Robust Face Recognition Using Siamese Ensemble Network In Adverse Climatic Conditions

This study presents a robust face verification approach utilizing a Siamese neural network architecture augmented by an ensemble of embedding strategies. The dataset used was the original one provided by the organizing committee; however, for experimental purposes, the dataset was selectively modified to include a focused subset. In the folder named TASK_B, a total of 50 unique identities were chosen, with each person's clear image stored in the Reference folder. Correspondingly, for each of the 50 identities, five distorted versions were manually curated and stored in the Distorted folder. These distortions included challenging real-world variations such as blurring, fog, resizing, Gaussian noise, and simulated rain, allowing the model to learn robust matching features even under adverse visual conditions. There are no subfolders; all images were named consistently to retain identity information, with filenames structured in a way that allowed positive and negative pair generation. Although the distorted images vary in their noise types, the task of matching them to their clean counterparts benefits significantly from targeted preprocessing. To enhance feature extraction, each image was resized to 160×160 pixels and converted to grayscale. A series of filtering operations were applied—median filtering for salt-and-pepper noise removal, Gaussian blurring for smoothness, and normalization to scale pixel intensities into the [0, 1] range. Each image was also reshaped to include a single channel.The architecture is centered around a Siamese network that takes image pairs as input and learns to predict whether they represent the same identity. The core embedding module combines three parallel streams: a convolutional feature extractor, a frozen ResNet50 (adapted for grayscale inputs), and a shallow FaceNet-inspired network. The feature vectors from these streams are concatenated and passed through a dense projection layer. The absolute difference between two such embeddings is fed into a sigmoid classifier for final decision-making. Training was done using binary cross-entropy loss and the Adam optimizer. The generated pairs included both positive (same person) and negative (different person) samples. Evaluation metrics revealed a Top-1 Accuracy of 0.852 and a macro-averaged F1 Score of 0.8633. The model achieved a macro-averaged precision of 0.90, recall of 0.93, and F1-score of 0.90. These results confirm the model’s strong discriminative capability, even under severe image distortions.