Bias Assessment and Explainable(XAI) Driven Region-Based Prediction of Human Age and Gender Using Facial Images.

Presented By-

MD Rufsan Jani Shanto(221-35-1064)

Supervised by:

Ms. Ishrat Sultana

Lecturer, Department of Software Engineering

Outlines

- Research Gaps and Problem Statement
- Objectives and Contribution
- Research Scope and Literature Review
- Methodology
- Result and Discussion
- Conclusion
- References
- Publications

Research Gaps & Problem Statement

Research Gaps

- ✓ Existing facial age–gender studies mostly focus on accuracy/MAE and rarely analyze fairness across gender, age-group or race, especially on UTKFace.

- ✓ Very few works combine XAI with bias analysis: they show a few Grad-CAM heatmaps but do not quantify which regions (forehead, mid-face, jaw) drive predictions.

Problem Statement

- ✓ Automatic facial age–gender prediction is highly sensitive to pose, lighting, occlusion, ethnicity, and image quality, which makes reliable prediction difficult.

- ✓ We need fair and explainable models that can check demographic bias and clearly show which parts of the face influence the predictions.

Objectives

➔ **RO1:** To evaluate and compare 4 deep learning architectures on the UTKFace dataset.

➔ **RO2:** To quantify demographic biases across gender groups using TPR, demographic parity.

➔ **RO3**: To identify which facial regions (forehead, mid-face, jaw) contribute most to predictions using Grad-CAM explainability.

Research Scope

- ✓ The study focuses on both early-aged and aged people.
- ✓ The study uses Grad-CAM to show which parts of the face the model focuses on when making decisions.
- ✓ The study is limited to binary human gender (male and female) and does not cover non-binary gender identities.

Literature Review

Title	Authors	Year	Applied Method	Key finding
Age Estimation from Facial Images using Transfer Learning and K-fold Cross-Validation.	S. M. Shihab Uddin et al.	2021	Used pre-trained CNNs (VGG16, ResNet50, and SENet50 with VGGFace weights), fine-tuned with custom layers for 8 age groups. Applied transfer learning, 5-fold cross-validation, data augmentation, and layer freezing, trained on Google Colab GPUs (~5 hrs/model). **Dataset:** UTKFace dataset—comprising over 20,000 images with ages from 0 to 116 years, labeled for age, gender, and race	**Result:** ResNet50 performed best with 88.03% validation accuracy, surpassing FaceNet (56.9%), VGG16 (64%), and distillation-based models. **Limitation:** Limited data, design, and real-world variability reduce model generalization.
Age and Gender Detection using Facial Images	Esmat Mohamed et al.	2023	The system uses Haar classifiers or YOLOv8 for face detection and CNNs for classification, with preprocessing like alignment and augmentation to enhance robustness, optimized for real-time inference. **Dataset:** IMDB-Wiki, FairFace, UTKFace, Age BD	**Result:** YOLOv8 achieved 94.2% gender and 62.5% age accuracy, with 4,700+ correct predictions, showing strong performance. **Limitation:**Age imbalance and limited evaluation metrics may bias and weaken the model's reliability.

Title	Authors	Year	Applied Method	Key finding
Age and Gender Prediction using Deep CNNs and Transfer Learning	Vikas Sheoran et al.	2023	They trained deep models from scratch and with pre-trained networks (VGGFace, ResNet50, SENet50), using transfer learning and linear regression for predictions. **Dataset:** They used the UTKFace dataset with 20,000+ labeled face images under varied conditions.	**Result:** Custom CNN achieved 5.67 MAE (age) and 94.5% (gender). Transfer learning with SENet50 gave best age MAE (4.58) and 94.94% gender accuracy **Limitation:** Faces with occlusion, extreme angles, and low lighting pose challenges.
Age and Gender Prediction Using Machine Learning	Bhavana B. Helwate et al.	2024	Uses OpenCV, CNNs, and TensorFlow with attention and multi-task learning for real-time age and gender prediction. **Dataset:** Not explicitly specified, but the approach aims for performance even with limited data.	**Result:** The proposed method outperforms existing techniques with higher accuracy in age and gender prediction. **Limitation:** Limited details suggest issues with diverse faces, accessories, or distorted images.

Title	Authors	Year	Applied Method	Key finding
1 Human Age and Gender Prediction from Facial Images Using Deep LearningMethods	Puja Dey et al.	2024	The method uses a CNN with preprocessing, data augmentation, and regularization, trained on an 80/20 split and benchmarked against pre-trained models. **Dataset: 1.**Adience: Unfiltered facial images with age groups and gender labels. **2.**UTKFace: Diverse facial images labeled by age and gender, widely used for prediction tasks.	**Result:** The CNN outperformed existing methods with age accuracies of 86.42% and 81.96% and gender accuracies of 97.65% and 96.32%. **Limitation:** Image variability, demographic gaps, and high compute needs limit model performance.
2 Face-based Age and Gender Classification Using Deep Learning Model	Rajiv Kumar et al.	2024	A deep CNN pre-trained on IMDb-WIKI and fine-tuned on OIU-Adience uses dropout, augmentation, and hyperparameter tuning to classify age and gender with accuracy and MAE. **Dataset: 1.**IMDb-WIKI: Used to pre-train CNN on facial features. **2.**OIU-Adience: Benchmark with real-world face images labeled by age and gender.	**Result:**Achieved 84.8% age group accuracy and 2.26 MAE, outperforming CNN2ELM and generalizing well on unconstrained images. **Limitation:** Extreme poses, dataset noise, and limited gender accuracy analysis may affect performance.

Title	Authors	Year	Applied Method	Key finding
Human Age and Gender Prediction from Facial Images Using Deep LearningMethods	Puja Dey et al.	2024	The method uses a CNN with preprocessing, data augmentation, and regularization, trained on an 80/20 split and benchmarked against pre-trained models. **Dataset: 1.**Adience: Unfiltered facial images with age groups and gender labels. **2.**UTKFace: Diverse facial images labeled by age and gender, widely used for prediction tasks.	**Result:** The CNN outperformed existing methods with age accuracies of 86.42% and 81.96% and gender accuracies of 97.65% and 96.32%. **Limitation:** Image variability, demographic gaps, and high compute needs limit model performance.
Face-based Age and Gender Classification Using Deep Learning Model	Rajiv Kumar et al.	2024	A deep CNN pre-trained on IMDb-WIKI and fine-tuned on OIU-Adience uses dropout, augmentation, and hyperparameter tuning to classify age and gender with accuracy and MAE. **Dataset: 1.**IMDb-WIKI: Used to pre-train CNN on facial features. **2.**OIU-Adience: Benchmark with real-world face images labeled by age and gender.	**Result:**Achieved 84.8% age group accuracy and 2.26 MAE, outperforming CNN2ELM and generalizing well on unconstrained images. **Limitation:** Extreme poses, dataset noise, and limited gender accuracy analysis may affect performance.

Title	Authors	Year	Applied Method	Key finding
Research on Deep Learning-based Image Processing and Classification Techniques for Complex Networks	Jiangli Liu et al.	2025	The proposed method designs an encoder using DCNN, ECANet, and DSA_ASPP modules to effectively capture multi-scale features. It integrates SIFT features as network nodes, with correlation coefficients as edge weights, to construct an image feature network for detailed local information extraction. **Dataset:** CUB-200-2011 and Stanford Dogs for classification, and CamVid and Cityscapes for segmentation evaluation.	**Result:** The proposed algorithm outperforms existing models in detection accuracy, achieving 69.6% intersection and 73.6% concurrency ratios. **Limitations:** data demand, high complexity, poor generalization, and need for encoder optimization.
Gender Classification from Human Face Images Using Deep Learning Based on MobileNetV2 Architecture	Nisreen Ryadh Hamza	2025	The study used a fine-tuned pre-trained MobileNetV2 for male/female classification, with data augmentation and transfer learning. Training employed the Adam optimizer, and performance was evaluated using the F1-score. **Dataset:**Biggest Gender/Face Recognition Dataset (Kaggle, 2021): 27,167 images (17,678 male, 9,489 female).	**Result:** The proposed system achieved a 96% F1 score, showing high accuracy and strong precision-recall balance for practical use. **Limitation:** Limited facial variation and real-world conditions may reduce performance.

Title	Authors	Year	Applied Method	Key finding
FaceMe: An Explainable AI-Powered Web-Based Face Recognition System for Age, Gender, and Skin Tone Estimation	Rahma Alliouche et al.	2025	MobileNetV2 (pre-trained, fine-tuned) vs custom CNNSpecialized branches for age, gender, skin tone outputSHAP explainability for model transparencyPre-processing: resizing, normalizationFlask/Python used for the live web appMobileNetV2 trained 5 epochs; custom CNN trained 20 epochs. **Dataset:** Collections include UTKFace and similar public datasets.	**Result:** MobileNetV2: 93% test accuracyCustom CNN: steady accuracy over 20 epochsF1-score: 0.78; Recall: 0.80; Precision: 0.90Successful real-time inference with webcamSHAP clearly identifies facial regions for each prediction. **Limitations:** Dataset bias remains a challengeMulti-output learning can risk overfitImage quality and variation affect results. Ethical risks with demographic predictions.

