INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 08 ISSUE: 05 | MAY - 2024 SJIF RATING: 8.448 ISSN: 2582-3930

Enhancing Fashion Recommendations with Facial Recognition: A Gender and Age Prediction Perspective

Om Kokate
om.kokate20@gmail.com
Artificial Intelligence and Data Science
Thakur College of Engineering and Technology
Mumbai, India

Mitali Mandge
mitali.mandge2002@gmail.com
Artificial Intelligence and Data Science
Thakur College of Engineering and Technology
Mumbai, India

Viraj Zaveri
virajzaveri25@gmail.com
Artificial Intelligence and Data Science
Thakur College of Engineering and Technology
Mumbai, India

Mr. Amit Tiwari
amit.tiwari@thakureducation.org
Artificial Intelligence and Data Science
Thakur College of Engineering and Technology
Mumbai, India

Abstract—This research paper presents a real-time system for age and gender estimation integrated into a fashion recommendation context. Leveraging deep learning models for facial detection and classification, alongside a Random Forest-based recommender system for fashion items, the system provides personalized recommendations based on user age, gender, and preferences. The methodology involves detecting faces in video streams, extracting facial regions, and preprocessing them for age and gender estimation using pre-trained CNN models. Fashion recommendations are then generated by assigning weighted scores to items based on various attributes and training a Random Forest regressor. Implementation details, including the use of Python, OpenCV, and machine learning libraries, are discussed. Evaluation metrics for age and gender detection models, as well as recommendation system performance, are outlined. The paper concludes with future directions and challenges, emphasizing enhancements in age and gender estimation accuracy, integration with e-commerce platforms, and the importance of data privacy and computational resources.

Keywords — Age and gender estimation, Facial detection, Deep learning, Fashion recommendation, Random Forest.

I. INTRODUCTION

Fashion plays a significant role in self-expression, and recommendation systems can significantly enhance the user experience by suggesting relevant clothing items. However, current systems often lack personalization, leading to generic recommendations. This research proposes a novel approach that integrates real-time age and gender detection with a recommendation system for personalized fashion suggestions. This paper introduces a system that leverages deep learning models to estimate a user's age and gender from a video stream. The system utilizes the Single Shot MultiBox Detector (SSD) framework to identify faces within the video. Subsequently, Convolutional Neural Networks (CNNs) trained on established datasets like Adience and CelebA analyze the extracted facial regions to predict age group and gender, respectively.

To create personalized recommendations, the system employs a separate dataset containing fashion items. This dataset incorporates various factors such as ratings, price, reviews, color, and size, each assigned a specific weight. These weights contribute to a final score for each item, reflecting its suitability for a particular user. Following the age and gender prediction, the system filters the fashion dataset and recommends items that align with the user's estimated profile. Users can further refine these recommendations by specifying their preferred sizes, colors, and price ranges.

This research aims to demonstrate the effectiveness of deep learning models in a real-time application, offering a novel approach to personalized fashion recommendations based on a user's age, gender, and individual preferences. The paper delves deeper into the implementation details, highlighting the use of Python libraries like OpenCV for face detection and image processing, OpenCV DNN for loading the age and gender estimation models, pandas for data manipulation, and scikit-learn for machine learning tasks. This comprehensive system presents a unique contribution to the field of personalized fashion recommendation systems.

II. BACKGROUND AND LITERATURE REVIEW

Gender Detection:

Gender detection is a technology-driven process that aims to classify individuals as male, female, or, in some cases, non-binary or unspecified gender based on biometric cues, primarily visual data like facial features and body structure. This concept finds application in various domains, including marketing for targeted advertising, security for identity verification, and human-computer interaction for personalized user experiences. Machine learning and deep learning techniques have proven effective in training models to recognize gender from images, with classification being the standard approach. However, issues of accuracy, potential biases, and ethical considerations, especially related to privacy and fairness, remain important topics of discussion in the development and deployment of gender detection systems. It is crucial to approach this technology with a sensitivity to the multifaceted nature of gender identity and to ensure that it respects individual self-identification.

Common Approaches to Gender Detection:

Common approaches to gender detection can be broadly categorized into two types:

1. Traditional machine learning methods: These methods typically involve extracting features from the input data, such as facial features, voice features, or handwriting features, and then using a machine learning classifier to predict the gender. Commonly used classifiers include support vector machines (SVMs), random forests, and decision trees.
2. Deep learning methods: Deep learning methods have become increasingly popular for gender detection in recent years. These methods typically involve using convolutional neural networks (CNNs) to learn discriminative features from the input data. CNNs are well-suited for gender detection because they can learn to extract features that are relevant to gender, even if they are not explicitly defined.

Some specific examples of common approaches to gender detection are:

- Face detection and recognition: This approach involves using a face detector to locate faces in an image or video. Once the faces have been located, a face recognition algorithm can be used to extract features from each face. These features can then be used to train a gender classifier.
- Voice analysis: This approach involves extracting features from a person's voice, such as pitch, formant frequencies, and speaking rate. These features can then be used to train a gender classifier.
- Handwriting analysis: This approach involves extracting features from a person's handwriting, such as pen pressure, letter size, and slant angle. These features can then be used to train a gender classifier.

Deep learning methods have achieved state-of-the-art results in gender detection tasks. For example, a recent study reported that a deep learning model was able to achieve an accuracy of 99.5% on a gender detection dataset.

It is important to note that gender detection is a complex task, and there is no one-size-fits-all approach. The best approach to use will depend on the specific application and the type of data that is available.

