

Gender Classification

Data Science Mini project

January 29, 2024



Rameez Akther M

2022510025

B.Tech AI & DS

Anna University

MIT Campus

Chromepet

# Problem Statement

Develop a gender detection system using deep learning and transfer learning techniques to classify facial images as either "man" or "woman." The goal is to create an accurate and robust model capable of analyzing diverse facial features and making gender predictions. The system should be trained on a dataset of facial images, considering variations in facial expressions, poses, and lighting conditions. The primary objectives include achieving high classification accuracy, optimizing model performance, and ensuring the system's applicability in real-world scenarios.

# Introduction

Gender detection plays a crucial role in various applications, from user experience personalization to demographic analysis. This project focuses on leveraging deep learning to create an effective gender detection system using a dataset of facial images.

# Model Used

In this gender classification project, three distinct models are employed to predict the gender of individuals from facial images. The first model is a custom-designed Convolutional Neural Network (CNN), comprising multiple convolutional and fully connected layers with batch normalization and dropout for regularization. Additionally, transfer learning is employed with two popular pre-trained models, VGG16 and ResNet50, both originally trained on ImageNet. Custom fully connected layers are added on top of these base models to adapt them for gender classification. The models are trained using an Adam optimizer with varying learning rates, and the training process is enhanced by a learning rate scheduler. The project aims to explore the effectiveness of both custom and transfer learning models in accurately predicting gender from facial images.

# Dataset Used

**The UTKFace dataset** is a comprehensive facial dataset containing over 20,000 labeled images, making it a valuable resource for research in computer vision and machine learning. Each image in the dataset is annotated with information regarding the individual's age, gender, and ethnicity. The age labels cover a broad range from 0 to 116 years, facilitating tasks such as age estimation. Gender information is binary, indicating whether the person is male or female. Additionally, the dataset provides ethnicity labels, contributing to its diversity. UTKFace images exhibit a wide variety of facial expressions, poses, and lighting conditions, ensuring the dataset's suitability for robust facial analysis. Researchers commonly use UTKFace for tasks such as age estimation, gender classification, and facial recognition due to its extensive and diverse content. The dataset is publicly available, organized in a structured format with folders containing images and corresponding metadata files. Proper citation is encouraged when using UTKFace in research to acknowledge the dataset's contributors and maintain ethical practices in data usage.

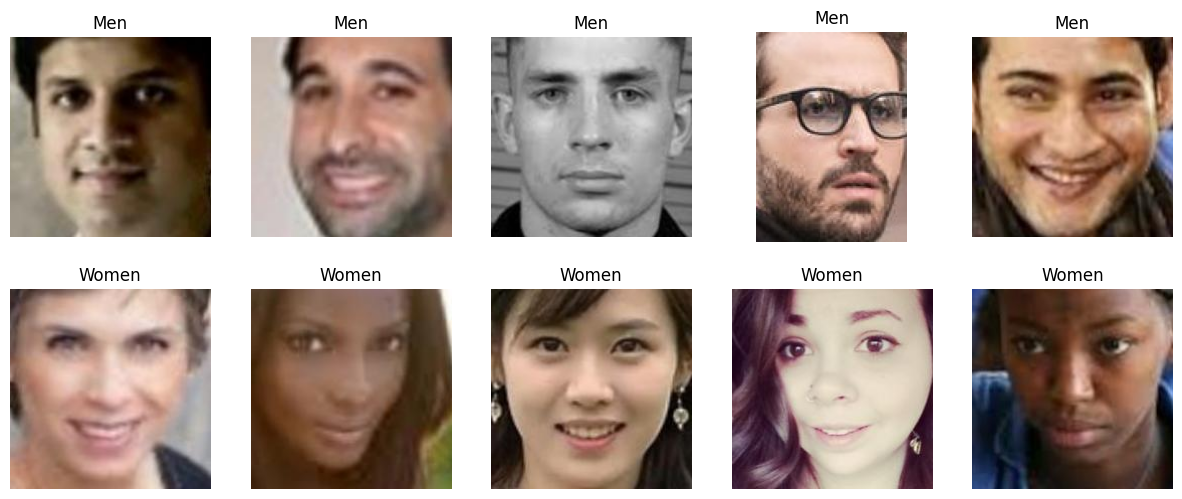
The gender classification dataset utilized in this project is a comprehensive compilation that spans age groups from 16 to 60. The primary source is the UTK dataset, encompassing a wide age range to ensure diversity in the training and testing sets. Additionally, the Kaggle dataset provided in the base code supplements the UTK data, contributing further variety to the model training process. To enrich the dataset, images from Google were acquired and meticulously cropped using Python, expanding the collection with diverse facial representations. This augmentation strategy aims to enhance the model's ability to generalize across a broad spectrum of ages and facial characteristics, ensuring robust gender classification performance. The combination of UTK, Kaggle, and custom Google-acquired images collectively forms a comprehensive and diverse dataset for training and evaluating the gender classification models in this project. Additionally 496 images from CelebA Dataset has been taken for model evaluation.