Methodology

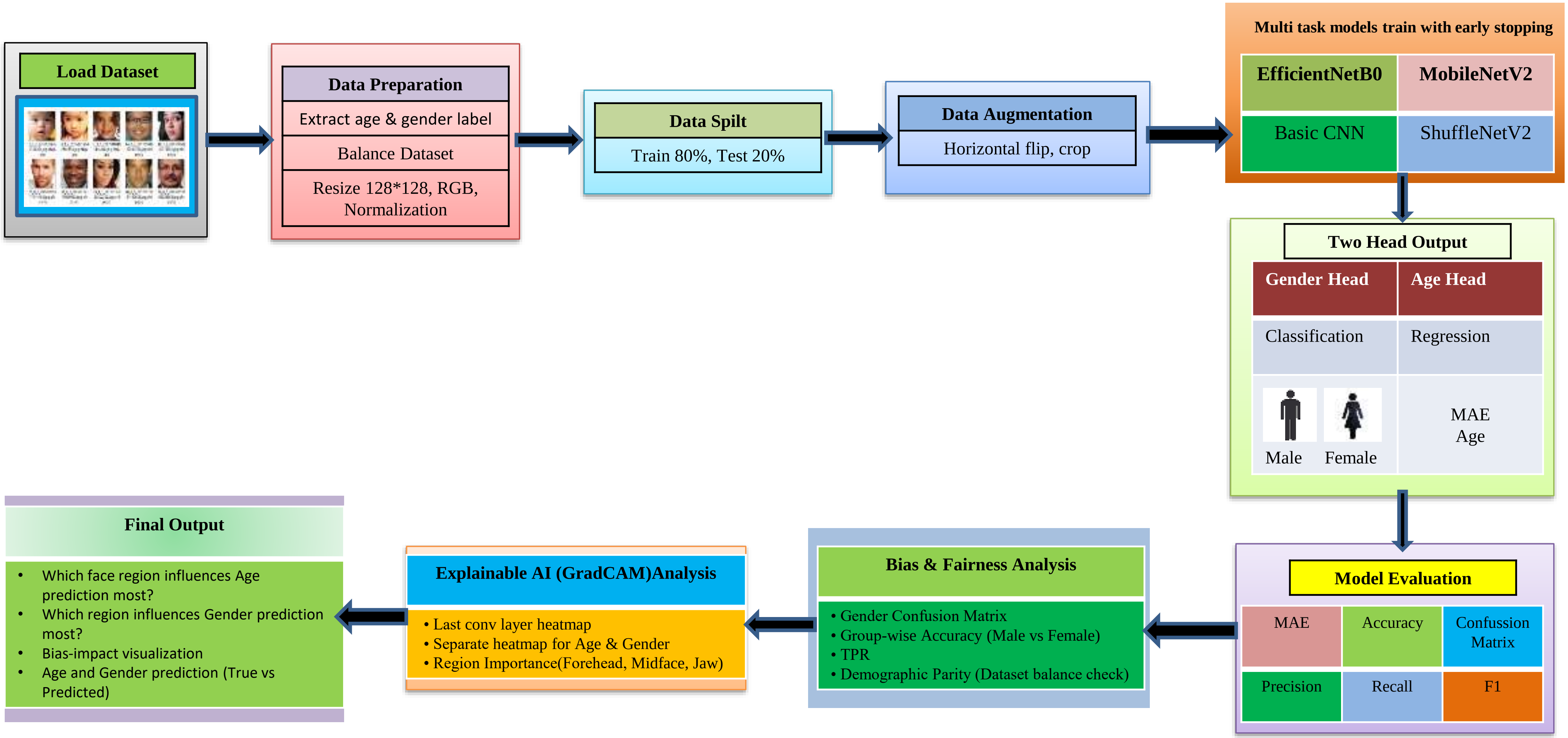


Figure: Work-flow diagram

Result and Discussion

Table View of Dataset

Image File Name	Age	Gender	Ethnicity	Image (Preview ideal)
9_0_4_20170110225238472.jpg.chip.jpg	9	0 (male)	4 (mixed)	
20_1_0_20170103175416510.jpg.chip.jpg	20	1 (female)	0 (white)	
30_1_0_20170117091918162.jpg.chip.jpg	30	1 (female)	0 (white)	
50_0_3_20170119204542975.jpg.chip.jpg	50	0 (male)	3 (indian)	
72_0_3_20170105180912247.jpg.chip.jpg	72	0 (male)	3 (indian)	

Performances of Proposed Models

Comparison of models

Model	Age MAE	Gender Acc.	Precision	Recall	F1	XAI Optimization (zone importance)
EfficientNetB0(main)	**7.07**	**90.71**	**90.40**	**91.20**	**90.80**	**High(forehead)**
MobileNetV2	10.27	88.14	91.80	88.20	89.00	Low (forehead & midface)
Custom CNN	11.71	85.95	86.90	85.90	14.1	Low (forehead)
ShuffleNetV2	15.56	74.15	79.00	74.65	74.60	Mid(forehead)

Performances of Proposed Models

Bias and Fairness of models						
Model	Demographic Parity (Dataset balance Checks)		Demographic Bias In Classification - TPR		Group Fairness (Age MAE)	
	Male	Female	Male	Female	Male	Female
EfficientNetB0(main)	49.80%	52.2%	90.20%	91.20%	7.07	8.37
MobileNetV2	49.80%	50.2%	91.80%	84.50%	10.27	11.92
Custom CNN	46.40%	53.6%	84.90%	86.9%	11.78	12.72
ShuffleNetV2	41.10%	53.90%	78.30%	70.50%	15.56	18.42

Gender and Age Prediction result Proposed Models

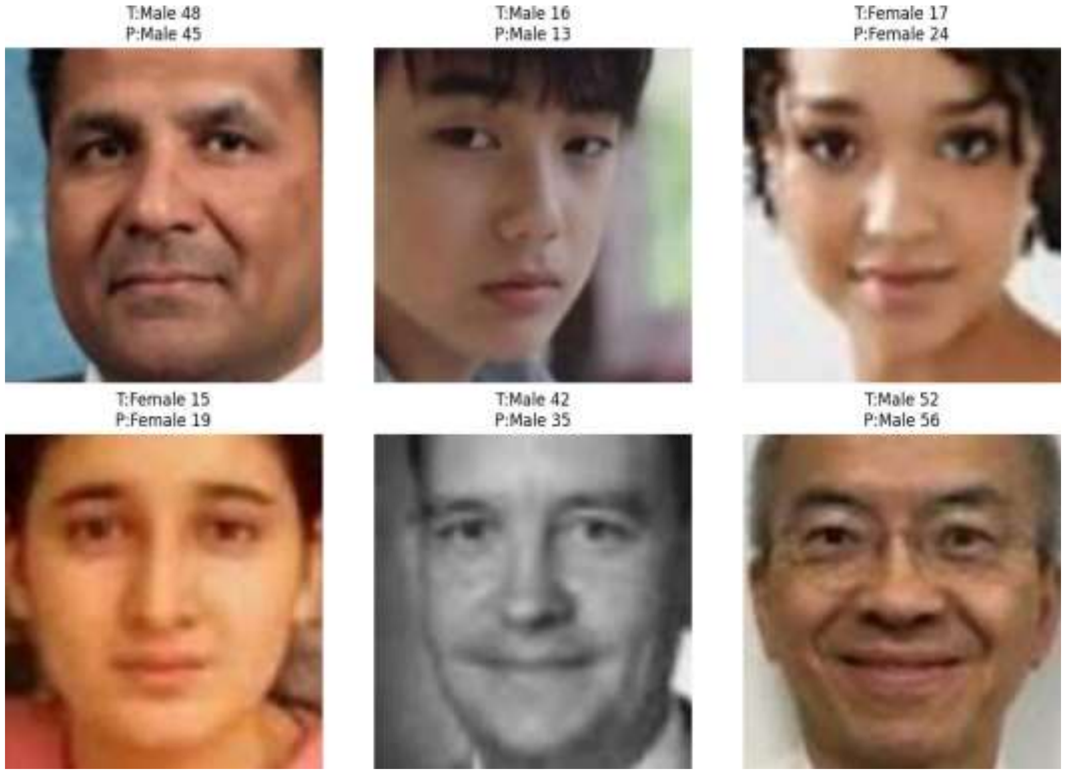


Figure : Predction output of EfficientNetB0 Model .

Figure : Predction output of MobileNetV2 Model .

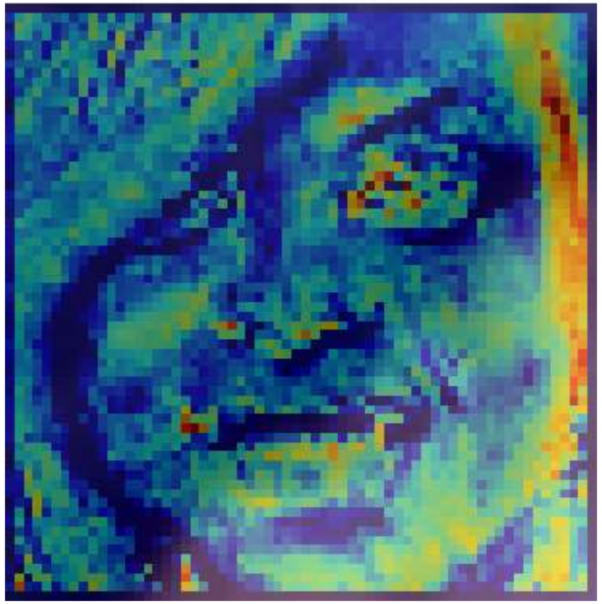
Figure : Predction output of Basic CNN Model .

Figure : Predction output of ShuffleNetV2 Model .

Result of Proposed Model‘s Gradcam/XAI optimized zones for Gender prediction)

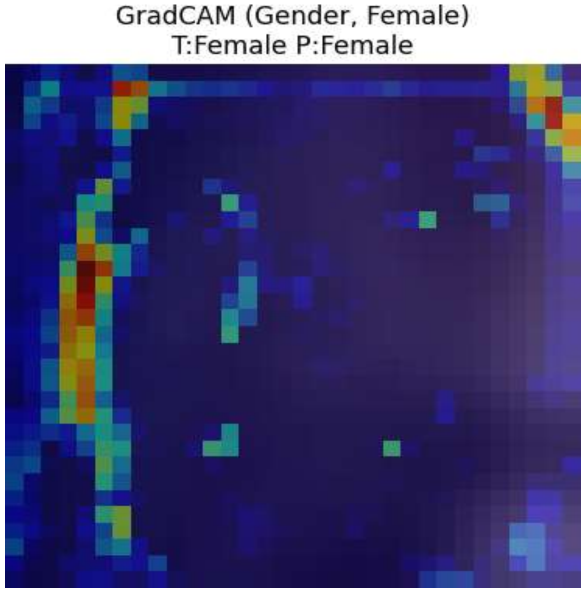
Zone importances: {'Forehead':0.60815, 'Midface': nan, 'Jaw': nan}

Figure : EfficientNetB0 Model’s Gradcam/XAI optimazation .