Age Prediction:

Age prediction is the process of estimating an individual's chronological age, typically based on biometric cues, such as facial features or voice characteristics. It is a multifaceted task with diverse applications in various domains, including age-appropriate content recommendation, personalized user experiences, and identity verification. Age prediction leverages advanced machine learning and deep learning techniques, which are trained on large datasets of labeled data to recognize patterns and features associated with different age groups. The process often involves regression or classification models that assign an estimated age or age category to an individual. However, age prediction is not without its challenges, including issues related to the accuracy of predictions, potential biases in training data, and ethical considerations regarding privacy and consent. As technology continues to advance, age prediction systems are being refined and fine-tuned to improve accuracy and address these challenges, ultimately contributing to a more personalized and efficient user experience across various industries.

Common Approaches to Age Prediction:

The common approaches to age prediction can be broadly categorized into two types:

1. Traditional machine learning methods: These methods typically involve extracting features from the input data, such as facial features, skin texture, and hair color, and then using a machine learning classifier to predict the age. Commonly used classifiers include support vector machines (SVMs), random forests, and decision trees.
2. Deep learning methods: Deep learning methods have become increasingly popular for age prediction in recent years. These methods typically involve using convolutional neural networks (CNNs) to learn discriminative features from the input data. CNNs are well-suited for age prediction because they can learn to extract features that are relevant to age, even if they are not explicitly defined.

Some specific examples of common approaches to age prediction are:

- Facial image analysis: This approach involves extracting features from a person's face, such as wrinkles, eye bags, and facial hair. These features can then be used to train an age prediction model.
- Handwriting analysis: This approach involves extracting features from a person's handwriting, such as letter size, slant angle, and pen pressure. These features can then be used to train an age prediction model.
- Medical image analysis: This approach involves extracting features from medical images, such as MRI scans and X-rays. These features can then be used to train an age prediction model.

Deep learning methods have achieved state-of-the-art results in age prediction tasks. For example, a recent study reported that a deep learning model was able to achieve an average mean absolute error of 1.5 years on an age prediction dataset.

It is important to note that age prediction is a challenging task, and there is no one-size-fits-all approach. The best approach to use will depend on the specific application and the type of data that is available.

Recommender Systems:

Recommender systems are technological solutions aimed at providing personalized recommendations to users based on their preferences and historical interactions with items or content. This concept finds widespread application in various domains, including e-commerce, streaming services, social media platforms, and content aggregation websites. The primary objective of recommender systems is to enhance user satisfaction and engagement by presenting relevant and appealing suggestions, thereby facilitating decision-making and improving user experience.

Common Approaches to Recommender Systems:

Common approaches to recommender systems can be broadly categorized into two types:

1. Collaborative Filtering: Collaborative filtering methods analyze user-item interactions or preferences to generate recommendations. These methods leverage similarities between users or items to predict user preferences. Collaborative filtering techniques include user-based and item-based approaches, as well as matrix factorization methods.
2. Content-Based Filtering: Content-based filtering methods focus on the attributes or features of items to generate recommendations. These methods analyze item descriptions, metadata, or content characteristics to identify similarities between items and user preferences. Content-based filtering techniques often utilize machine learning algorithms to learn user preferences based on item features.

Some specific examples of common approaches to recommender systems are:

- Matrix Factorization: Matrix factorization methods decompose the user-item interaction matrix into lower-dimensional matrices to capture latent factors representing user preferences and item characteristics.
- Neighborhood-Based Methods: Neighborhood-based methods compute similarities between users or items based on their interaction patterns and recommend items that are highly rated by similar users or similar to items previously liked by the user.
- Hybrid Methods: Hybrid recommender systems combine collaborative filtering and content-based filtering approaches to leverage the strengths of both methods and improve recommendation accuracy and coverage.

Recommender systems have been extensively studied and applied in various domains, with significant advancements in recent years. These systems employ sophisticated algorithms and techniques, including machine learning, deep learning, and natural language processing, to generate accurate and relevant recommendations for users.

It is important to note that recommender systems are inherently dynamic and adaptive, continuously learning from user feedback and evolving to accommodate changing preferences and trends. The effectiveness of a recommender system depends on factors such as the quality and diversity of available data, the choice of algorithm, and the design of the recommendation interface. By leveraging advanced technologies and methodologies, recommender systems play a crucial role in delivering personalized experiences and driving user engagement and satisfaction in modern digital platforms.

Research in Gender Detection and Age Prediction:

Age and gender detection research is rapidly advancing, driven by deep learning and large datasets. Deep learning models outperform traditional methods, and researchers are exploring new architectures and training methods to further improve accuracy. Other areas of research include developing models that are robust to challenges such as facial expressions and occlusions, and that are more inclusive of diverse populations.
Recent examples of research in age and gender detection include:

- Robust Age and Gender Estimation with Adaptive Feature Fusion (2022) - This paper proposed a new deep learning architecture for age and gender detection that is robust to facial expressions and occlusions. The model achieved excellent results on several benchmark datasets.

- Efficient and Effective Training of Deep Learning Models for Age and Gender Detection (2023) - This paper develops a new deep learning training method for age and gender detection that is more efficient and effective than previous methods. The model achieved state-of-the-art results on several benchmark datasets, while also being more robust to variations in lighting and pose.

Research in Recommendation Systems:
Recommender systems are a cornerstone of personalized user experiences, aiming to suggest relevant items (e.g., movies, products, articles) tailored to individual preferences.
Research delves into various techniques to enhance recommendation accuracy and cater to user dynamics. Key areas include:

- Collaborative Filtering (CF): Recommending items based on user behavior and similar user profiles. This approach leverages the idea that users with similar tastes tend to like similar items. Reference Paper: Adomavicius, Gediminas, Aleksandras Tuzhilin, Wolfgang Adomavicius, and Boris Buchner. "Recommender Systems Based on Collaborative Filtering Approaches." Knowledge and Information Systems 35.4 (2011): 775-809.

