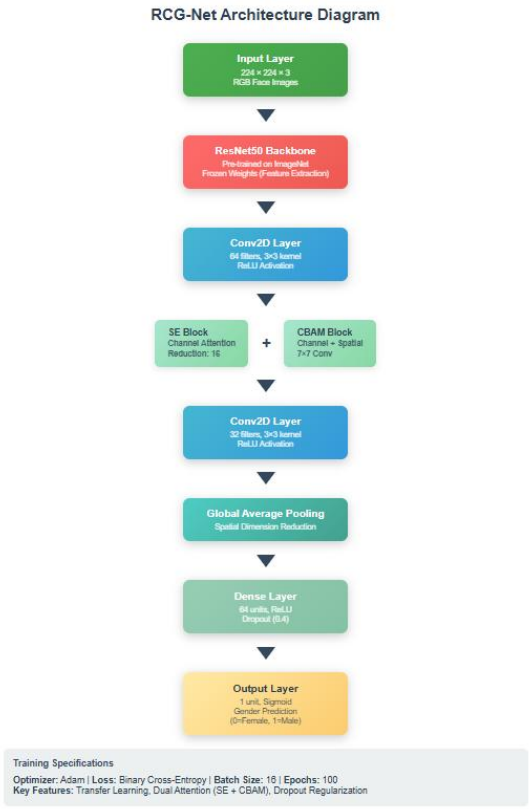
COMSYS-2025 - The 6th International Conference on Frontiers in Computing and Systems (COMSYS-2025)

Theme: Robust Face Recognition and Gender Classification under Adverse Visual Conditions

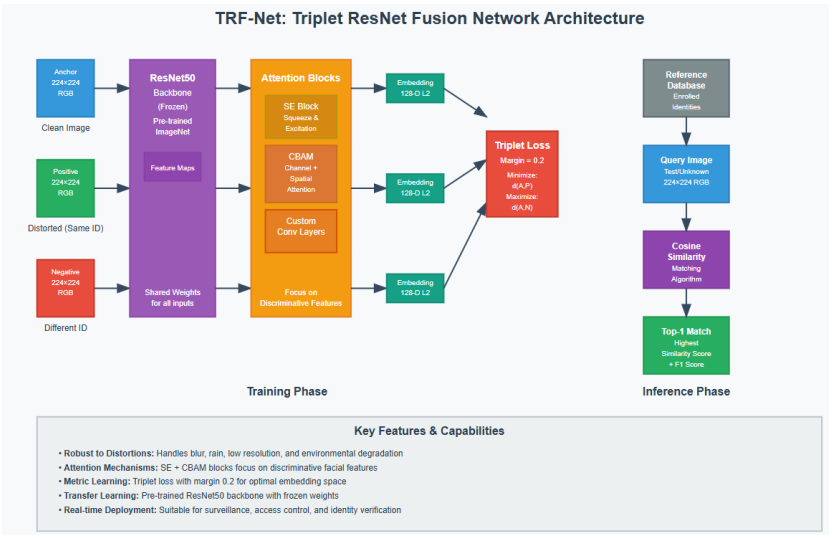
1-page technical summary:

Task A: RCG-Net - ResNet50-Based Gender Classification Network



RCG-Net presents a sophisticated deep learning architecture for binary gender classification from facial images, combining transfer learning with advanced attention mechanisms in a unified framework. The model employs a frozen ResNet50 backbone pre-trained on ImageNet as a robust feature extractor for 224×224×3 RGB input images, followed by custom enhancement layers including a 64-filter convolutional layer, a Squeeze-and-Excitation (SE) block for channel-wise attention with reduction ratio 16, a Convolutional Block Attention Module (CBAM) for dual channel and spatial attention using 7×7 convolutions, another 32-filter convolutional layer, global average pooling, a 64-unit dense layer with 0.4 dropout regularization, and finally a single-unit sigmoid output layer for binary classification (0=Female, 1=Male). The model is trained using Adam optimizer with binary cross-entropy loss over 100 epochs with batch size 16, incorporating ResNet50-specific preprocessing with mean centering and scaling. The dual attention mechanism significantly enhances feature selection by emphasizing important channels through SE blocks while simultaneously focusing on relevant spatial regions via CBAM, resulting in improved classification accuracy and demonstrating the effectiveness of combining established CNN backbones with modern attention mechanisms for robust gender classification tasks in computer vision applications, with comprehensive evaluation including confusion matrix analysis, classification reports with precision/recall/F1-score metrics, and training curve visualization.

Task B: TRF-Net: Triplet ResNet Fusion Network for Robust Face Recognition



TRF-Net is a face recognition framework built for real-world scenarios where image quality can be degraded due to environmental or acquisition issues. It combines the strength of transfer learning with advanced attention mechanisms and distance-based metric learning for high-precision identity verification and classification. The architecture is based on a frozen ResNet50 backbone pre-trained on ImageNet for feature extraction, followed by custom convolutional layers enhanced with Squeeze-and-Excitation (SE) blocks and Convolutional Block Attention Modules (CBAM). These attention blocks allow the network to focus on the most discriminative facial features by applying both channel-wise and spatial attention refinement. TRF-Net is trained using a triplet loss function with a margin of 0.2, where each training sample consists of: an anchor image (usually a clean image), a positive image (a distorted version of the same person), a negative image (an image of a different person). This training strategy teaches the network to minimize intra-class distance (anchor-positive) and maximize inter-class distance (anchor-negative) within a learned 128-dimensional L2-normalized embedding space. All images are resized to 224×224 RGB and normalized for input. During inference, TRF-Net computes embeddings for both reference identities (enrolled faces) and query/test images. Identification is performed via cosine similarity, where a test image is matched to the reference identity with the highest similarity score. A robust test pipeline has been designed where: 1. One image per identity is used to create reference embeddings.2. Test images (including unseen identities) are matched against the reference database. 3. Top-1 prediction and macro-averaged F1 scores are computed to evaluate performance. Optionally, a similarity threshold can be applied to reject unknown faces (e.g., when similarity is too low). TRF-Net demonstrates strong generalization in handling image distortions such as blur, rain, low resolution, etc., and is well-suited for real-time deployment scenarios like surveillance, access control, and identity verification in unconstrained environments.