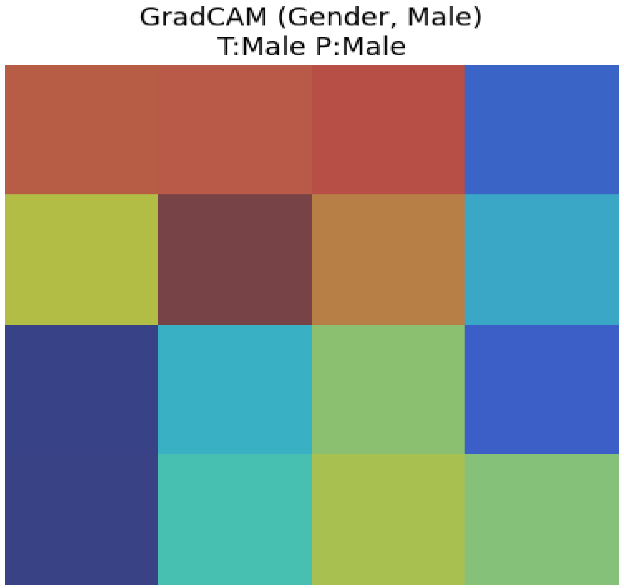
Zone importances:{'Forehead': 0.17574986815452576, 'Midface': 0.1725803166627884, 'Jaw': nan}

Figure : MobileNetV2 Model’s Gradcam/XAI optimazation .



Gender zone: {'Forehead': 0.07564841210842133, 'Midface': nan, 'Jaw': nan}

Figure : Basic CNN Model's Gradcam/XAI optimazation .



Gender zone: {'Forehead': 0.5097659826278687, 'Midface': nan, 'Jaw': nan}

Figure : ShuffleNetV2 Model's Gradcam/XAI optimazation .

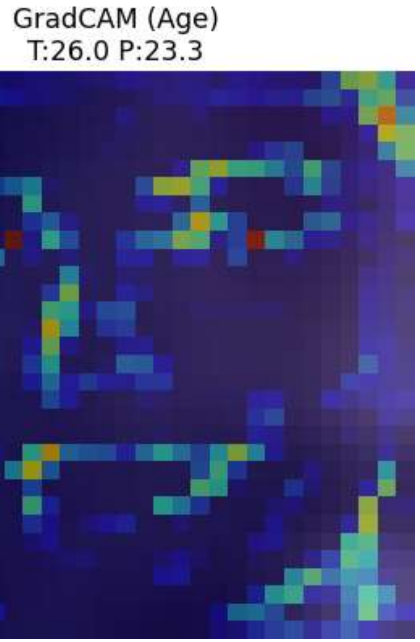
Result of Proposed Model's Gradcam/XAI optimized zones for Age prediction

Zone importances: {'Forehead': 0.53339884, 'Midface': nan, 'Jaw':nan}

Figure : EfficientNetB0 Model's Gradcam/XAI optimazation .

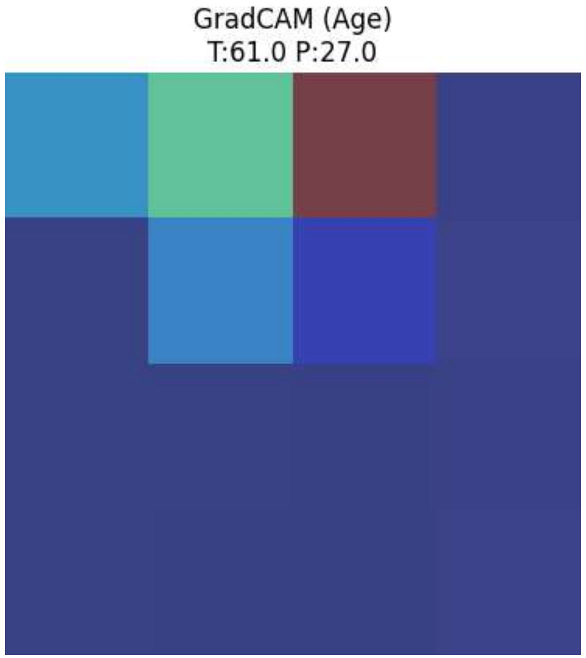
Age zone: {'Forehead': 0.007737392093986273, 'Midface': 0.03990013152360916, 'Jaw': nan}

Figure : MobileNetV2 Model's Gradcam/XAI optimazation .



Age zone: {'Forehead': 0.10249076038599014, 'Midface': nan, 'Jaw': nan}

Figure : Basic CNN Model’s Gradcam/XAI optimazation .



Age zone: {'Forehead': 0.1294628381729126, 'Midface': nan, 'Jaw': nan}

Figure : ShuffleNetV2 Model’s Gradcam/XAI optimazation .