\*The base code is taken from github repository named Gender-Detection made by BalajiSrinivasan [Github rep](https://github.com/balajisrinivas/Gender-Detection)

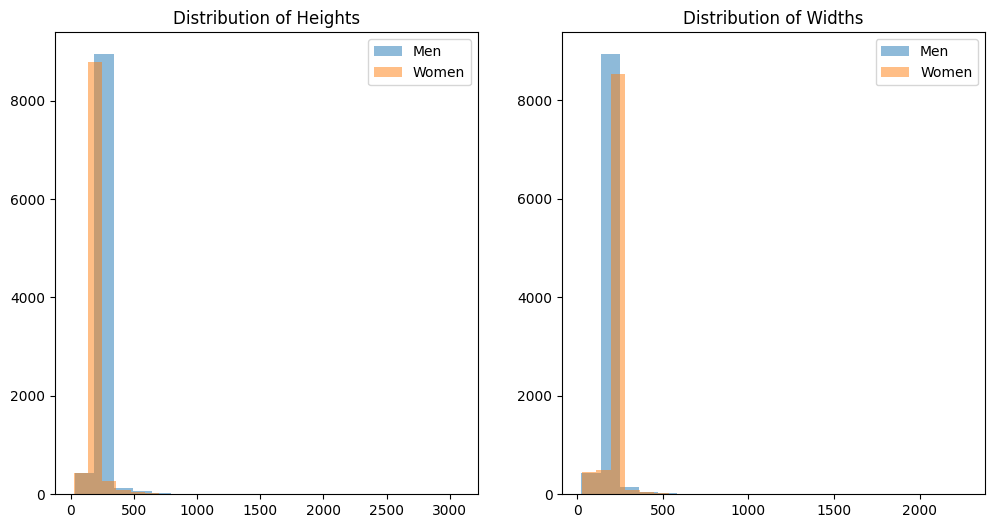
# EDA

## Viewing the dataset

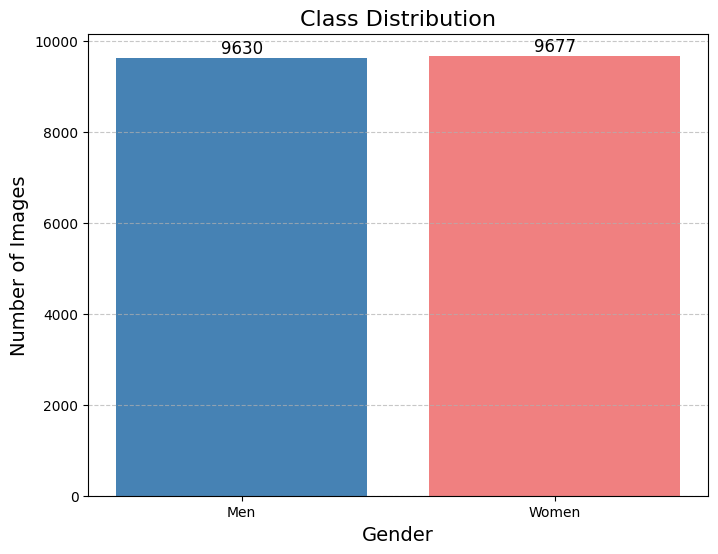


## 

## Distribution of Image Dimensions in the dataset



## Class Distribution



# CNN Model Architecture

## Input Layer

* Input shape: (img\_width, img\_height, 3), representing the width, height, and color channels of the input image.

