing approaches in terms of both accuracy and the balance between precision and recall.

One of the key strengths of our approach is the effective use of a robust convolutional neural network (CNN) architecture optimized forage and gender classification tasks. For instance, the method in [1] achieves an age classification accuracy of 44.26% and a gender classification accuracy of 85.16%. Another study [2] reports an age classification accuracy of 61.3% and a gender classification accuracy of 96.2%. Compared to these, our model's accuracy rates of 70% for age and 88% for gender demonstrate significant improvements. This indicates our model's ability to generalize better across diverse age groups and gender categories.

Additionally, our method employs an 80/20 data split, ensuring that the model is trained on diverse examples while being rigorously tested on a substantial portion of the dataset. This strategy enhances the model’s learning capability and provides a robust evaluation framework, leading to high sensitivity and specificity across all age groups. By achieving higher performance metrics and utilizing an efficient data-splitting strategy, our approach is more reliable and effective for real-world applications, making it the best choice compared to existing state-of-the-art methods detailed in Table 9.

**Chapter 5: CONCLUSION**

This report comprehensively evaluated our model's performance in age and gender classification using a dataset of 23,708 images from the Adience database. Through meticulous data splitting, employing an 80/20 ratio for training and testing, we ensured a robust framework for evaluating the model’s effectiveness. Our approach demonstrates significant advancements over existing methods, achieving 70% accuracy in age group classification and 88% in gender classification, as detailed in Table 9.

The model’s performance is particularly notable in its high sensitivity and specificity across various age groups, with exceptional precision in identifying Toddlers/Preschoolers. Compared to state-of-the-art methods, our approach consistently outperforms age and gender classification tasks, indicating a superior ability to generalize and accurately predict across diverse age categories.

Our method’s robust architecture and comprehensive evaluation framework underline its reliability and effectiveness for real-world applications. The significant improvements in accuracy and balanced performance metrics highlight the potential for practical deployment in systems requiring precise age and gender classification. Future work will refine the model to enhance its performance and expand its applicability across different datasets and demographic variations.

**BIBLIOGRAPHY**

[1] P. Jeevan and N. V. Sailaja, “Age and Gender Classification using CNN,” \*2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)\*, Hyderabad, Telangana, India, pp. 1233–1237, 2022.

[2] A. Ghildiyal and E. P. Ijjina, “Gender and Age Detection using Deep Convolutional Neural Networks,” \*2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)\*, Hyderabad, Telangana, India, pp. 952–955, 2022

[3] S. Ghosh and S. K. Bandyopadhyay, “Gender Classification and Age Detection Based on Human Facial Features Using Multi-Class SVM,” \*British Journal of Applied Science & Technology\*, vol. 10, no. 4, pp. 1-15, Jul. 2015.

[4] K. R. Hassan and I. H. Ali, “Age and Gender Classification using Multiple Convolutional Neural Network,” \*IOP Conference Series: Materials Science and Engineering\*, vol. 928, pp. 032039, 2020.

[5] A. Saha, N. K. S, and N. P, “Age and Gender Prediction using Adaptive Gamma Correction and Convolutional Neural Network,” \*2023 International Conference on Computer Communication and Informatics (ICCCI)\*, Coimbatore, India, pp. 1–7, Jan. 2023.

[6] Md. N. I. Opu, T. K. Koly, A. Das, and A. Dey, “A Lightweight Deep Convolutional Neural Network Model for Real-Time Age and Gender Prediction,” \*2020 Third International Conference on Advances in Electronics, Computers, and Communications (ICAECC)\*, Bangalore, India, pp. 1–6, Dec. 2020.

[7] R. Thaneeshan, K. Thanikasalam, and A. Pinidiyaarachchi, “Gender and Age Estimation From Facial Images using Deep Learning,” \*2022 IEEE International Conference on Information Technology Research (ICITR)\*, Peradeniya, Sri Lanka, pp. 1–7, Apr. 2022.

[8] S. T. Rahman, A. Arefeen, S. S. Mridul, A. I. Khan, and S. Subrina, “Human Age and Gender Estimation using Facial Image Processing,” \*2020 IEEE Region 10 Symposium (TENSYMP)\*, Dhaka, Bangladesh, pp. 1–5, Jun. 2020.