**Face Classification using MobileNetV3-Small with SE Attention**

**Project Overview**

This project aims to build a lightweight and accurate deep learning model for gender classification from facial images using the MobileNetV3-Small architecture enhanced with Squeeze-and-Excitation (SE) attention. The model is trained and evaluated on a dataset organized into gender-based subfolders ("male" and "female").

**1. Model Architecture**

**1.1 MobileNetV3-Small**

MobileNetV3-Small is a lightweight convolutional neural network designed for mobile and edge devices. It uses efficient building blocks like **depth wise separable convolutions**, **hard-swish activation**, and **squeeze-and-excitation (SE) attention** to achieve a balance between **speed and accuracy**, making it ideal for real-time facial classification tasks. The 'Small' variant is ideal for low-resource environments.

**1.2 SE Attention**

SE (Squeeze-and-Excitation) blocks improve feature representation by allowing the network to recalibrate channel-wise feature responses adaptively.

**Squeezing**: Aggregating spatial information using global average pooling.

**Exciting**: Re-weighting channel importance through a gating mechanism.  
 This helps the model prioritize the most informative features (e.g., facial contours, eyes, etc.) for classification.

**1.3 Final Classifier**

After SE-enhanced features are extracted, a Global Average Pooling layer followed by a fully connected classifier outputs the gender class (male/female).

* **Global Average Pooling (GAP)** to flatten the spatial feature maps.
* A **fully connected (linear) layer** that maps the feature vector to the target classes (e.g., male/female).
* This produces class probabilities used for the final prediction.

**2. Dataset**

The dataset is structured as follows:

Task\_A/

├── train/

│ ├── female/

│ └── male/

├── val/

├── female/

└── male/

* **Total Classes**: 2
* **Image Size**: Resized to 224x224
* **Normalization**: Standard ImageNet mean and std

**3. Training Details**

**3.1 Optimizer and Loss**

* **Optimizer**: Adam
* **Learning Rate**: 0.0001
* **Loss Function**: CrossEntropyLoss

**3.2 Metrics Tracked**

* Training and Validation Loss
* Accuracy
* Precision, Recall, F1-score (macro and weighted)

**3.3 Training Summary**

* **Epochs**: 50
* **Batch Size**: 16

**3.4 Resource Usage**

* **CPU Memory Used**
* **Training Time**: ~95.0 mins (varies by hardware)

**4. Evaluation Results**

**4.1 Validation Classification Report**

| **Class** | **Precision** | **Recall** | **F1-Score** | **Support** |
| --- | --- | --- | --- | --- |
| Female | 0.89 | 0.68 | 0.77 | 79 |
| Male | 0.93 | 0.98 | 0.95 | 343 |
|  |  |  |  |  |
| **Accuracy** |  |  | **0.92** | **422** |
| **Macro Avg** | 0.91 | 0.83 | 0.86 | 422 |
| **Weighted Avg** | 0.92 | 0.92 | 0.92 | 422 |

**4.2 Confusion Matrix**

Visualized using seaborn to show true positives, false positives, etc.

A blue squares with black numbers

AI-generated content may be incorrect.

**5.Training vs Validation plot (loss and accuracy)**

Plot curves against Training and Validation output to visualize loss and accuracy.

**A graph of a graph of a graph

AI-generated content may be incorrect.**

**6. Grad-CAM Visualization**

Grad-CAM heatmaps were used to visualize where the model focuses when predicting gender from a face. Results showed attention around eyes, nose, and jawline.



**7. Feature Selection**

SE attention dynamically emphasizes informative channels. This adaptive feature selection contributed to performance without manual feature engineering.

**8. Test Script**

A separate test.py script is provided that:

* Accepts a test dataset path (same folder structure)
* Loads the pretrained model
* Outputs evaluation metrics: Accuracy, Precision, Recall, F1-score

**9. GitHub Repository Structure**

**Face-Gender-Classification**

**├── model.py # MobileNetV3 + SE definition**

**├── train.py # Training script**

**├── test.py # Evaluation script**

**├── utils.py # Utility functions**

**├── README.md # Project explanation**

**├── pretrained.pth # Trained model weights**

**└── requirements.txt # Libraries needed**

**10. Submission Summary**

* Training + validation results
* Classification report
* Confusion matrix
* Grad-CAM analysis
* Test script
* GitHub repo with pretrained weights
* Folder structure followed

**11. Conclusion**

This project successfully demonstrates the design and implementation of a lightweight yet powerful deep learning model for face-based gender classification using **MobileNetV3-Small enhanced with Squeeze-and-Excitation (SE) attention**.

By combining MobileNetV3’s efficiency with SE’s adaptive feature selection capabilities, the model achieves **high accuracy (92%)** while maintaining **computational efficiency**, making it suitable for deployment on mobile and edge devices.

Key accomplishments include:

* Accurate classification with clear performance metrics (Precision, Recall, F1-Score)
* Visual interpretability through **Grad-CAM heatmaps**
* Efficient training pipeline with resource usage tracking
* Clean dataset integration and evaluation-ready architecture

The project also adheres to all submission requirements, including test script support, pretrained weights, and structured reporting. This solution offers a solid foundation for real-world face classification tasks and can be extended to age detection, emotion recognition, or multi-label facial analysis with minimal modifications.

**In essence**, this project balances accuracy, speed, and explainability—delivering a practical and deployable AI solution for facial classification.

For submission, please include:

* This report (PDF or link)
* GitHub repo with code + weights
* Google Form link with repo + results