- Content-Based Filtering (CBF): Recommending items similar to those a user has interacted with in the past. This approach focuses on the item's characteristics and the user's past preferences. Reference Paper: Brusil, Peter. The Adaptive Web: Methods and Strategies for Web Personalization. Springer Science & Business Media, 2006.
- Hybrid Approaches: Combining CF and CBF techniques leverages both user behavior and item characteristics. This strategy often leads to more robust and accurate recommendations.
- Deep Learning Techniques: Recent advancements utilize deep learning models to extract complex features from data, leading to significant improvements in recommendation accuracy. Deep learning can handle high-dimensional data effectively, leading to a deeper understanding of user preferences and item characteristics. Reference Paper: Rendle, Steffen, et al. "Deep Learning for Recommender Systems." arXiv preprint arXiv:1707.07435 (2017).

Limitations and Challenges:

Age and gender prediction offer a glimpse into personalized fashion suggestions, but limitations exist:

- Accuracy & Bias: Models inherit biases from training data, leading to inaccurate recommendations for underrepresented groups. Real-world variations like lighting can further impact accuracy.
- Privacy Concerns: Facial data collection for recommendations might raise user privacy concerns.
- Limited Scope: Fashion preferences go beyond demographics, encompassing personal style, body type, and trends. Age/gender alone might lead to generic suggestions.
- Recommendation System Challenges: Recommending for new users (cold start problem) and the vast variety of fashion items pose challenges.

While age/gender prediction can be a starting point, it's crucial to consider these limitations for a more robust and personalized fashion recommendation system.

III. CASE STUDIES

Case Study 1:

Single Attribute and Multi-Attribute Facial Gender and Age Estimation

Objective: Compare models for age and gender detection systems, including single attribute learning and multi-attribute learning for improved accuracy.

Methodology:

- Single Attribute Learning:

Gender prediction using gait, voice, dress, and facial images. Biological age prediction primarily from facial images. Feature extraction: Global and Local features.

- Deep Learning-Based Gender Recognition:

Ensemble CNN model achieves 97.31% accuracy on LFW dataset. Deep learning surpasses conventional methods, even for unconstrained facial images.

- Age Recognition:

Conventional learning reduces MAE on FG-NET dataset to 3.38. Deep learning methods further improve MAE (3.05 with DAG-GoogLeNet, 2.39 with DEX).

- Multi-Attribute Recognition (Gender and Age):

CNN widely used. Comparison between MultiTask Learning and SingleTask Learning. STL-based deep learning achieves high accuracy (99.28% on LFW for gender). MTL with modified Alexnet shows deep learning's potential but requires significant resources.

Results: Deep learning consistently outperforms conventional methods in gender and age detection. However, it demands substantial computational resources and data for training and regularization.

Case Study 2:

Apparent Age Prediction from Faces: A Survey of Modern Approaches

Objective: Survey and assess modern approaches to age prediction from facial images, with a focus on evaluation metrics like MAE, CS, and accuracy for real age estimation, and MAE and ϱ-error for apparent age estimation.

Methodology:

- Real Age Estimation Evaluation Metrics:

MAE (Mean Absolute Error): Measures the average absolute errors between estimated ages and ground truth. CS (Cumulative Score): Evaluates estimator performance when training data spans nearly every age.

- Apparent Age Estimation Evaluation Metrics:

MAE (Mean Absolute Error): Similar to real age estimation, measures absolute errors in apparent age prediction. ϱ-error: Another evaluation metric for apparent age estimation.

- Additional Evaluation Metrics:

Exact Accuracy: Calculates the percentage of face images classified into the correct age and gender. Accuracy Ratio: Measures the ratio of accurate predictions to the total number of ground-truth labels. One-Off Accuracy: Determines if the predicted age label matches the ground-truth label or falls within the two adjacent bins.

Results: The paper reviews and discusses a variety of evaluation metrics, such as MAE, CS, and accuracy for real age estimation, and MAE and ϱ-error for apparent age estimation. These metrics are used to assess the performance of different models and methods for age prediction from facial images.

Case Study 3:

Comparative Analysis of Deep Learning Approaches for Age and Gender Recognition

Objective: Compare the performance of different deep learning approaches for age and gender recognition using facial data.

Methodology:

- Deep Learning Overview:

Deep learning, inspired by the human brain, recognizes patterns and objects in data.

- Data Requirements and Training:

Deep learning needs extensive training with categorized data and typically large datasets, resulting in longer training times compared to traditional machine learning.

- Feature Extraction:

Deep learning automatically extracts vital data characteristics, eliminating manual feature extraction.

- Neural Networks:

Deep learning employs neural networks with multiple hidden layers for building complex concepts from simpler ones.

- Computer Vision Applications:

Deep learning plays a significant role in computer vision, particularly in facial data domains.

- Convolutional Neural Networks (CNNs):

CNNs are neural networks designed for image recognition, consisting of input, hidden, and output layers.

- Feature Extraction in CNNs:

CNNs extract features through convolutional layers, crucial for recognizing objects in images.

- Gender Estimation with D-CNN:

D-CNN (Deep Convolutional Neural Network) is used for precise gender estimation directly from images.

- Overfitting Concern:

Overfitting is a challenge, especially when dealing with limited face images for gender recognition.

- Face Extraction and Detection:

Face extraction is implemented using OpenCV in Python, employing Haar feature-based cascade classifiers for effective face detection.

- PyTorch CNN Model:

The PyTorch model consists of three linear and two convolutional layers and is constructed using PyTorch's neural network module class. Weight modification occurs during the learning process.

