**Face Recognition Attendance System Framework**

**PART A: Face Recognition Attendance System Framework**

**1. Introduction**

Face recognition technology generally divides into two primary tasks: **face classification** and **face verification**.

* **Face Classification**: A closed-set recognition task that assigns a person's face to a known identity.
* **Face Verification**: An open-set task, determining if two images depict the same individual without prior knowledge of identity.

In this project, we implemented face verification leveraging a Convolutional Neural Network (CNN) to create a face recognition attendance system for enterprise settings.

**2. Face Verification Overview**

The verification task requires evaluating whether two images depict the same individual by calculating a numerical similarity score between pairs. Direct comparisons using flattened image vectors are computationally expensive and insufficiently discriminative. To address this, we employed CNNs for compact, discriminative face embeddings, significantly reducing computation and enhancing verification accuracy.

**3. Implementation Approaches**

We explored two distinct CNN training strategies:

**3.1 Face Embeddings**

The CNN backbone architecture, termed **FaceCNN**, produces embeddings (fixed-size vectors representing faces). It comprises:

* Three convolutional layers (32, 64, and 128 filters respectively), each followed by ReLU activations and max-pooling.
* Adaptive average pooling and a linear embedding layer (128-dimensional embedding).

**Insights:**

* Compact and efficient structure ideal for real-time verification tasks.
* Embeddings facilitate similarity computations, essential for face verification.

**3.2 Training Strategies**

**(a) Supervised Learning (Classification-based)**  
A classification-based approach using softmax loss trains the CNN to classify among known identities. The final embedding layer (128-dimensional) precedes a fully-connected softmax classification head predicting 4000 identities.

**Architecture:**

* CNN backbone (FaceCNN) + Fully connected (128 → 4000 identities).
* Total parameters: 625,760.

**Insights:**

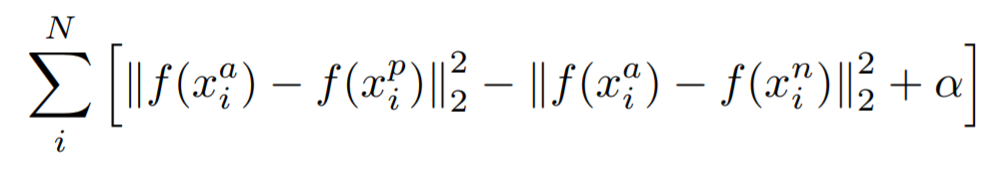
* Although effective for classification, embeddings extracted here tend to reflect class distinctions rather than intra-identity similarity, possibly limiting discriminative verification performance.

**(b) Metric Learning (Triplet Loss)**  
Triplet loss explicitly optimizes embedding separability by pulling same identities (anchor-positive pairs) closer and pushing different identities (anchor-negative pairs) apart.

**Architecture:**

* Triplet Network (FaceCNN shared weights) with anchor-positive-negative embeddings.
* Total parameters: 109,760 (shared FaceCNN backbone).

**Triplet Loss Formulation:**

**Insights:**

* Directly optimizes embedding space for verification, theoretically superior to the supervised approach for embedding discrimination.
* Requires effective sampling of triplets (positive, negative) to learn effectively.

**4. Embedding Analysis**

To visually assess embedding quality, we conducted dimensionality reduction using t-SNE plots:

* **Classifier-CNN Embeddings:** Limited intra-class clustering, suggesting weak verification separability.
* **Triplet-CNN Embeddings:** Slightly better cluster formations, aligning with the design objective of metric learning.

**Insights:**

* t-SNE plots indicated the Triplet-CNN embeddings marginally improved discriminative ability compared to the classifier approach, though still relatively weak.

**5. Similarity Distance Metric**

We compared two metrics widely used in face verification:

* **Cosine Similarity** (angle between embeddings).
* **Euclidean Distance** (direct distance between embeddings).

Experiments indicated negligible performance differences between metrics, likely due to the embeddings lacking normalization.

**6. Evaluation Metrics (ROC and AUC)**

We evaluated the verification task using Receiver Operating Characteristic (ROC) curves and Area Under Curve (AUC):

|  |  |  |
| --- | --- | --- |
| Model | Metric | AUC Score |
| Supervised Classifier | Cosine | 0.51 |
| Supervised Classifier | Euclidean | 0.51 |
| Triplet Embedding | Cosine | 0.52 |
| Triplet Embedding | Euclidean | 0.52 |

**Insights:**

* All models performed close to random (AUC ≈ 0.5).
* Triplet embedding slightly outperformed classification embeddings, aligning with theoretical expectations.

**7. Threshold-based Verification Accuracy**

Complementing ROC-AUC analysis, we employed a threshold-based decision system using Youden’s J statistic to select an optimal similarity threshold, simulating real-world deployment conditions:

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Metric | Optimal Threshold | Accuracy |
| Supervised Classifier | Cosine | 0.9985 | 51.49% |
| Supervised Classifier | Euclidean | -0.0975 | 49.01% |
| Triplet Embedding | Cosine | 0.9981 | 52.33% |
| Triplet Embedding | Euclidean | -0.0429 | 47.62% |

**Insights:**

* Threshold-based accuracy mirrors AUC findings, suggesting minimal discriminative power in the current embeddings.
* Performance clustered around random guess accuracy (50%), insufficient for practical verification deployment.

**8. Limitations and Possible Improvements**

**Current Limitations:**

* **Undertrained Models:** Limited epochs or inadequate dataset size.
* **Embedding Space:** Weak discriminative capability suggests embeddings aren't sufficiently trained.
* **Suboptimal Triplet Mining:** Randomly selected triplets may have limited effectiveness in training discriminative embeddings.
* **Data Challenges:** Poor image quality, variable lighting, and poses adversely impact embeddings.

**Suggested Improvements:**

* **Enhanced CNN Architectures:** Employ deeper, proven backbones (e.g., MobileFaceNet, ResNet).
* **Optimized Triplet Mining:** Implement semi-hard or hard-negative mining strategies to enhance training effectiveness.
* **Dataset Expansion:** Increase training set size and diversity.
* **Normalization Techniques:** Implement embedding normalization (L2) to better utilize cosine similarity.

**9. Conclusions and Reflections**

Our implementation explored classification and metric learning for face verification, utilizing compact CNN embeddings evaluated using ROC-AUC and threshold-based accuracy metrics.

* **Findings:** Both methods yielded minimal effectiveness, with metric learning showing slight but insignificant superiority.
* **Metric Selection:** Cosine similarity and Euclidean distance performed similarly, suggesting embeddings must be normalized for meaningful comparison.
* **Practical Implications:** Current embeddings are insufficiently discriminative for reliable attendance verification systems.

Future enhancements, particularly adopting deeper architectures, improved triplet selection strategies, and expanded datasets, are crucial to achieving practically viable face recognition systems.

**PART B: Anti-Spoofing**

**PART C: Emotion Detection**

**PART D: User Interface**