## Convolutional Layers

* First Conv2D layer with 32 filters, kernel size (3, 3), 'same' padding, and ReLU activation.
* Batch Normalization layer for normalization.
* MaxPooling2D layer with pool size (3, 3) for spatial down-sampling.
* Dropout layer with a dropout rate of 0.25 for regularization.
* Second Conv2D layer with 64 filters, kernel size (3, 3), 'same' padding, and ReLU activation.
* Batch Normalization layer.
* Third Conv2D layer with 64 filters, kernel size (3, 3), 'same' padding, and ReLU activation.
* Batch Normalization layer.
* MaxPooling2D layer with pool size (2, 2).
* Dropout layer with a dropout rate of 0.25.
* Fourth Conv2D layer with 128 filters, kernel size (3, 3), 'same' padding, and ReLU activation.
* Batch Normalization layer.
* Fifth Conv2D layer with 128 filters, kernel size (3, 3), 'same' padding, and ReLU activation.
* Batch Normalization layer.
* MaxPooling2D layer with pool size (2, 2).
* Dropout layer with a dropout rate of 0.25.

## Fully Connected Layers

* Flatten layer to convert the 3D feature map to 1D.
* Dense layer with 1024 units and ReLU activation.
* Batch Normalization layer.
* Dropout layer with a dropout rate of 0.5 for regularization.

## Output Layer

Dense layer with 1 unit and a sigmoid activation function for binary classification (gender prediction).

This architecture combines convolutional and fully connected layers to effectively capture hierarchical features in facial images, facilitating gender classification. The model is compiled using the Adam optimizer with a binary crossentropy loss function.

# VGG16 Model Architecture

The VGG16 model used in this gender classification project is a popular pre-trained deep neural network architecture originally designed for image classification tasks. The customized architecture for gender classification by adding additional layers on top of the VGG16 base is as follows:

## Base VGG16 Model

* Input shape: (img\_width, img\_height, 3), representing the width, height, and color channels of the input image.
* VGG16 includes multiple blocks of convolutional layers with max pooling.

## Custom Fully Connected Layers

* Flatten layer to convert the 3D feature map from the VGG16 base into a 1D vector.
* Dense layer with 512 units and ReLU activation.
* Dropout layer with a dropout rate of 0.5 for regularization.
* Dense layer with 1 unit and a sigmoid activation function for binary classification (gender prediction).

The VGG16 architecture, with its deep convolutional structure, provides a powerful feature extractor. The additional custom fully connected layers adapt the model to the specific gender classification task. The model is compiled using the Adam optimizer with a binary crossentropy loss function for training. Transfer learning with VGG16 leverages pre-trained weights, enabling effective feature extraction and classification on the gender dataset.

# ResNet Model Architecture

The ResNet (Residual Network) model used in this gender classification project is based on the ResNet50 architecture. The customized architecture for gender classification, where additional layers are added on top of the ResNet50 base, is as follows:

## Base ResNet50 Model

* Input shape: (img\_width, img\_height, 3), representing the width, height, and color channels of the input image.
* ResNet50 consists of a deep residual network architecture with skip connections and bottleneck layers.

## Custom Fully Connected Layers

* Flatten layer to convert the 3D feature map from the ResNet50 base into a 1D vector.
* Dense layer with 512 units and ReLU activation.
* Dropout layer with a dropout rate of 0.5 for regularization.
* Dense layer with 1 unit and a sigmoid activation function for binary classification (gender prediction).