Result in Graphs or Charts

Gender Accuracy Time Series of the models

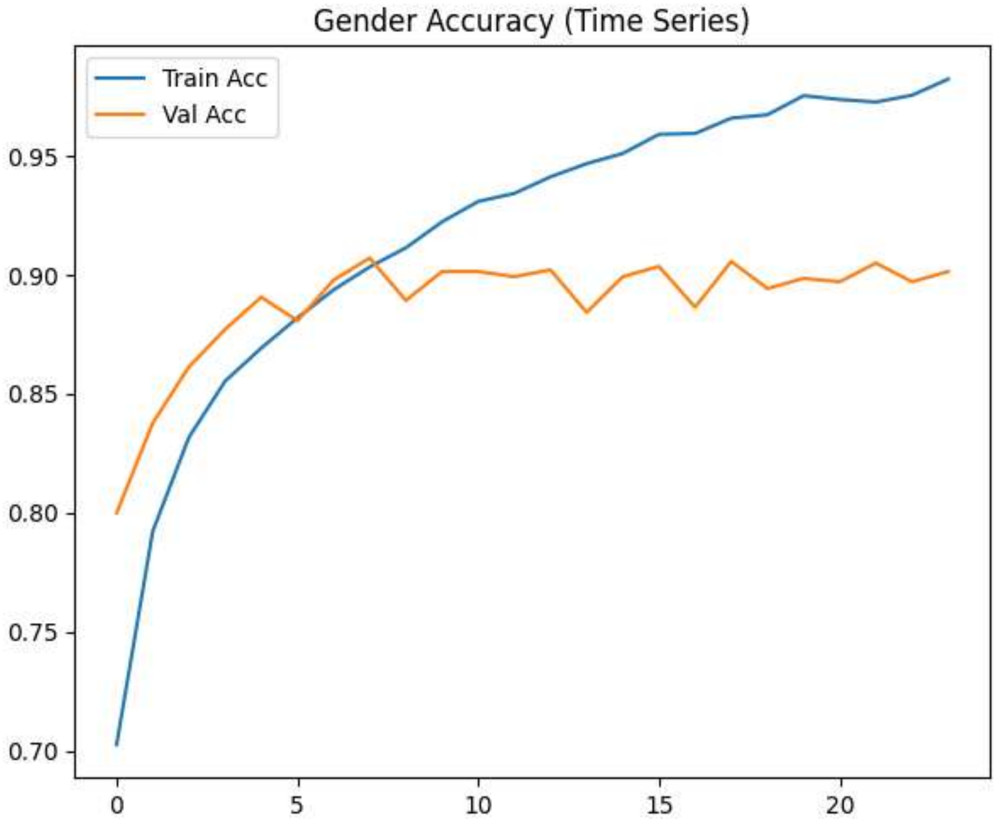


Figure: Gender Accuracy Time Series of the EfficientNetB0 model

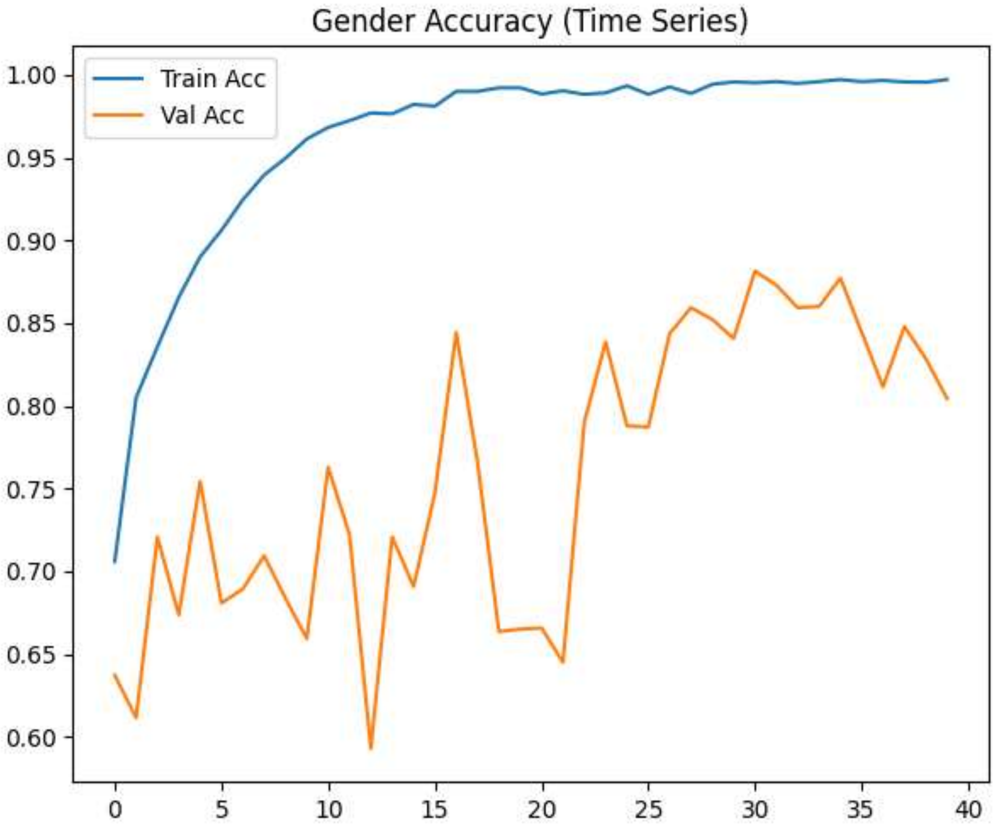


Figure: Gender Accuracy Time Series of the MobileNetV2 model

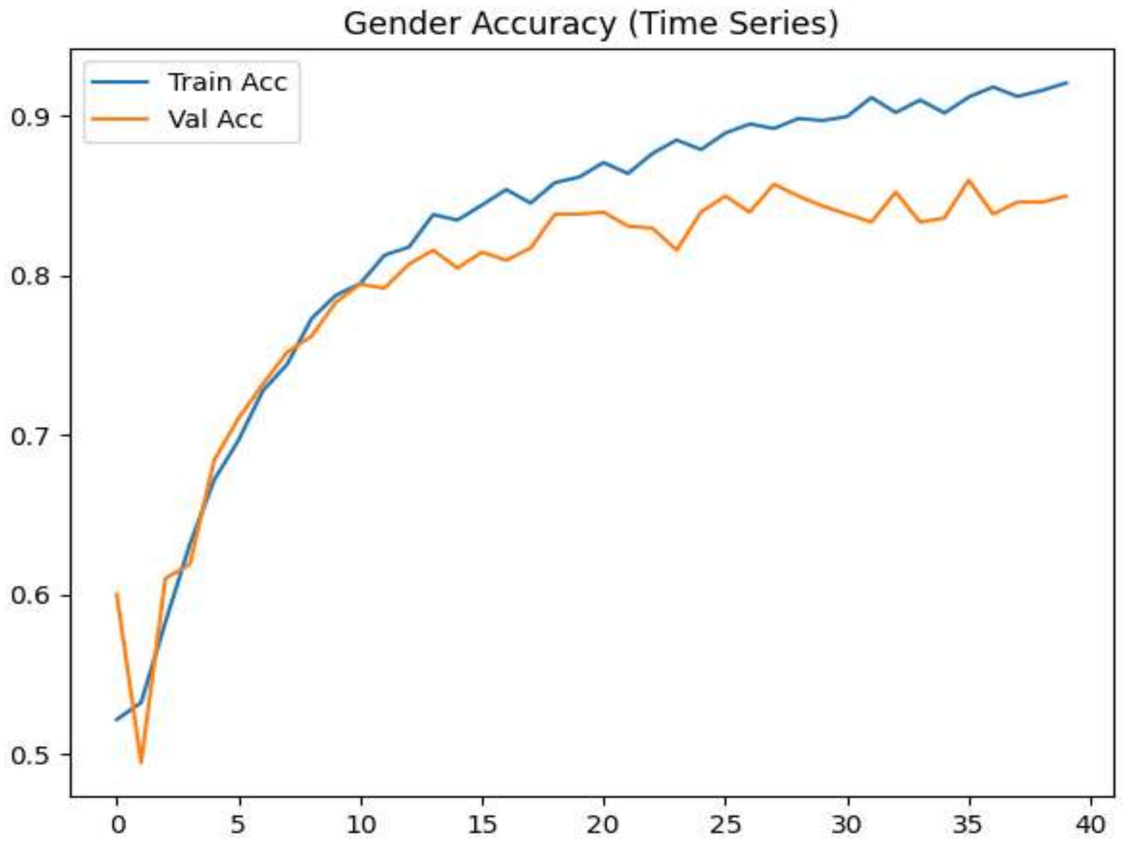


Figure: Gender Accuracy Time Series of the Basic CNN model

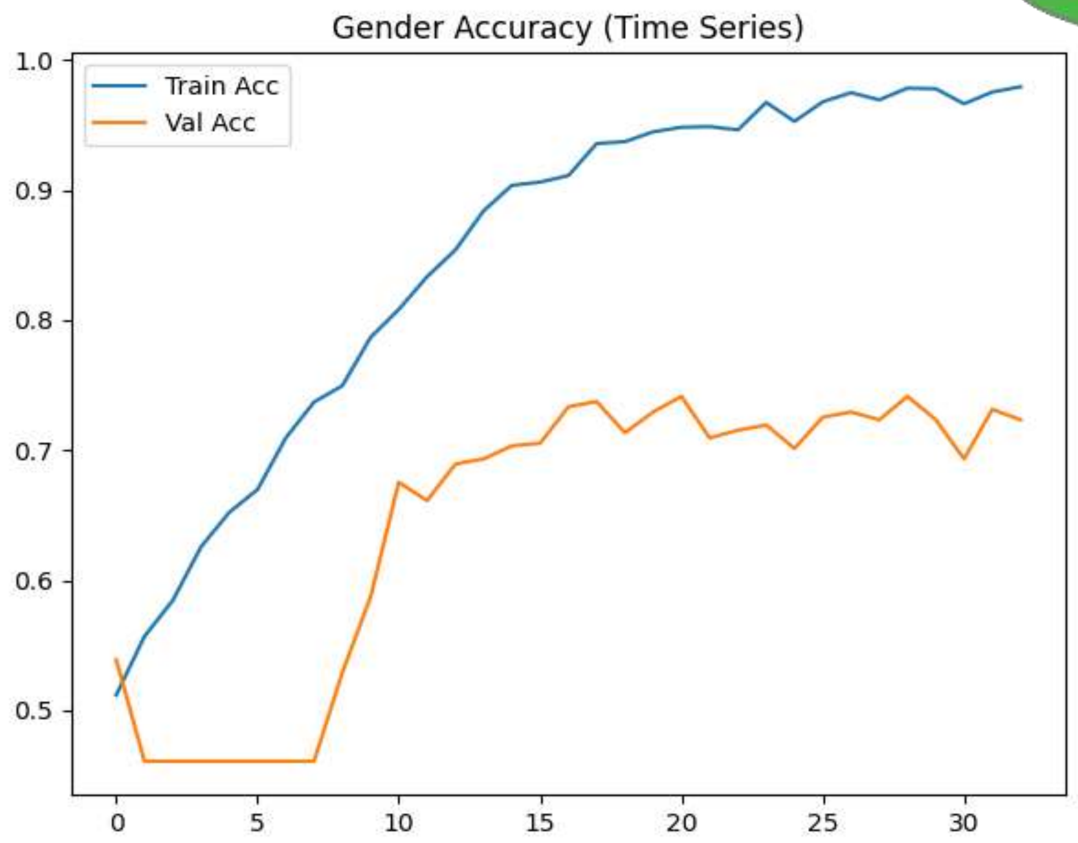


Figure: Gender Accuracy Time Series of the ShuffleNetV2 model

Age MAE Time Series of the models