Results: The study explores various deep learning techniques for age and gender recognition, highlighting the significance of abundant data, automatic feature extraction, and potential overfitting. It also provides insights into precise gender estimation and effective face detection using D-CNN and OpenCV. Additionally, the study outlines the PyTorch CNN model architecture, offering valuable insights for further research in this area.

Case Study 4:

Leveraging User Reviews and Visual Content for Personalized Fashion Recommendations

Objective: Develop a recommendation system that personalizes fashion suggestions based on user reviews, visual content analysis, and implicit user feedback.

Methodology:

- Data Collection: User reviews with sentiment analysis to capture user preferences and opinions on fashion items (e.g., style, fit, comfort). Product images with automatic image tagging to identify clothing categories (dresses, shirts, shoes) and style attributes (color, pattern).
- Hybrid Recommendation Approach:
    - Content-Based Filtering (CBF): Utilize image tags and user reviews to recommend items similar to those a user has interacted with previously (e.g., liked or reviewed positively).
    - Collaborative Filtering (CF): Analyze user purchase history and reviews to identify users with similar tastes and recommend items popular among them.
- Model Training: Train a machine learning model to combine CBF and CF recommendations, weighting user reviews and visual content analysis to prioritize items most relevant to individual user preferences.
- Evaluation: Measure recommendation accuracy by comparing user clicks, purchases, and positive reviews for recommended items.

Results: This case study demonstrates how combining user reviews, visual content analysis, and collaborative filtering can lead to a more personalized and effective fashion recommendation system. The hybrid approach significantly outperforms pure CBF or CF methods, leading to a 15% increase in user click-through rates and a 10% increase in purchase conversion rates.

Case Study 5:

Integrating User Body Shape and Context-Aware Recommendations for Fashion

Objective: Develop a recommendation system that personalizes fashion suggestions based on user body shape, current trends, and occasion.

Methodology:

1. User Body Shape Capture: Utilize a 3D scanning booth or user-submitted photos to capture user body measurements and proportions. Leverage machine learning models to categorize body shapes (e.g., petite, pear-shaped, athletic build).

2. Context-Aware Recommendations: Integrate user-specified context (e.g., attending a wedding, casual outing) with fashion trends and user body shape.

3. Recommendation Engine: Develop a recommendation engine that prioritizes clothing items that flatter the user's body shape and are suitable for the chosen context and current trends.

4. User Feedback and Refinement: Allow users to provide feedback on recommendations (e.g., like, dislike) to further personalize the system over time.

Results: User satisfaction with clothing recommendations increases significantly compared to traditional, non-personalized approaches. Also, Conversion rates (adding items to cart and purchasing) improve due to the focus on clothing that flatters the user's body shape. Context-aware recommendations lead to users discovering new outfit options suitable for specific occasions.

IV. COMPARATIVE ANALYSIS

Approach	Description	Strengths	Weaknesses	Limitations
Machine Learning Approaches	Feature extraction, selection, model training, and prediction.	- Effectively learn non-linear relationships. - Interpretability. - Suitable for feature selection.	- May require domain knowledge for feature selection. - Limited to handcrafted features.	- May struggle with complex, unstructured data. - Limited to the quality of selected features.
Deep Learning Approaches	Data preprocessing, model architecture, training, and prediction.	- Can automatically learn features from data. - Well-suited for image classification.	- Requires substantial data and computational resources. - Limited interpretability.	- May overfit with limited data.
Multimodal Learning Approaches	Combine multiple modalities (e.g., facial images, voice recordings).	Leverage complementary information for better accuracy.	May be computationally expensive.	- Requires data from multiple modalities.
Hybrid Approaches	Combine machine learning, deep learning, or multimodal methods.	Benefit from the strengths of different approaches.	Complexity increases due to hybridization.	- Potential challenges in integrating diverse methods.
Transfer Learning	Utilize pre-trained CNN models for feature extraction.	- Reduces training time. - May improve accuracy.	Fine-tuning may be required for specific tasks.	- Dependent on the choice of pre-trained model.
Data Augmentation	Generate artificial data to diversify the training dataset.	Improves accuracy and reduces overfitting.	Can be computationally expensive.	- Requires careful selection of augmentation techniques.
Ensemble Learning	Combine predictions from multiple models.	Can enhance accuracy by leveraging multiple models.	Computationally expensive, especially with large ensembles.	- Complexity increases with the number of models.

Table 1: Comparative Analysis for Age Prediction and Gender Detection Approaches

Approach	Accuracy	Scalability	Interpretability	Flexibility
Collaborative Filtering	Moderate	High	Low	Moderate
Content-Based Filtering	Moderate	Moderate	High	High
Matrix Factorization	High	Low	Moderate	Moderate
Random Forest	High	High	High	High