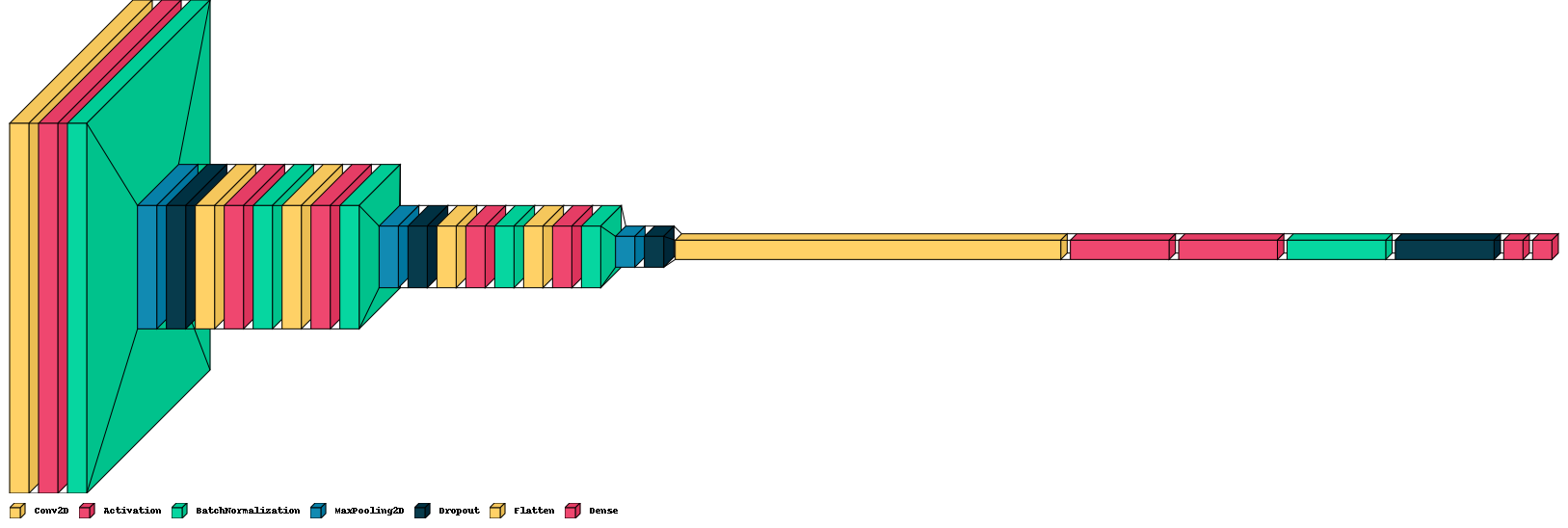
The ResNet architecture introduces skip connections, enabling the effective training of very deep networks. The additional custom fully connected layers adapt the model to the specific gender classification task. The model is compiled using the Adam optimizer with a binary crossentropy loss function for training. Transfer learning with ResNet50 leverages pre-trained weights, enhancing the model's ability to capture complex features in facial images for gender classification.

# Model Architecture Diagram

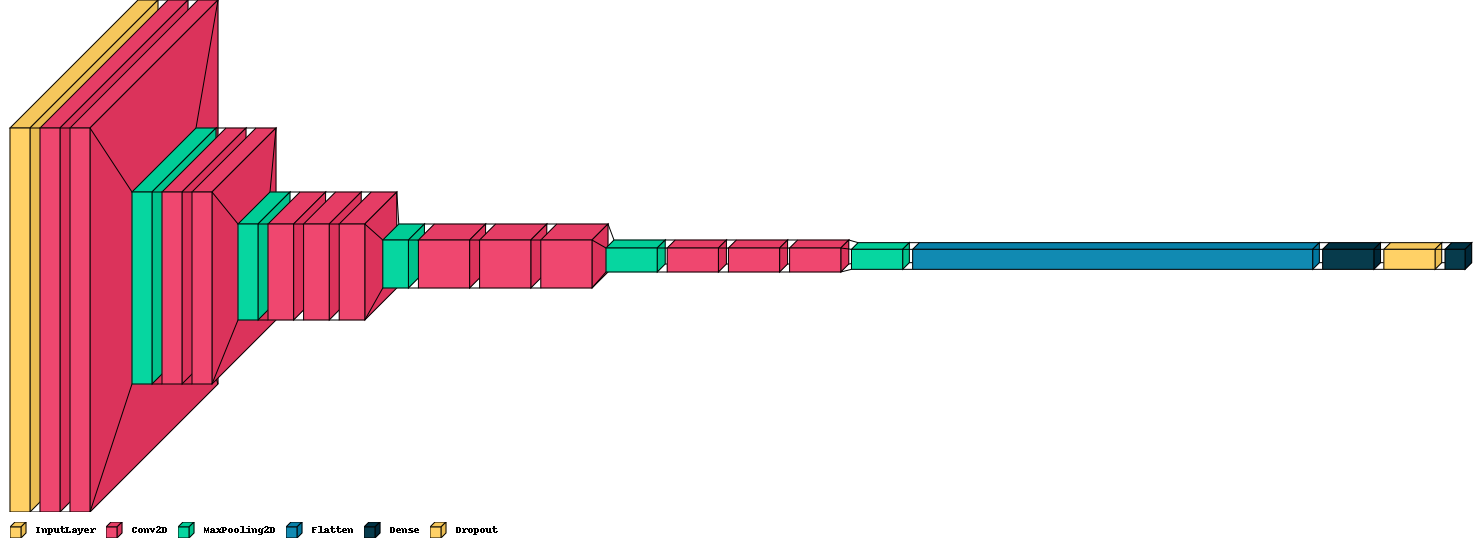
Used Visual Keras to plot the model architecture diagram