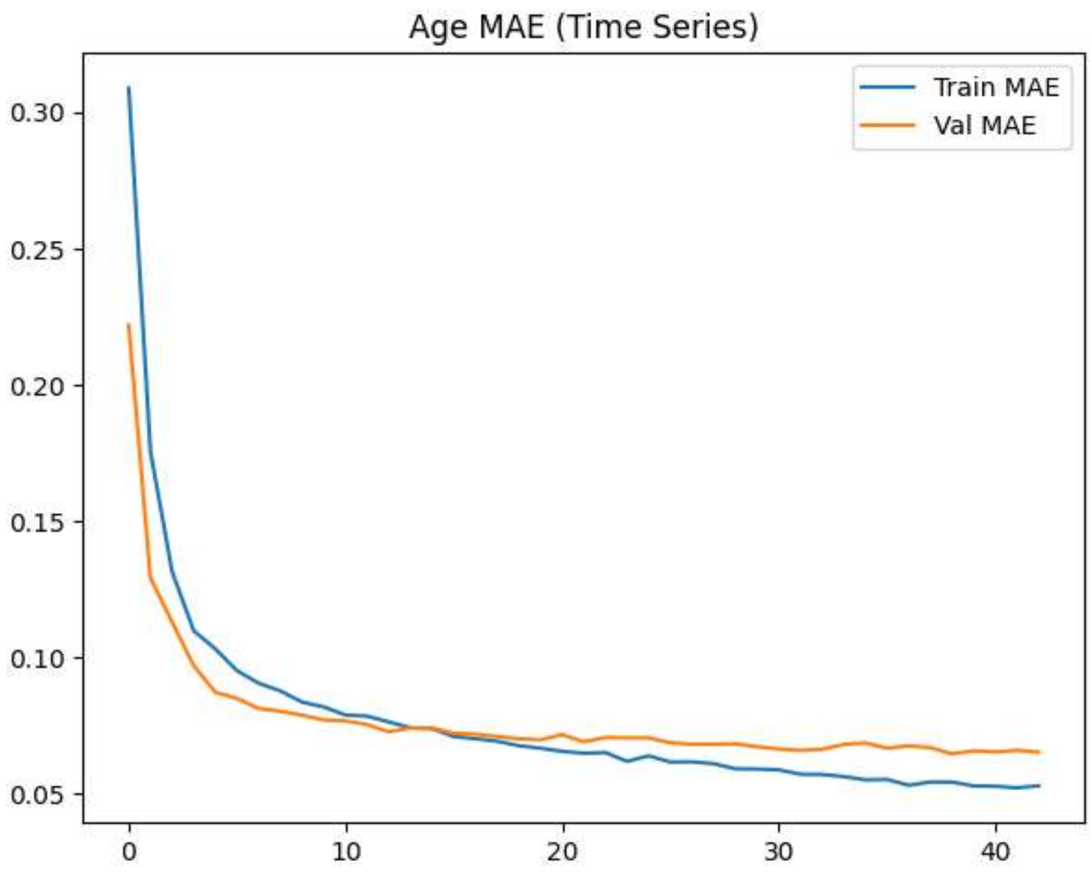


Figure: Age MAE Time Series of the EfficientNetB0 model

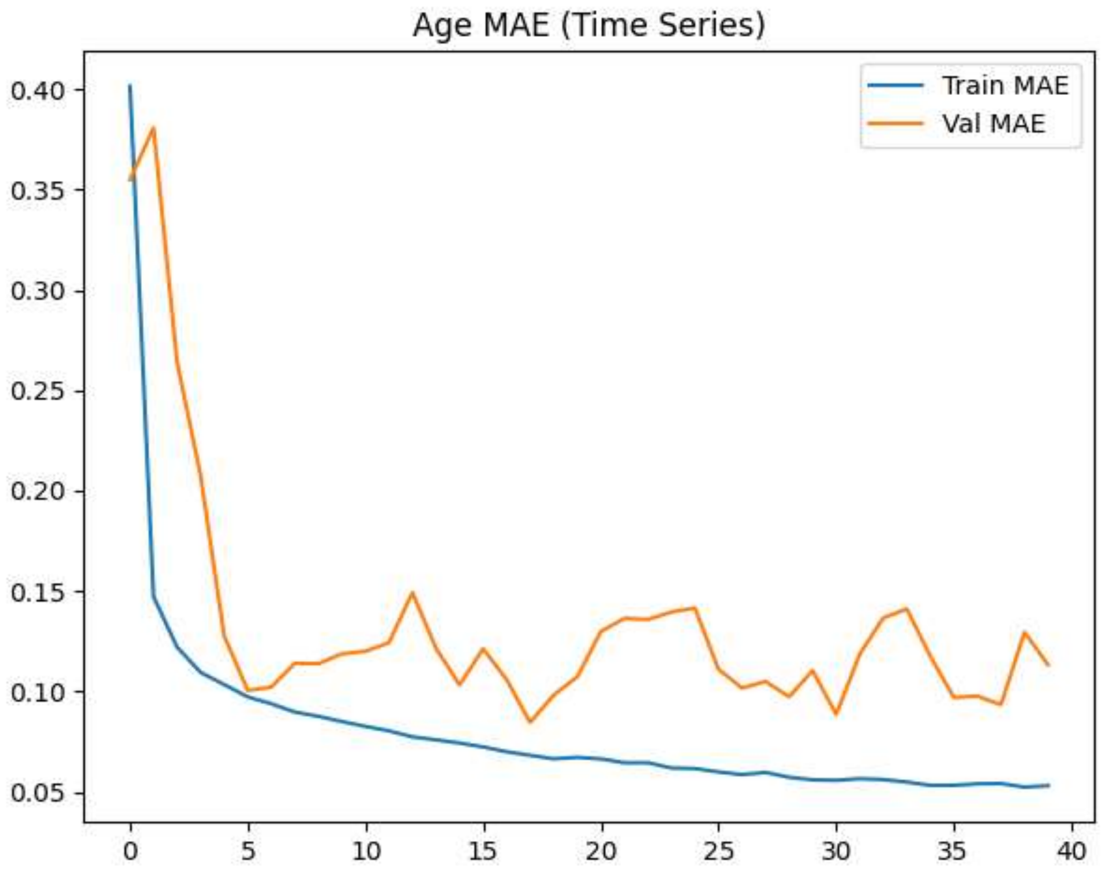


Figure: Age MAE Time Series of the MobileNetV2 model

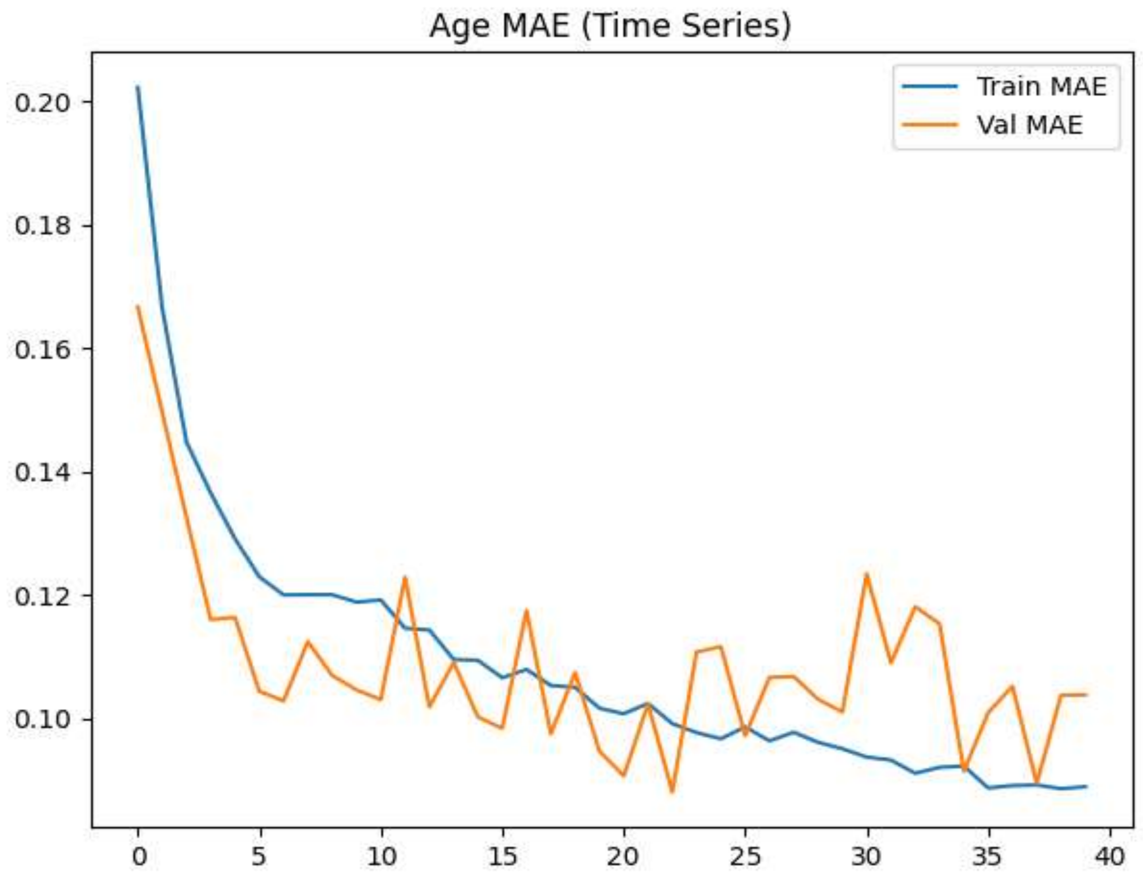


Figure: Gender Accuracy Time Series of the Basic CNN model

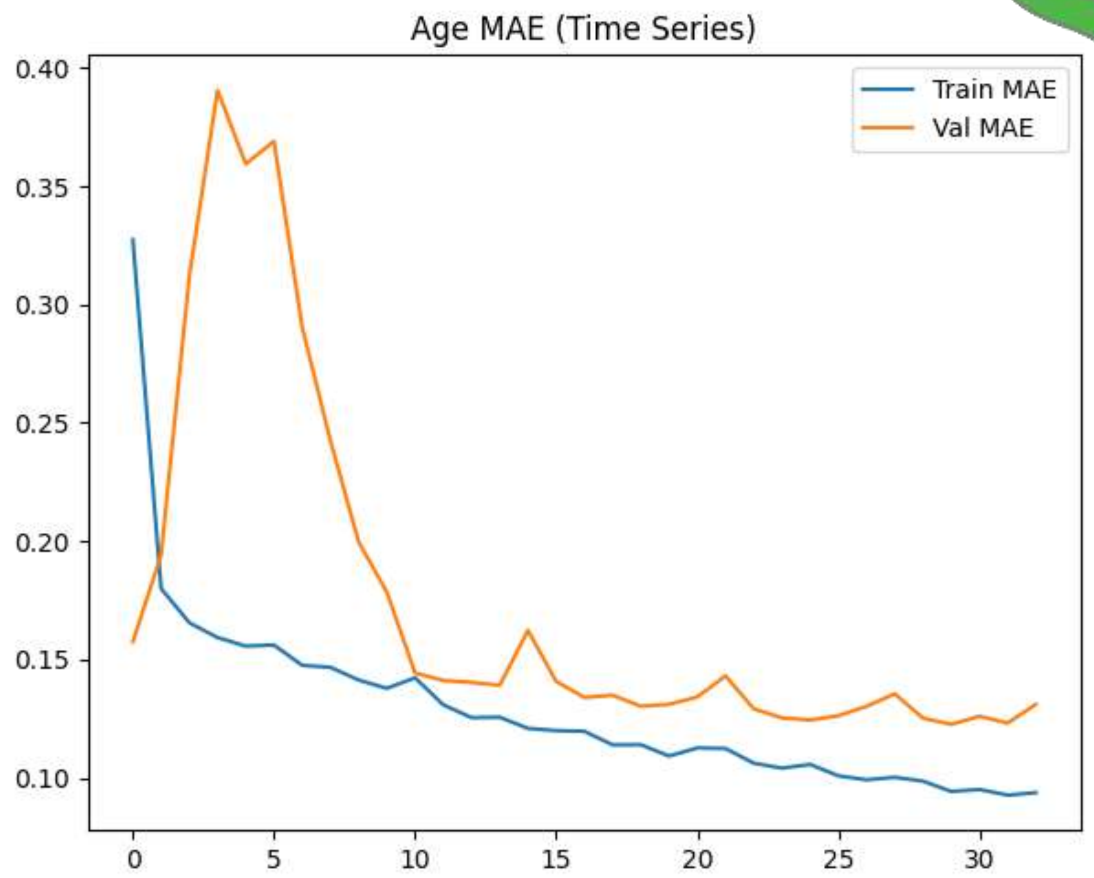


Figure: Gender Accuracy Time Series of the ShuffleNetV2 model

Bias & Fairness Analysis Graph of the models

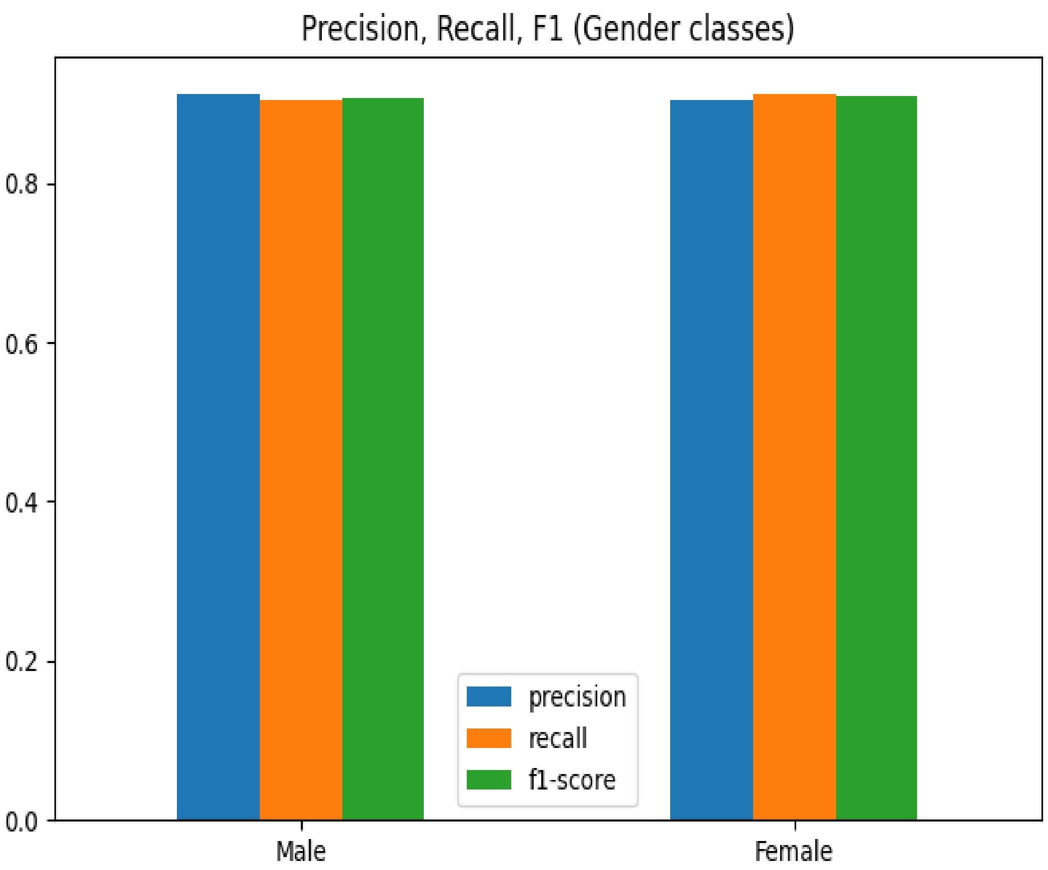


Figure: Output of EfficientNetB0 Model Gender Bias Analysis Graph.