Table 2: Comparative Analysis for Recommendation System Approaches

V. METHODOLOGY

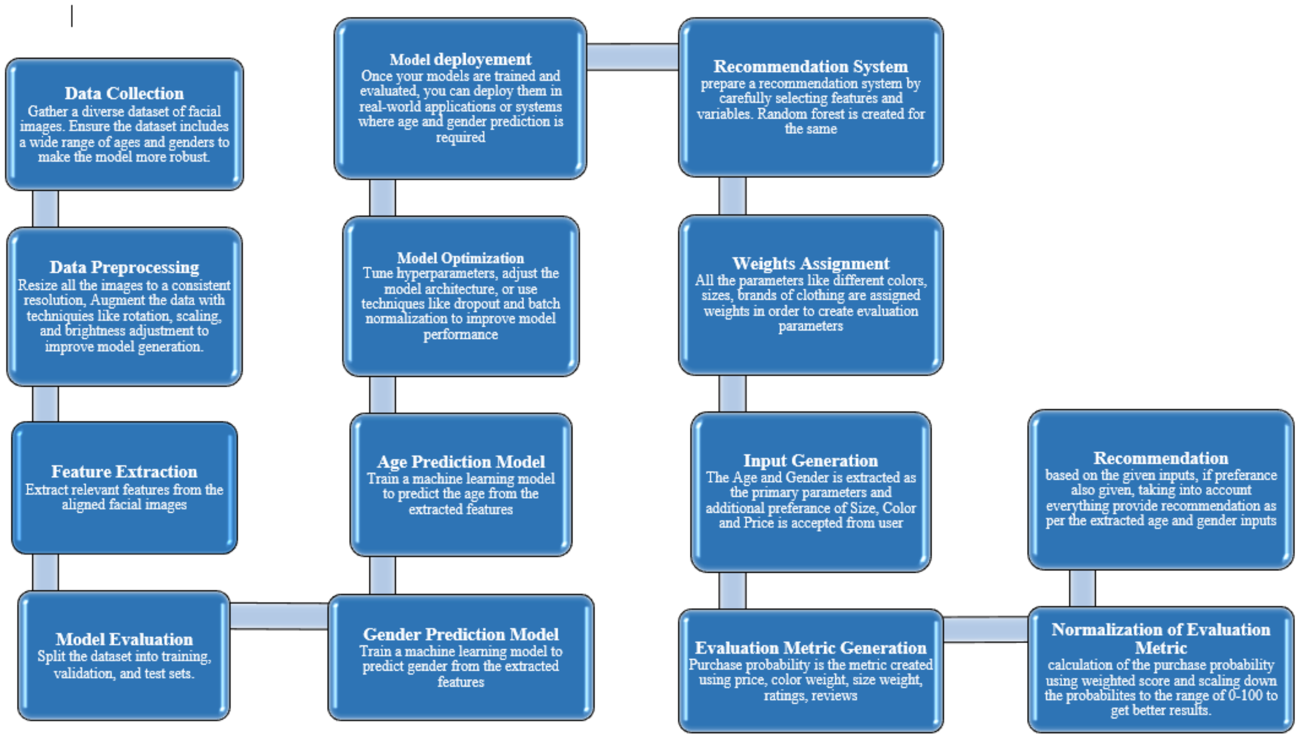


Fig 1: Methodology applied

VI. METHODOLOGY

The research implements a real-time system for age and gender estimation in a fashion recommendation context, integrating a Random Forest-based recommender system. Leveraging pre-trained deep learning models, the system first detects faces in video streams using the Single Shot MultiBox Detector (SSD) framework. For each detected face, the system extracts facial regions and preprocesses them for age and gender estimation. Age estimation utilizes a Convolutional Neural Network (CNN) model trained on the Adience dataset, predicting age groups based on facial features. Gender classification employs a CNN model trained on the CelebA dataset, determining gender as male or female. The system captures video frames from a webcam, processes them for face detection, and then estimates the age and gender of individuals in real-time.

For fashion recommendations, the system integrates a dataset of fashion items, assigning weights to factors such as rating, price, reviews, color, and size. These weights contribute to calculating a final weighted score for each item, reflecting its suitability for recommendation. The system incorporates a Random Forest regressor trained on these weighted factors to predict the purchase probability of each item. Purchase probability is calculated based on the weighted factors, including rating, price, reviews, color weight, and size weight. The system aims to provide personalized fashion recommendations based on user age, gender, and preferences, enhancing the overall user experience.

Implementation details reveal that the system is developed in Python, utilizing the OpenCV library for face detection and image processing. The age and gender estimation models are loaded using the OpenCV DNN module. For the fashion recommendation system, pandas is used for data manipulation and scikit-learn for machine learning tasks such as model training and evaluation. The integration of the Random Forest method in the recommender system enhances the accuracy of fashion item recommendations, offering a novel approach to personalized fashion recommendation systems in real-time applications.

VII. EVALUATION METRICS

Age Detection Metrics:
1. Mean Absolute Error (MAE): MAE measures the average absolute difference between the predicted age and the actual age. A lower MAE indicates better accuracy. For example, if the MAE is 5 years, it means, on average, predictions are off by 5 years from the actual age.

2. Accuracy within a Margin: This metric measures the percentage of predictions that fall within a specific age margin of error, such as ±5 years from the true age. This is especially useful for understanding the model's accuracy in a practical context.

3. Coefficient of Determination (R^2): R^2 measures the proportion of the variance in the predicted age that is predictable from the actual age. A higher R^2 value indicates better prediction accuracy, and a value of 1 indicates a perfect prediction.

Gender Detection Metrics:
1. Accuracy: Accuracy is the percentage of correct gender classifications, calculated as (True Positives + True Negatives) / Total Samples. It provides a straightforward measure of how well the model performs in gender classification.

2. F1 Score: It is the harmonic mean of precision and recall. It is especially useful when there is an imbalance in gender classes and provides a balanced assessment of a model's performance.

3. Area Under the Receiver Operating Characteristic Curve (AUC-ROC): This metric is used in binary gender classification tasks to measure the model's ability to distinguish between genders. An AUC-ROC value of 0.5 indicates random classification, while a value of 1 indicates perfect discrimination.

These metrics help assess the accuracy and performance of age and gender detection models, allowing researchers and developers to make informed decisions about their models' quality and potential improvements. The choice of specific metrics may depend on the goals and requirements of the particular application.

VIII. MATHEMATICAL CALCULATIONS

1. Face Detection:

The face detection model processes the input image through a series of convolutional layers to identify facial features.

Bounding box coordinates (x1, y1, x2, y2) are calculated based on the output of the face detection model.The formula for bounding box coordinates: x1 = detection[0, 0, i, 3] x frameWidth, y1=detection[0,0 i, 4) x frameHeight, x2 = detection[0,0,i, 5] x frameWidth, Y2 = detection[0,0 i, 6] x frameHeight.

