Technical Summary

Team Objective:

We developed two deep learning pipelines using the FACECOM dataset: (A) for gender classification under relatively clean conditions, and (B) for face recognition under severe visual degradations such as blur, overexposure, low-light, fog, and rain. The main challenge was to ensure reliable performance, especially in Task B, across these adverse real-world scenarios.

Approach & Architecture:

For Task A, we used the CoAtNet-0 architecture from the TIMM library—a hybrid of convolution and self-attention for capturing both local and global features. We replaced its final layer with a binary classifier for gender prediction. CoAtNet-0 was selected for its efficiency on small datasets and strong generalization, even when trained from scratch.

Training Strategy:

- Preprocessing included resizing, cropping, and ImageNet normalization.
- The model was trained from scratch using the training subset of FACECOM with a weighted cross-entropy loss to account for the high class imbalance (Male >> Female).
- Data augmentation was deliberately avoided to preserve degradation characteristics.
- Early stopping and validation accuracy guided model selection.

Innovation:

We observed significant class imbalance—females were underrepresented. To address this, weighted metrics (precision, recall, F1-score) were prioritized. Our model achieved 94% F1-score on validation, despite adversarial image conditions.

Task B: Face Recognition (Few-Shot Multi-Class Classification)

Approach & Architecture:

For identity recognition, we implemented a Prototypical Network (ProtoNet) using ResNet-18 as the feature extractor, followed by a projection head that maps embeddings to a 256-dimensional latent space. This approach is ideal for low-shot learning with limited clean/distorted face samples per identity.

Training Strategy:

- Episodic training simulated N-way K-shot tasks to encourage few-shot generalization.
- Each episode included distorted (support) and clean (query) images for robust prototype formation.
- Embeddings were compared using Euclidean distance to compute class prototypes and predictions.
- No fine-tuning on the test set ensured model generalization to unseen identities.

Sampling Innovation:

Our FewShotDataset class enforced balance across episodes by requiring a minimum number of clean and distorted images per identity, reducing overfitting to well-represented classes.

Results:

The ProtoNet achieved a Top-1 Accuracy and Macro F1-score of 96% on the training set and 87% on validation, confirming strong generalization in degraded settings.

Conclusion:

Both models were designed for real-world conditions. Task A's CoAtNet-based gender classifier is optimized for imbalanced, degraded data, while Task B's ProtoNet excels in generalizing across identities with minimal supervision. Together, our approach provides a resilient AI solution for facial analytics under adverse visual conditions.