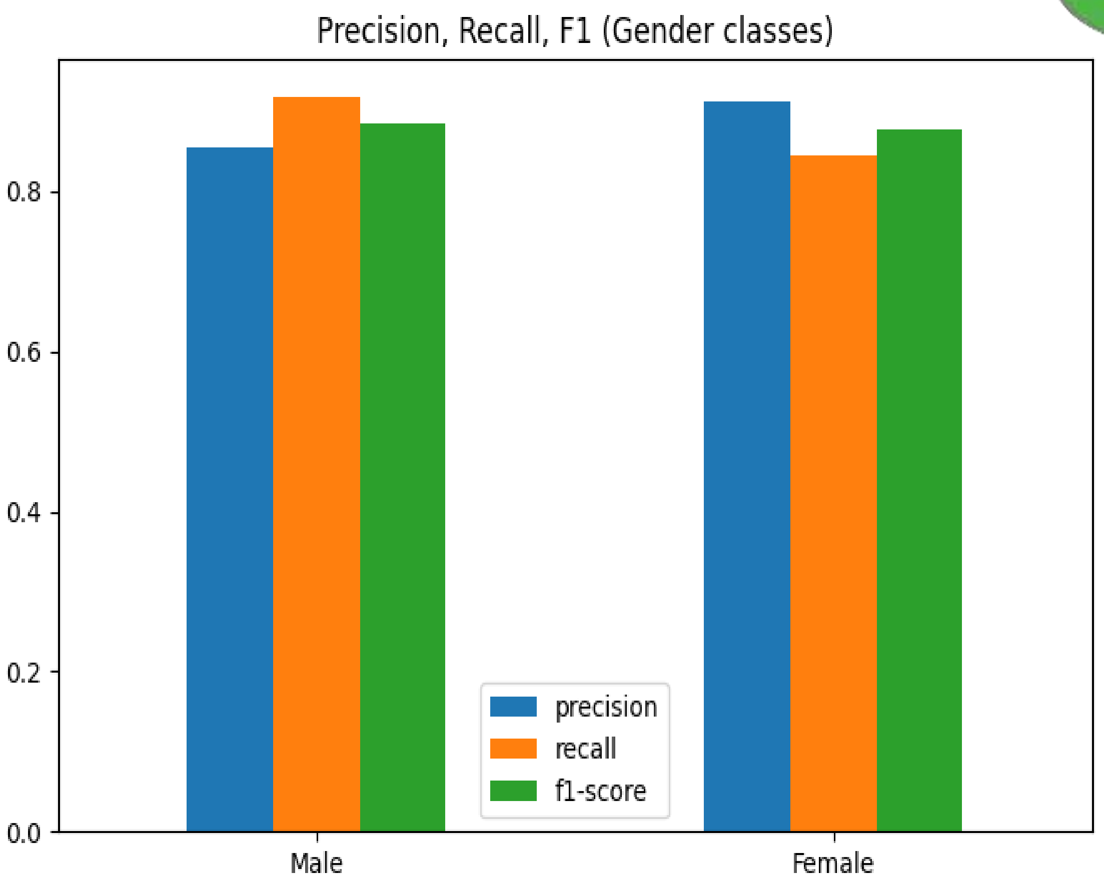


Figure: Output of MobileNetV2 Model Gender Bias Analysis Graph.

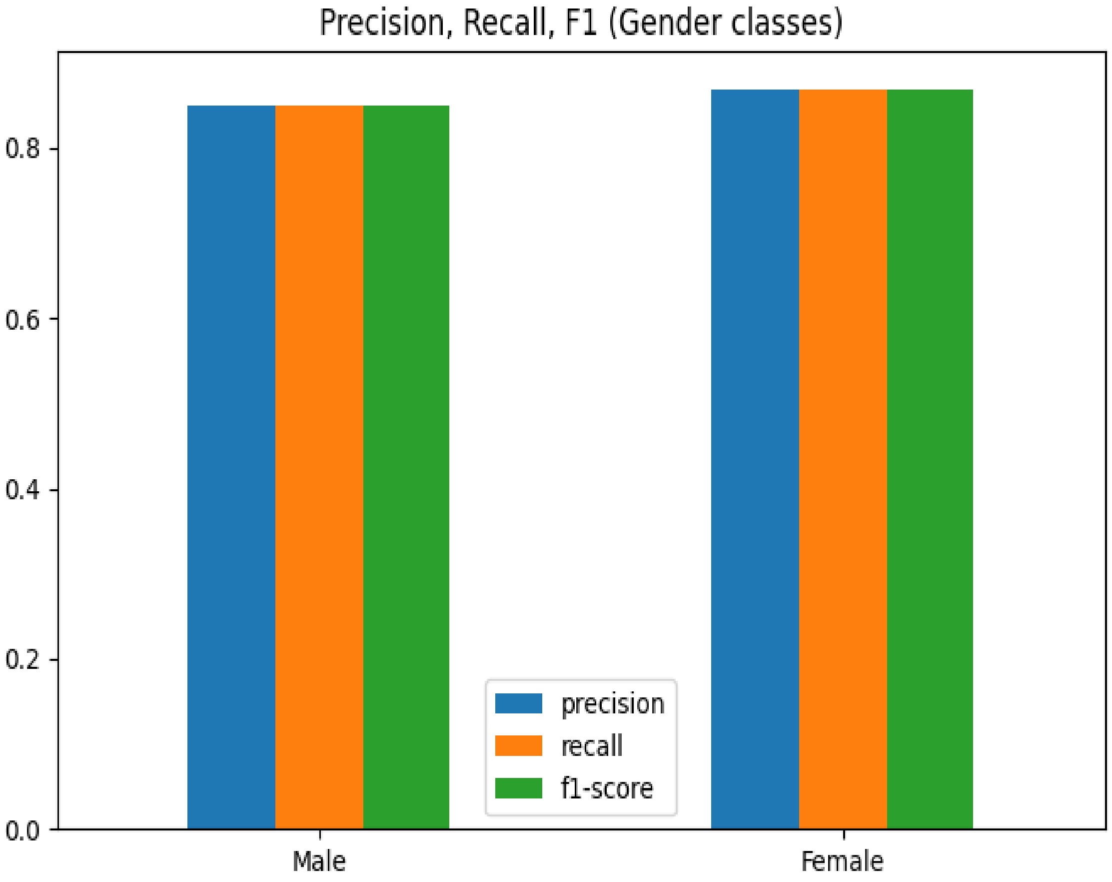


Figure: Output of Basic CNN Model Gender Bias Analysis Graph.

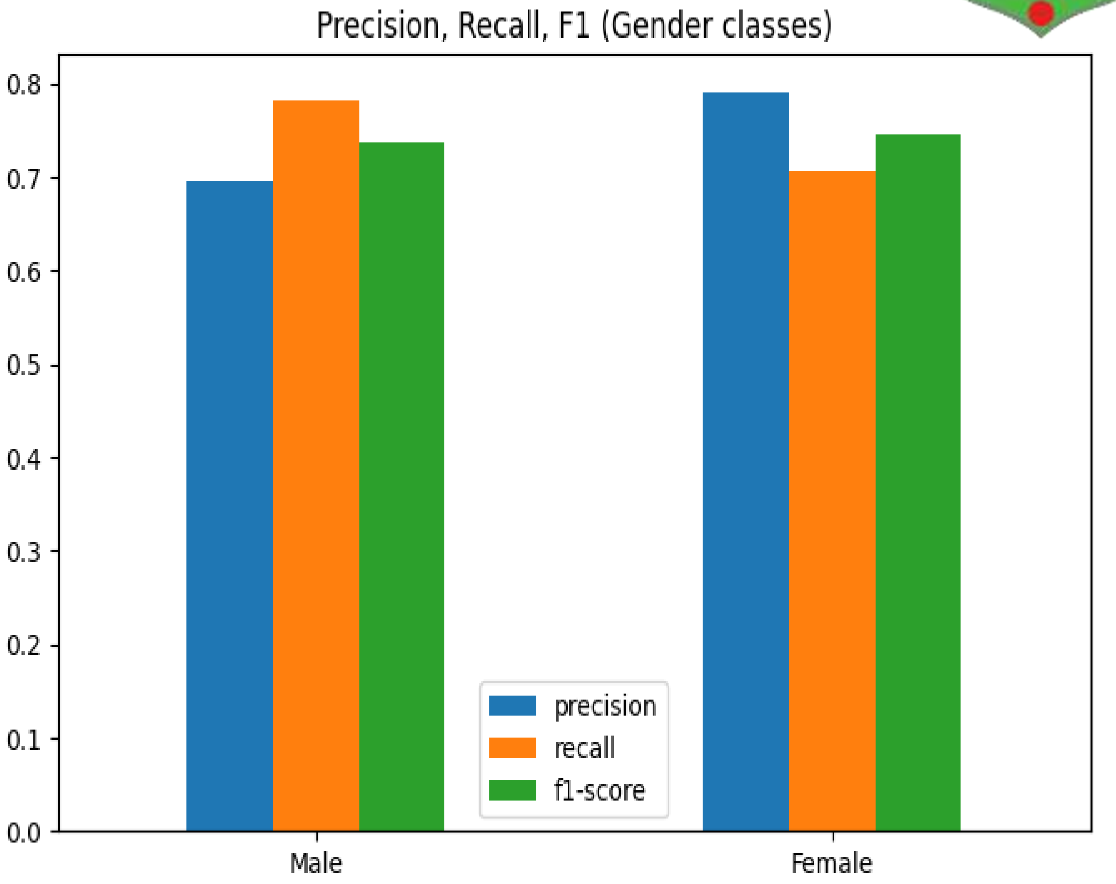


Figure: Output of Basic ShuffleNetV2 Model Gender Bias Analysis Graph.