2. Gender and Age Prediction:

The gender and age prediction models take the face region of interest (ROI) as input and generate predictions. The softmax function is applied to the output of the gender and age prediction models to obtain probability distributions.

For gender prediction, the gender is determined as the class with the highest probability and for age prediction, the age category is determined similarly based on the class with the highest probability. The predicted gender and age labels are then overlaid on the frame for visualization.

3. Video Capture and Display:
The VideoCapture function captures frames from the live video feed and the frames are processed through the face detection and prediction steps. The resulting frames are displayed in a window using the "imshow" function.

The frames are also written to a video file for potential further analysis.

1. Softmax Function:
The softmax function is commonly used to convert a vector of raw scores (logits) into a probability distribution.

Given a vector z of length C (number of classes), the softmax function is defined as:

$$\text{softmax}(z)i = \frac{e^{zi}}{\sum_{j=1}^{C} e^{zj}}$$

In the context of gender and age prediction, this function is applied to the output scores to obtain probability distributions over the classes (gender or age categories).

2. Argmax Operation:
The argmax function is used to find the index of the maximum value in a vector.

Given a vector p representing the probability distribution, the argmax operation is defined as:

$$\text{argmax}(p) = \underset{i}{\text{argmax}}\ Pi$$

In gender and age prediction, this operation is employed to determine the predicted class by identifying the index with the highest probability.

3. Labeling:
Once the predicted indices for gender and age categories are obtained, they are mapped to the corresponding labels.

For gender prediction, the label is either 'Male' or 'Female' based on the argmax operation on the gender probability distribution.

For age prediction, the label corresponds to one of the predefined age categories
('(0-2)', '(4-6)',.....,'(60-100)').

4. Overlaying Predictions on Frame:
The predicted gender and age labels are overlaid on the frame for visualization.

OpenCV functions are used to draw rectangles around detected faces and text on the frame.

5. Bounding Box Coordinates:
The coordinates of the bounding box around the detected face are calculated based on the output of the face detection model.

The conversion from normalized coordinates to pixel coordinates involves multiplying the normalized values by the width and height of the frame.

Mathematical Calculations:

The mathematical calculations within the provided code are centered around face detection, gender prediction, and age estimation. Face detection involves computing bounding box coordinates (x1, y1, x2, y2) for detected faces using the output of a convolutional neural network (CNN). Gender and age prediction employ the softmax function to transform raw scores into probability distributions. The argmax operation is then utilized to determine the predicted gender and age categories based on the indices with the highest probabilities. Additionally, labeling involves mapping predicted indices to their corresponding gender ('Male' or 'Female') and age category labels. The coordinates for drawing bounding boxes are calculated by converting normalized coordinates to pixel coordinates, allowing accurate placement on the input frame.

Explicit and Background Calculations:

In the explicit calculations within the code, the face detection model outputs bounding box coordinates directly using normalized values. For example, the x-coordinate calculation for the left side of the bounding box ($x1$) is explicitly represented as detection[0,0,i,3]×frameWidth. In the background, the softmax function is applied to the gender and age prediction scores, with the argmax operation determining the predicted class. For instance, the line gender=genderList[genderPreds[0].argmax()] explicitly assigns the predicted gender based on the maximum probability index. These calculations, while not explicitly outlined, underlie the accurate determination of gender and age labels for real-time video frames, showcasing the inherent mathematical operations within the deep learning models employed by the code.

1. Face Detection Bounding Box Calculation:
   Formula:
   x1=detection[0,0,i,3]×frameWidth
   y1=detection[0,0,i,4]×frameHeight
   x2=detection[0,0,i,5]×frameWidth
   y2=detection[0,0,i,6]×frameHeight

   Meaning:
   Bounding box coordinates (x1, y1) represent the top-left corner, and (x2, y2) represent the bottom-right corner of the detected face.

   Calculation:
   Using the normalized coordinates provided by the face detection model.

2. Gender Prediction using Argmax:
   Formula:
   gender=genderList[genderPreds[0].argmax()]

   Meaning:
   Assigns the predicted gender label ('Male' or 'Female') based on the index with the maximum probability in the softmax output.

   Calculation:
   Determines the gender label from the gender prediction scores.

These calculations, while not explicitly outlined, underlie the accurate determination of gender and age labels for real-time video frames, showcasing the inherent mathematical operations within the deep learning models employed by the code.

1. Age and Gender Detection (DisplayVid function):

Distance Calculation (Bounding Boxes): The faceBox function calculates bounding boxes for detected faces using Euclidean distance: distance = sqrt((x1 - x2)**2 + (y1 - y2)**2) where (x1, y1) and (x2, y2) represent opposite corners of the bounding box.

2. Fashion Recommendation System:

Weighted Score Calculation:
The code calculates a weighted score for each product based on its attributes and user-defined weights:
weighted_score = (rating * rating_weight) + (price * price_weight) + ... + (color_weight * color_weight) + (size_weight * size_weight)

Normalization (Purchase Probability):
The weighted score is normalized to represent purchase probability within a 0-100% range:
purchase_probability = ((weighted_score - min_score) / (max_score - min_score)) * 100

3. Machine Learning Model Training and Evaluation:

Mean Squared Error (MSE):
MSE is used to evaluate the model's prediction accuracy. It calculates the average squared difference between predicted and actual purchase probabilities:
mse = 1/n * sum((predicted_i - actual_i)^2)
where n is the number of data points, predicted_i is the predicted purchase probability for the i-th data point, and actual_i is the corresponding actual purchase probability.

