SELF-LEARNING FACE RECOGNITION SYSTEM – DESIGN REPORT

Submitted by: Janani D
For: Zeex AI Internship Technical Task (Q4)

1. INTRODUCTION:

A self-learning face recognition system continuously improves its accuracy over time using user feedback, without retraining from scratch.
The system's goal is to remain adaptive, privacy-preserving, and resistant to forgetting previously learned faces.

2. BASE MODEL SET UP:

A pre-trained model such as **FaceNet** or **VGGFace** acts as a **feature extractor**, converting faces into embeddings (numerical vectors capturing facial identity). These embeddings are stored in a **face database** and compared during recognition.

3. MODEL UPDATING (SELF-LEARNING MECHANISM)

When new faces or feedback are received:

1. The new image is converted to an embedding and compared with stored ones.
2. If confidence is low, the user/admin provides the correct label.
3. The system updates incrementally using **Transfer Learning** -keeping the base feature extractor frozen while updating the **classifier (KNN/SVM)** or the **L2-distance threshold**. This lightweight approach follows **Few-Shot Learning / Metric Learning**, enabling the system to recognize new identities from very few examples.

4. AVOIDING CATASTROPHIC FORGETTING:

As new faces are added, the model must retain older knowledge. To prevent forgetting:

- **Experience Replay (Rehearsal Buffer):** Stores a small sample of past embeddings for occasional reuse during updates.
- **Elastic Weight Consolidation (EWC):** A regularization method that penalizes changes to parameters critical for previous tasks, ensuring stable feature representation.

These strategies preserve performance while enabling incremental updates.

5. PRIVACY PRESERVATION:

Because face data is sensitive, the system uses:

- **Anonymized Embeddings:** Only numeric features are stored, not raw images.
- **Encryption:** Protects all stored and transmitted data.
- **Federated Learning (Optional):** Model updates occur locally, sharing only weight updates — not personal data.

This ensures strong privacy compliance and secure deployment.

6. DATA VALIDATION AND QUALITY CONTROL:

Before adding new feedback:

- The system checks image quality, face clarity, and label authenticity.
- It monitors for **Data Drift** or **Concept Drift** — shifts in data patterns due to lighting, camera quality, or demographics.
  If significant drift occurs, the system alerts the admin or performs fine-tuning using new representative data.

7. CONTINUOUS IMPROVEMENT WORKFLOW:

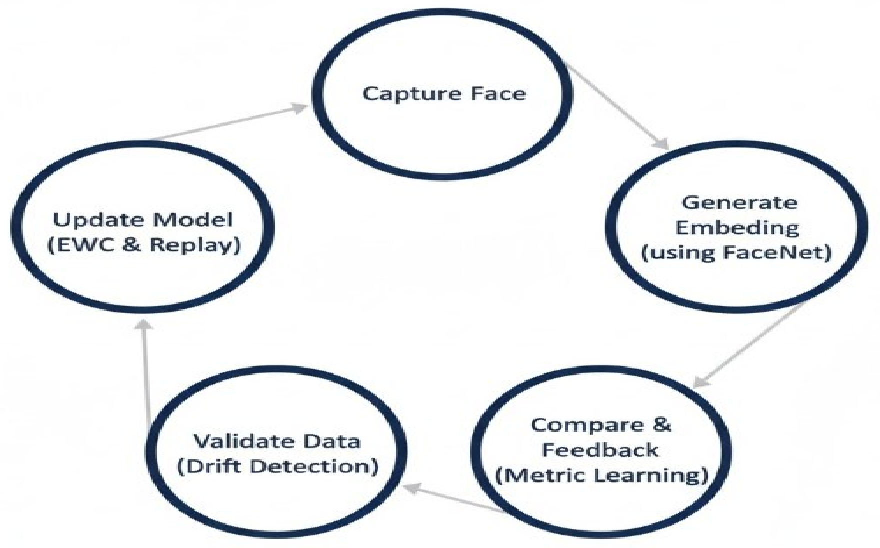
1. Capture face → Generate embedding
2. Compare with known database
3. Request feedback if confidence is low
4. Validate data and check for drift
5. Update classifier incrementally
6. Apply EWC / Replay for stability
7. Monitor accuracy and privacy metrics

This closed feedback loop ensures that the system evolves intelligently while preserving old knowledge.

8. CONCLUSION:

The proposed self-learning design enables continuous improvement through **transfer learning**, **few-shot adaptation**, and **regularized updates** like EWC. By combining incremental updates, drift detection, and privacy safeguards, the system achieves long-term adaptability without catastrophic forgetting — ideal for authentication, attendance, and surveillance applications.

Figure: Continuous Self-Learning Feedback Loop


Capture Face
Generate Embeding (using FaceNet)
Compare & Feedback (Metric Learning)
Validate Data (Drift Detection)
Update Model (EWC & Replay)