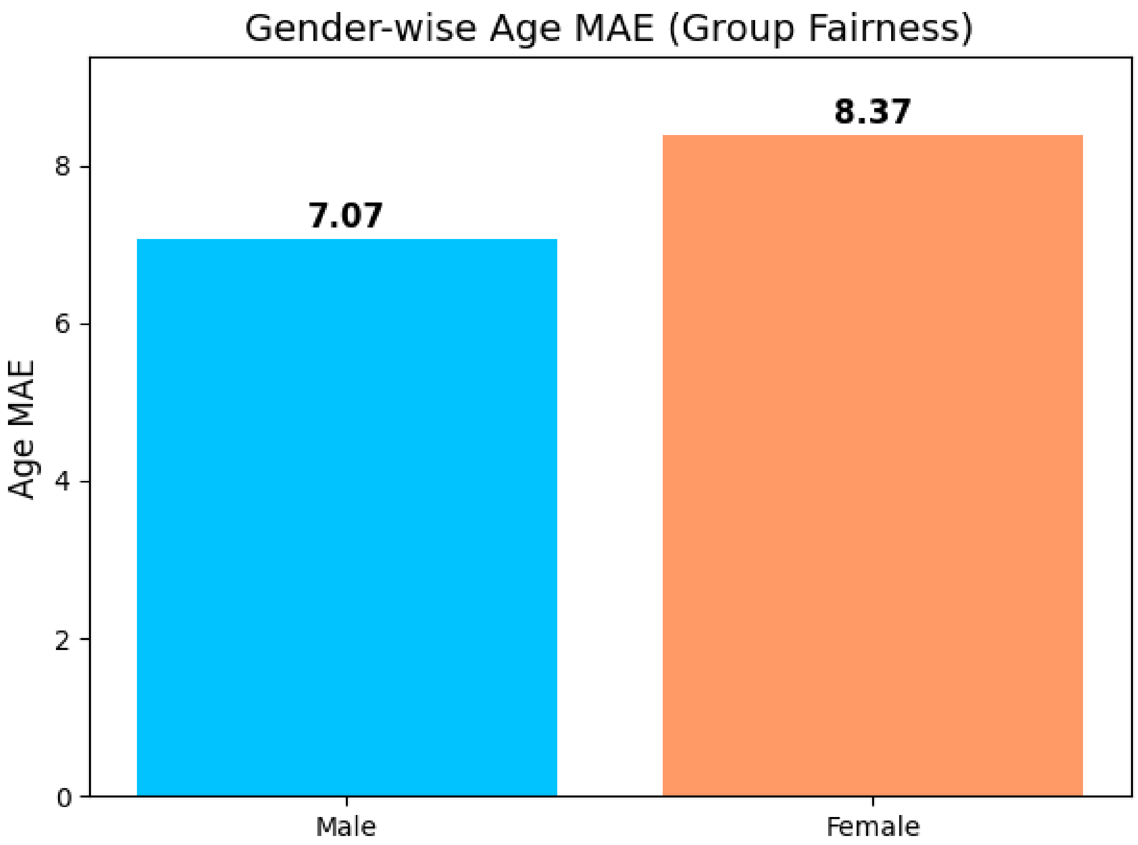


Figure: Output of EfficientNetB0 Model MAE fairness.

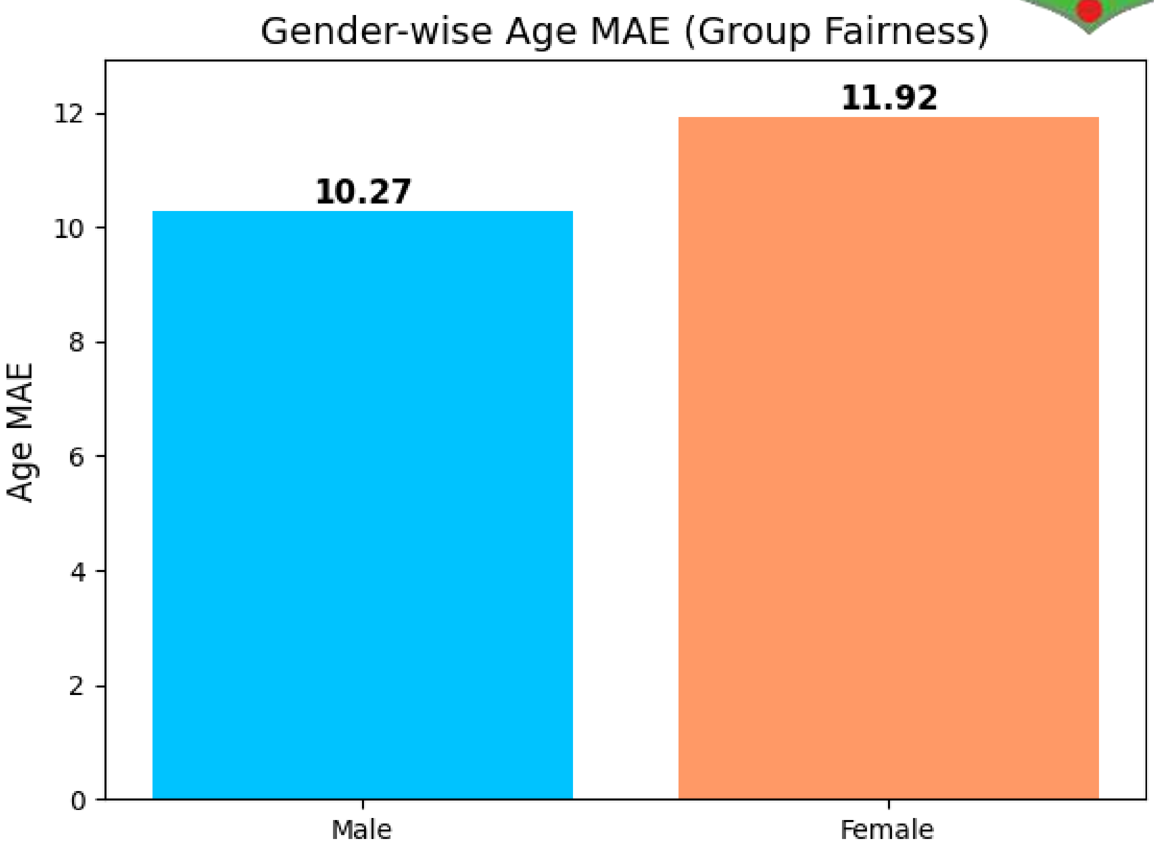


Figure: Output of MobileNetV2 Model MAE fairness.

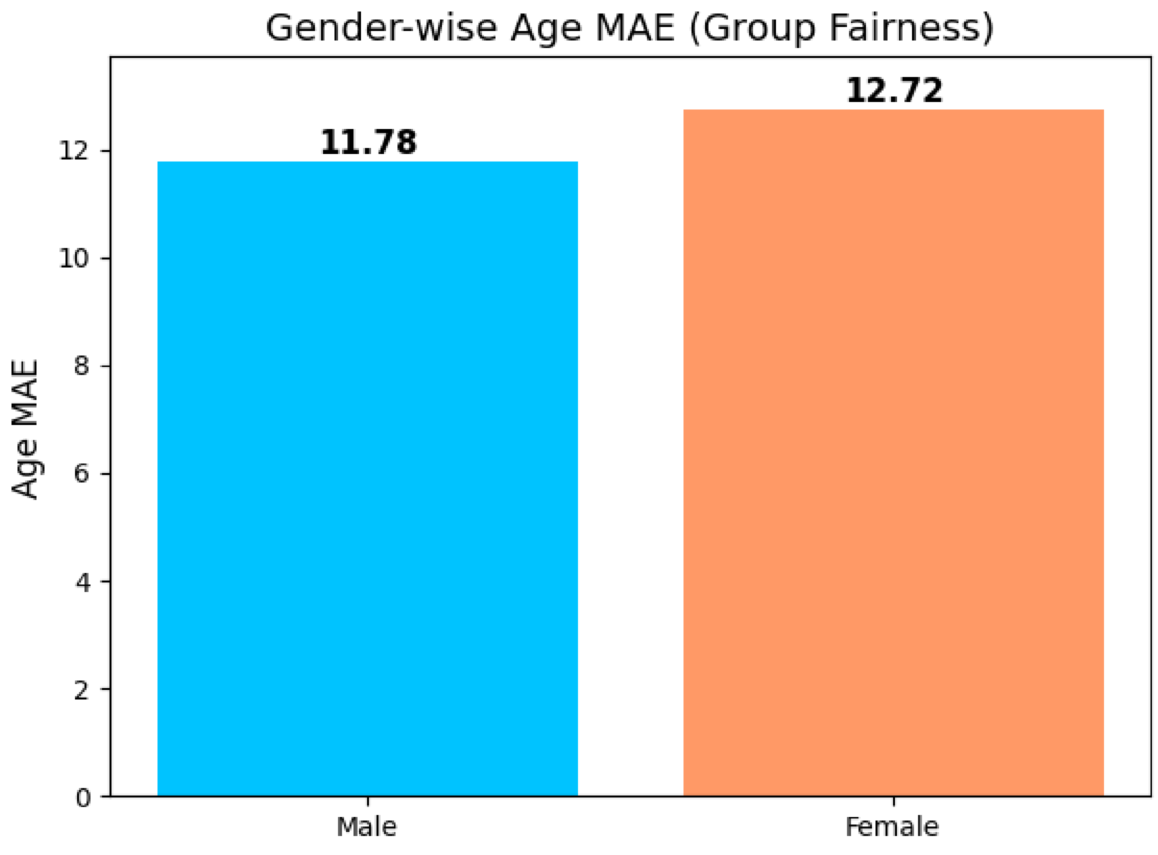


Figure: Output of Basic CNN Model MAE fairness.

