Milestone-2: Explanation of Embedding Comparison

The AI Guard system uses a laptop webcam to recognise trusted individuals and detect strangers entering the room. A key component of this system is face recognition, which relies on pre-trained deep learning models to convert facial images into numerical embeddings and compare them to a trusted database.

Face Embeddings-

- Each face detected by the camera is processed through a pre-trained network from the face_recognition library (which wraps dlib's face recognition model).
- The output is a 128-dimensional vector, called a face embedding, which encodes unique facial features.
- These embeddings allow the system to represent faces numerically in a high-dimensional space where:
    - Embeddings of the same person cluster closely together.
    - Embeddings of different people are far apart.

Enrollment of Trusted Faces-

1. For each trusted individual, multiple images are captured under varied lighting and angles.
2. Each image is converted into an embedding.
3. Embeddings are stored in a dictionary with the person's name as the key.
4. These embeddings are saved (np.save) for reuse, allowing the system to recognize trusted individuals without re-enrollment every time.

Comparing Embeddings

When a face is detected in real-time:

1. **Compute embedding**: Convert the detected face to a 128-dimensional embedding.
2. **Compute distances**: Calculate the **Euclidean distance** between the new embedding and all stored embeddings for each trusted person:

```
distances = face_recognition.face_distance(embeds, new_embedding)
```

3. **Find minimum distance**: Identify the closest match:

```
if np.min(distances) < tolerance:

recognized = True
```

4. **Decision logic**:
    - **Distance < tolerance** → face is recognized as that person.
    - **Distance ≥ tolerance** → face is unrecognized (potential intruder).

Conclusion-

By converting facial images into embeddings and comparing them with a trusted database, the system can accurately identify authorised individuals while detecting strangers. The use of multiple embeddings per person and a carefully chosen similarity threshold ensures robustness against variations in lighting, angles, and expressions. This embedding-based approach provides a reliable foundation for automated security, enabling the system to welcome trusted users and escalate appropriately for unknown individuals.