## 

## CNN Architecture Diagram

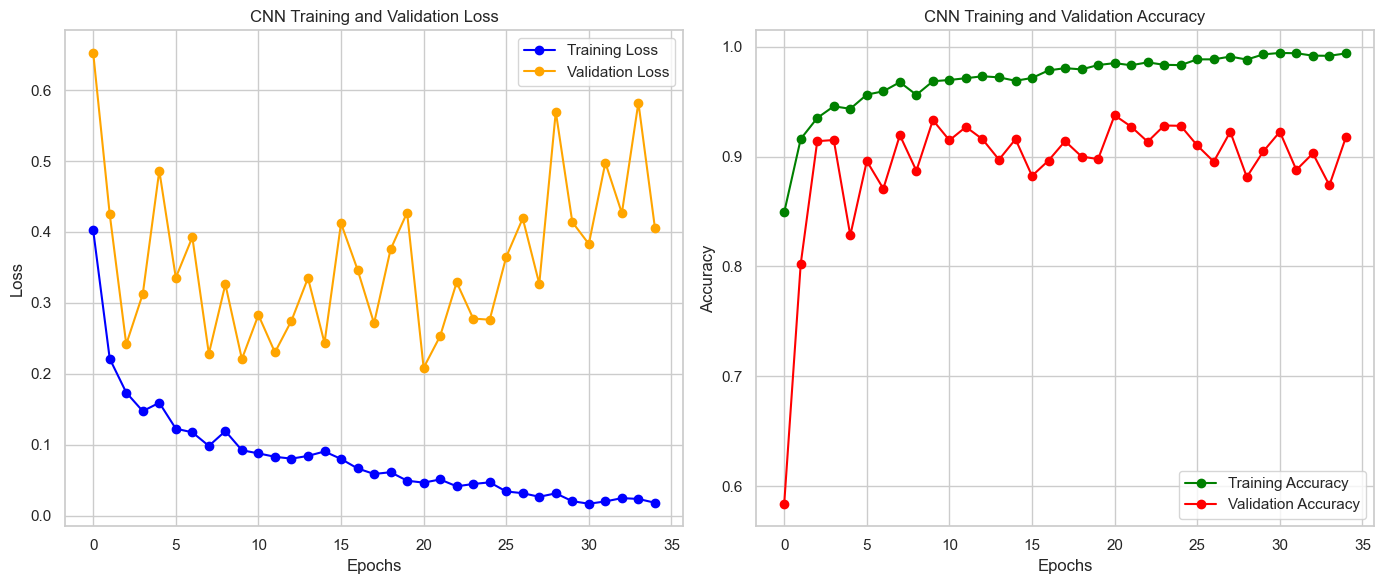


## VGG16 Architecture Diagram



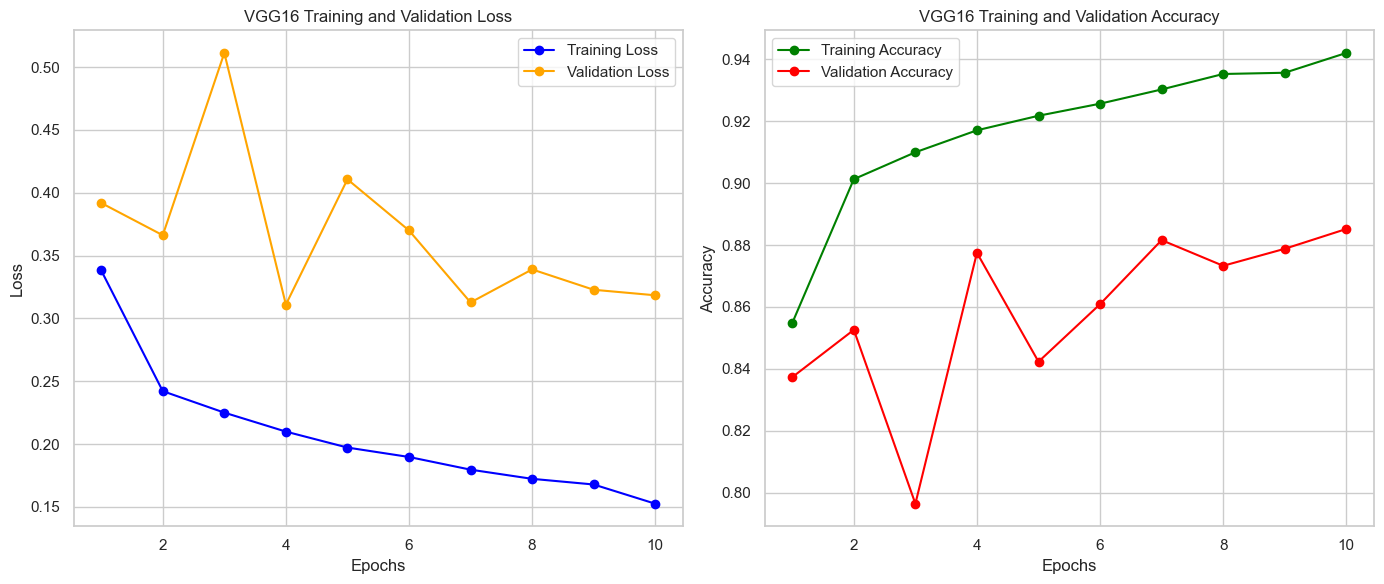
# Plot of training and validation accuracy and loss of models through the epoch

## CNN Model

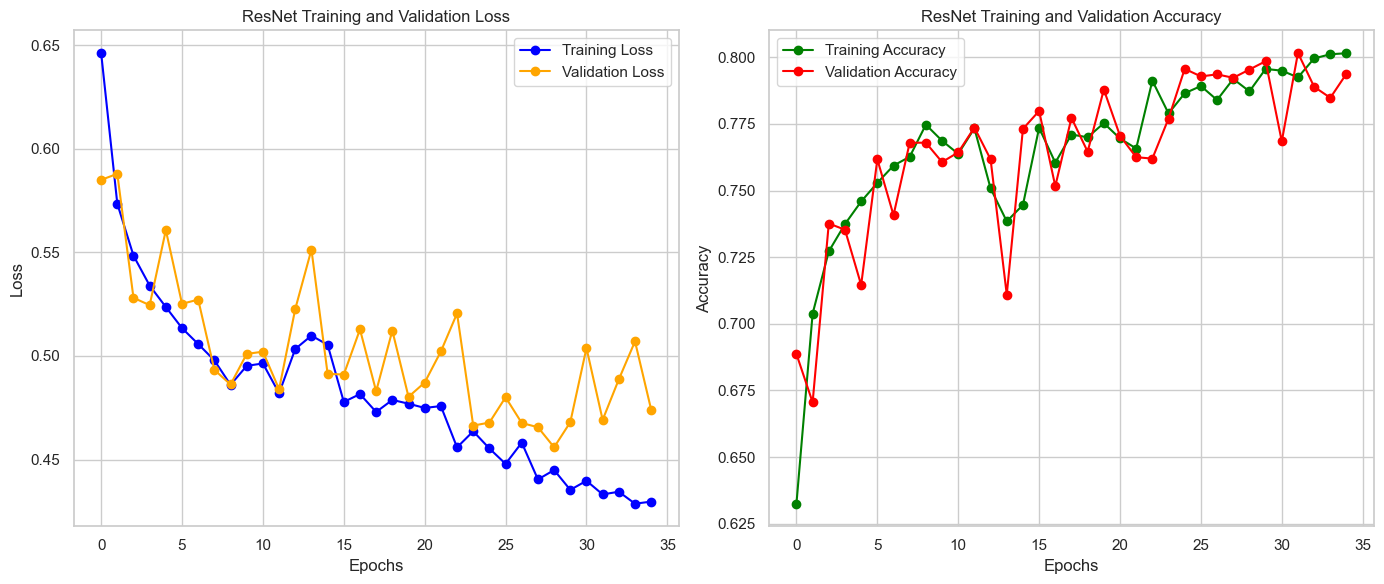


## 

## VGG16 Model



## ResNet Model



# Performance of the Model

For validation completely different images are used. The images are taken from the CelebA dataset. 496 images are taken from that dataset among which 249 images of women and 247 images of men.