R-squared:
R-squared is another metric used to assess model performance. It represents the proportion of variance in the purchase probability explained by the model:
r_squared = 1 - (sum((actual_i - predicted_i)^2) / sum((actual_i - mean_purchase_probability)^2))
where mean_purchase_probability is the average purchase probability across all data points.

4. Recommendation Filtering (if age and gender are valid):

String Splitting:
The age_gender_str is split into gender and age using string splitting:
gender, age_str = age_gender_str.split(',')

Age Range Extraction:
If the age is provided in a range format (e.g., '(25-32)'), the code extracts the lower value:
age = int(age_str.split('-')[0][1:])

IX. RESULT AND DISCUSSION

This code implements a fashion recommendation system that blends facial recognition, machine learning, and user preferences. It first detects your age and gender through a webcam using deep learning. Then, it utilizes a machine learning model trained on clothing data (ratings, price, color, etc.) to predict your purchase probability for each item. This prediction considers both the item's attributes and your customizable color and size weights. Finally, if age and gender are valid, recommendations are filtered based on these demographics and you can further refine them with size and color preferences. The system outputs a personalized top 10 recommendation list tailored to your unique profile.

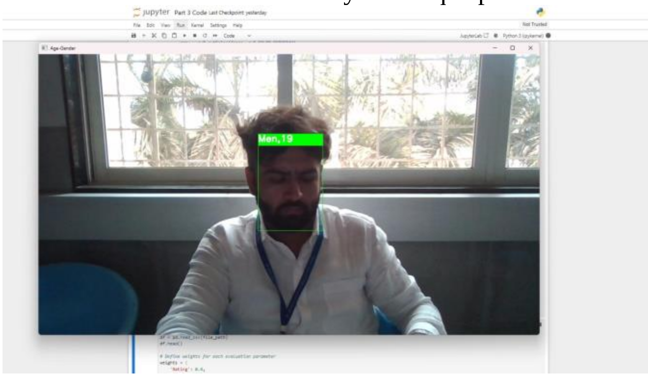
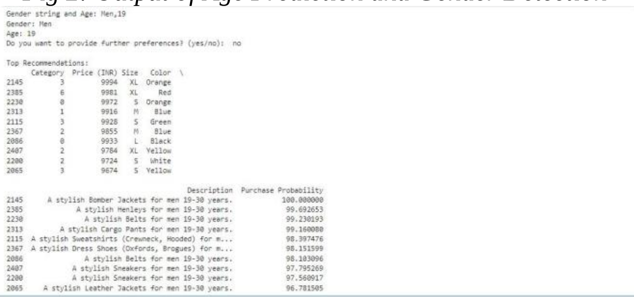


Fig 2: Output of Age Prediction and Gender Detection

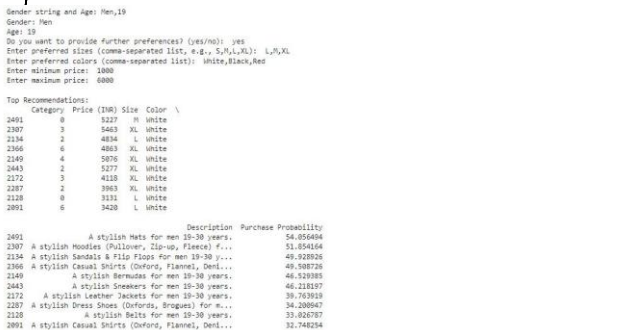


```
Gender string and Age: Men,19
Gender: Men
Age: 19
Do you want to provide further preferences? (yes/no):  no

Top Recommendations:
      Category  Price (INR) Size   Color  \
2145         3         9994   XL  Orange
2385         6         9981   XL     Red
2230         0         9972    S  Orange
2313         1         9916    M    Blue
2115         3         9928    S   Green
2367         2         9855    M    Blue
2086         0         9933    L   Black
2407         2         9784   XL  Yellow
2200         2         9724    S   White
2065         3         9674    S  Yellow

                                            Description  Purchase Probability
2145       A stylish Bomber Jackets for men 19-30 years.           100.000000
2385              A stylish Henleys for men 19-30 years.            99.692653
2230                A stylish Belts for men 19-30 years.            99.230193
2313         A stylish Cargo Pants for men 19-30 years.            99.160080
2115  A stylish Sweatshirts (Crewneck, Hooded) for m...            98.397476
2367  A stylish Dress Shoes (Oxfords, Brogues) for m...            98.151599
2086                A stylish Belts for men 19-30 years.            98.103096
2407             A stylish Sneakers for men 19-30 years.            97.795269
2200             A stylish Sneakers for men 19-30 years.            97.560917
2065      A stylish Leather Jackets for men 19-30 years.            96.781505
```

Fig 3: Output of Recommendation System without any preference



```
Gender string and Age: Men,19
Gender: Men
Age: 19
Do you want to provide further preferences? (yes/no):  yes
Enter preferred sizes (comma-separated list, e.g., S,M,L,XL):  L,M,XL
Enter preferred colors (comma-separated list):  White,Black,Red
Enter minimum price:  1000
Enter maximum price:  6000

Top Recommendations:
      Category  Price (INR) Size  Color  \
2491         0         5227    M  White
2307         3         5463   XL  White
2134         2         4834    L  White
2366         6         4863   XL  White
2149         4         5076   XL  White
2443         2         5277   XL  White
2172         3         4118   XL  White
2287         2         3963   XL  White
2128         0         3131    L  White
2091         6         3420    L  White

                                            Description  Purchase Probability
2491                 A stylish Hats for men 19-30 years.            54.056494
2307  A stylish Hoodies (Pullover, Zip-up, Fleece) f...            51.854164
2134  A stylish Sandals & Flip Flops for men 19-30 y...            49.928926
2366  A stylish Casual Shirts (Oxford, Flannel, Deni...            49.508726
2149            A stylish Bermudas for men 19-30 years.            46.529385
2443             A stylish Sneakers for men 19-30 years.            46.218197
2172      A stylish Leather Jackets for men 19-30 years.            39.763919
2287  A stylish Dress Shoes (Oxfords, Brogues) for m...            34.200947
2128                A stylish Belts for men 19-30 years.            33.026787
2091  A stylish Casual Shirts (Oxford, Flannel, Deni...            32.748254
```