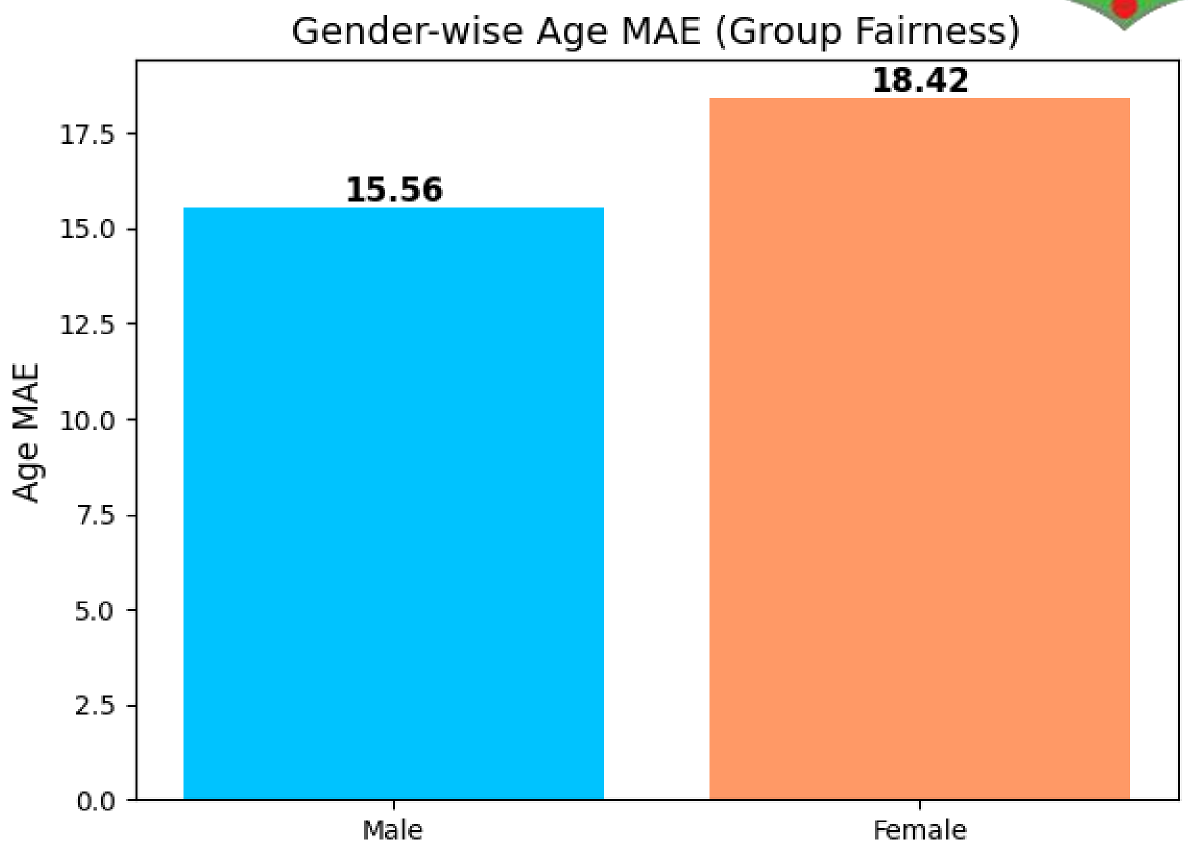


Figure: Output of ShuffleNet Model MAE fairness.

Conclusion

- ✓ Developed and compared four deep learning models on the UTKFace dataset for joint age and gender prediction.
- ✓ Designed a bias assessment pipeline to quantify performance differences across demographic groups using fairness metrics.
- ✓ Integrated Grad-CAM–based explainable AI to visualize region-wise contributions in facial age and gender predictions.

Limitations:

- **Model selection:** We compared only four deep learning models in this study. Performance could be different between other architectures.
- **Dataset coverage:** The conclusions are based on the only UTKFace dataset. For the reason, it might limit the generalization to other facial datasets or tasks.
- **XAI methods selection:** The explainability analysis relies mainly on Grad-CAM, which can be noisy and incomplete, and does not compare multiple XAI methods systematically.
- **Hardware:** Experiments were performed on some hardware, results (runtime, speed) may vary on different environments.
- **Testing in real-life examples:** This work does not contain direct spread, while model and bias analysis was fulfill .

Future works:

- ✓ Cross-dataset and cross-domain validation.
- ✓ Richer demographic and fairness modeling.
- ✓ Advanced XAI for facial analysis
- ✓ Bias evaluation focuses on a limited set of group metrics (e.g., TPR, MAE gaps) and does not yet link these disparities to downstream real-world harms or user studies.
- ✓ Create lightweight version of the models.

References

1. Akay, M., Du, Y., Sershen, C. L., Wu, M., Chen, T. Y., Assassi, S., Mohan, C., & Akay, Y. M. (2021). Deep learning classification of systemic sclerosis skin using the MobileNetV2 model. *IEEE Open Journal of Engineering in Medicine and Biology*, *2*, 104–110. https://doi.org/10.1109/ojemb.2021.3066097
2. Amin, H., Darwish, A., Hassanien, A. E., & Soliman, M. (2022). End-to-End Deep learning model for corn leaf disease classification. *IEEE Access*, *10*, 31103–31115. https://doi.org/10.1109/access.2022.3159678
3. Devi, V. S., Ramisetty, U., Ramisetty, K., & Thimmareddy, A. (2025). Real-Time age, gender and emotion detection using YOLOv8. *ITM Web of Conferences*, *74*, 01015. https://doi.org/10.1051/itmconf/20257401015
4. Dey, P., Mahmud, T., Chowdhury, M. S., Hossain, M. S., & Andersson, K. (2024). Human Age and Gender Prediction from Facial Images Using Deep Learning Methods. *Procedia Computer Science*, *238*, 314–321. https://doi.org/10.1016/j.procs.2024.06.030
5. *DSpace*. (n.d.). https://hdl.handle.net/10057/27838
6. Dwivedi, R., Kumar, R., Chopra, D., Kothari, P., & Singh, M. (2023). An Efficient Ensemble Explainable AI (XAI) Approach for Morphed Face Detection. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2304.14509

7. Gaya-Morey, F. X., Sanchez-Perez, J., Manresa-Yee, C., & Buades-Rubio, J. M. (2025). Bridging the gap in FER: addressing age bias in deep learning. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2507.07638
8. Hamza, N. R. (2025). Gender Classification from Human Face Images Using Deep Learning Based on MobileNetV2 Architecture. *Journal of Al-Qadisiyah for Computer Science and Mathematics*, *17*(1), 145–156. https://doi.org/10.29304/jqcsm.2025.17.11970
9. Hassan, B. a. R., & Dawood, F. a. A. (2024). Face-based gender classification using deep learning model. *Journal of Engineering*, *30*(01), 106–123. https://doi.org/10.31026/j.eng.2024.01.07
10. Ismail, W. N., Alsalamah, H. A., Hassan, M. M., & Mohamed, E. (2023). AUTO-HAR: An adaptive human activity recognition framework using an automated CNN architecture design. *Heliyon*, *9*(2), e13636. https://doi.org/10.1016/j.heliyon.2023.e13636
11. Kumar, R., Singh, K., Mahato, D. P., & Gupta, U. (2024). Face-based age and gender classification using deep learning model. *Procedia Computer Science*, *235*, 2985–2995. https://doi.org/10.1016/j.procs.2024.04.282
12. Levi, G., & Hassncer, T. (2015). Age and gender classification using convolutional neural networks. *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 34–42. https://doi.org/10.1109/cvprw.2015.7301352
13. Liu, J., Li, J., & Huang, G. (2025a). Research on deep learning-based image processing and classification techniques for complex networks. *Applied Mathematics and Nonlinear Sciences*, *10*(1). https://doi.org/10.2478/amns-2025-0351

14. Manresa-Yee, C., Ramis, S., & Buades, J. M. (2023). Analysis of gender differences in facial expression recognition based on deep learning using explainable artificial intelligence. *International Journal of Interactive Multimedia and Artificial Intelligence*, *9*(1), 18–27. https://doi.org/10.9781/ijimai.2023.04.003
15. Mohamed, E., Ashraf, A., Tarek, M., & Matar, W. (2025). Age and Gender Detection using Facial Images. *International Integrated Intelligent Systems*, *2*(1), 0. https://doi.org/10.21608/iiis.2025.292070.1003
16. Rasheed, J., Waziry, S., Alsubai, S., & Abu-Mahfouz, A. M. (2022). An Intelligent Gender Classification System in the Era of Pandemic Chaos with Veiled Faces. *Processes*, *10*(7), 1427. https://doi.org/10.3390/pr10071427
17. Sheoran, V., Joshi, S., & Bhayani, T. R. (2021). Age and Gender Prediction using Deep CNNs and Transfer Learning. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2110.12633
18. Singhal, A., Agrawal, K. K., Quezada, A., Aguiñaga, A. R., Jiménez, S., & Yadav, S. P. (2024a). Explainable Artificial Intelligence (XAI) model for cancer image classification. *Computer Modeling in Engineering & Sciences*, *141*(1), 401–441. https://doi.org/10.32604/cmes.2024.051363

List of Publication

1. Mausd, A., Ali, M. S., Ruku, N. I., Mallick, D., Parvez, M. S., Shanto, M. R. J., Al-Nuwaiser, W. M., & Akhund, T. M. N. U. (2025). EEG-Based Neurofeedback for ADHD in Children: Enhancing attention and reducing impulsivity using Machine learning. *International Journal of Computing and Digital Systems*, *18*(1), 1–12. https://doi.org/10.12785/ijcds/1571139415

List of Certification


aws academy

Md. Rufsan Jani Shanto
221-35-1064

Certificate of Completion for
AWS Academy Graduate - AWS Academy Machine Learning Foundations
Course hours completed
20 hours

Issued on
07/22/2025

Digital badge
https://www.credly.com/go/2cq2J87y


aws academy

Md. Rufsan Jani Shanto
221-35-1064

Certificate of Completion for
AWS Academy Graduate - AWS Academy Machine Learning for Natural Language Processing
Course hours completed
20 hours

Issued on
08/20/2025

Digital badge
https://www.credly.com/go/9FLT0CSt

THANK YOU

Question ???