## CNN Model

* Accuracy: 82%
* Precision-Recall Metrics: The model shows high precision and recall for both genders, particularly for women.
* Confusion Matrix: Higher false negatives for men, leading to a lower recall for men.

## VGG16 Model

* Accuracy: 89%
* Precision-Recall Metrics: The model exhibits high precision and recall for both genders, with a balanced performance.
* Confusion Matrix: Balanced distribution of true positives and true negatives.

## 

## ResNet Model

* Accuracy: 66%
* Precision-Recall Metrics: The model performs reasonably well in terms of precision and recall for both genders, with a better recall for women.
* Confusion Matrix: Higher false positives for men and false negatives for women.

## Inference

* The VGG16 model outperforms both ResNet and the custom CNN model in terms of overall accuracy (89%).
* VGG16 achieves a well-balanced precision-recall performance for both genders, indicating robustness.
* The custom CNN model, while achieving good accuracy (82%), exhibits a higher rate of false negatives for men, affecting the recall for men.
* When it comes to real time camera gender classification the CNN model provides faster classification and also performs considerably better classification

## Comparison Table

| **Model** | **Precision**  **(weighted average)** | **Recall**  **(weighted**  **average)** | **F-1 Score**  **(weighted**  **average)** | **Accuracy** |
| --- | --- | --- | --- | --- |
| CNN | 0.87 | 0.82 | 0.82 | 0.82 |
| VGG16 | 0.89 | 0.89 | 0.89 | 0.89 |
| ResNet | 0.68 | 0.66 | 0.64 | 0.66 |

# Applications

## 1.Ensuring Women's Safety in Public Transportation

In public transportation, instances may occur where individuals unintentionally enter women's compartments. To address such situations, a gender classification model can be implemented. The model detects the gender of passengers and prompts those in the wrong compartment to leave. This is done by identifying their gender and addressing them based on the color of their attire.

## 2.Human-Computer Interaction

In human-computer interaction, gender detection can be utilized to create personalized user experiences. For example, a system may adapt its interface based on the detected gender of the user.

## 3.Retail Analytics

Retailers can use gender detection to analyze customer demographics, helping them understand the gender distribution of shoppers. This information can be valuable for targeted marketing and product placement strategies.

## 4.Security and Access Control

Gender detection can enhance security systems by incorporating an additional layer of identity verification. It can be integrated into access control systems to regulate entry based on detected gender.

## 5.Event Planning

Event organizers can leverage gender detection for audience analysis during events. This information can guide event planning, including the design of facilities, amenities, and catering services.

# Limitations

## Data Bias and Representation

The models heavily rely on the diversity and representativeness of the training data. If the training data is biased or lacks diversity in terms of age, ethnicity, or other factors, the model's predictions may not generalize well to more diverse populations.

## Age Sensitivity

The age range of 16 to 60 might not cover all potential age groups, and the models may struggle with accurate predictions for individuals outside this range. The models might not generalize well to extreme age groups, such as children or elderly individuals.

## Limited Environmental Variability

The models are trained on facial images obtained from various sources, but they may not perform well in different environmental conditions, lighting situations, or with varying image qualities. Robustness to these factors may be limited.

## Sensitivity to Image Quality

The performance of the models may be sensitive to image quality and resolution. Noisy or low-quality images could affect predictions, and the models may struggle to generalize to images captured in real-world, uncontrolled settings.

## Limited Context Consideration

The models primarily focus on facial features and may not consider contextual information, such as clothing or accessories, which can also contribute to gender perception. This limitation might affect accuracy in scenarios where such context is crucial.

# Future Improvements

* Improved Child Face Gender Classification: Addressing the limitation related to child face gender classification is crucial for refining the model's performance across all age groups. Focused efforts on enhancing accuracy in this specific demographic will contribute to the model's overall effectiveness.
* Gender Count from Video: Implementing a mechanism to count genders from videos, while considering the uniqueness of individuals (ignoring multiple counts for the same person), would significantly improve the model's applicability in real-world scenarios. This enhancement ensures a more accurate representation of gender distribution in dynamic settings.

# Conclusion

In conclusion, the gender classification models, including the custom Convolutional Neural Network (CNN), VGG16, and ResNet50, exhibit varying degrees of effectiveness in predicting gender from facial images. Each model has its strengths and limitations, and the choice of the most suitable model depends on the specific application requirements.