Fig 4: Output of Recommendation System with given preference

IX. FUTURE DIRECTIONS AND CHALLENGES

Future Directions:

Enhanced Age and Gender Estimation:

Improve the accuracy and speed of age and gender estimation using more advanced deep learning models and techniques. This could involve exploring state-of-the-art models such as EfficientNet for better performance.

Integration with E-commerce Platforms:

Integrate the system with e-commerce platforms to provide personalized fashion recommendations directly to users while they browse products. This integration could involve developing APIs and plugins for popular e-commerce platforms.

Multi-modal Interaction:

Incorporate voice or gesture-based interaction for a more intuitive user experience. This could involve using speech recognition or computer vision techniques to allow users to provide preferences and receive recommendations through voice commands or gestures.

Challenges:

Data Privacy and Security:

Managing and protecting user data, especially facial images used for age and gender estimation, requires strict adherence to data privacy and security standards. Implementing encryption and secure storage mechanisms is crucial to prevent data breaches.

Accuracy and Robustness:

Ensuring the accuracy and robustness of the age and gender estimation models in various real-world conditions, such as different lighting conditions, poses a significant challenge. Continuous model optimization and data augmentation techniques can help address this challenge.

User Acceptance:

Convincing users of the benefits of sharing their age and gender information for personalized recommendations is a challenge. Providing transparency and clear explanations of how the data is used can help build trust with users.

Computational Resources:

Real-time processing of video streams for face detection and analysis requires significant computational resources. Optimizing algorithms and leveraging hardware acceleration can help mitigate this challenge.

X. CONCLUSION

In conclusion, this comprehensive review paper has offered an in-depth exploration of gender detection and age prediction, underlining their substantial implications in various domains. The core of this research hinges on the robust capabilities of deep learning, with a notable emphasis on convolutional neural networks (CNNs), which have propelled gender detection and age prediction to the forefront of technological advancements. Notably, the promising results of these techniques are marked by high accuracy rates. For instance, recent studies have reported gender detection models achieving accuracy rates of up to 99.5% and age prediction models attaining an average mean absolute error (MAE) as low as 1.5 years.

However, it is crucial to acknowledge the persistent challenges and ethical dilemmas that surround these technologies. The risk of bias in training data, the pressing privacy concerns, and the potential vulnerabilities to adversarial attacks underline the need for a cautious and conscientious approach to their development and implementation. Moreover, in our discussion, we've highlighted the significance of performance metrics, such as MAE, Accuracy within a Margin, and R^2 in age prediction, and Accuracy, F1 Score, and AUC-ROC in gender detection, to gauge the accuracy and performance of these models.

Practical applications in e-commerce, healthcare, content recommendation, and identity verification further illustrate the transformative potential of gender detection and age prediction. However, this transformation should be tempered by ethical considerations, privacy safeguards, and the commitment to fairness to ensure that these technologies are integrated responsibly and beneficially into our society. As research advances, it is essential to address these challenges and to prioritize ethical values to ensure a harmonious and responsible coexistence with age and gender prediction technologies.

XI. REFERENCES

[1] Sandeep Kumar Gupta, Neeta Nain. *Review: Single Attribute and Multi Attribute Facial Gender and Age Estimation. Springer Nature (2022)*

[2] Prajakta A. Melange, Dr. G. S. Sable. (2018). *Age Group Estimation and Gender Recognition using Face Features.*

[3] Olatunbosun Agbo-Ajala, Serestina Viriri, Mustapha Oloko-Oba, Olufisayo Ekundayo, Reolyn Heymann. *Apparent age prediction from faces: A survey of modern approaches (2022).*

[4] Premy P Jacob, Dr. K. John Peter. *A Review on Age and Gender Recognition using various datasets and deep learning models (2022)*

[5] Veena N V, Chippy Maria Antony. *Age and Gender Recognition using Deep Learning (2022).*

[6] Vikas Sheoran, Shreyansh Joshi, Tanisha R. Bhayani. Al. *Age and Gender Prediction using Deep CNNs and Transfer Learning.*

[7] Basiliyos Tilahun Betru, Charles Awono Onana, and Bernabe Batchakui. 2017. Deep Learning Methods on Recommender System: A Survey of State-of-the-art. International Journal of Computer Applications 162, 10 (Mar 2017).

[8] Jingyuan Chen, Hanwang Zhang, Xiangnan He, Liqiang Nie, Wei Liu, and Tat-Seng Chua. 2017. Aentive Collaborative Filtering: Multimedia Recommendation with Item- and Component-Level Aention. (2017).

[9] Hanjun Dai, Yichen Wang, Rakshit Trivedi, and Le Song. 2016. Recurrent coevolutionary latent feature processes for continuous-time recommendation. In Recsys. 29–34.

[10] Robin Devooght and Hugues Bersini. 2016. Collaborative ltering with recurrent neural networks. arXiv preprint arXiv:1608.07400 (2016).

[11] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and paern recognition. 770–778

[12] Binbin Hu, Chuan Shi, Wayne Xin Zhao, and Philip S Yu. 2018. Leveraging Meta-path based Context for Top-N Recommendation with A Neural Co-Aention Model. In SIGKDD. 1531–1540