TWPE EXERCISE

DeepFace : Closing the Gap to Human-Level Performance in Face Verification

- Summary:

The research paper titled "DeepFace: Closing the Gap to Human-Level Performance in Face Verification" by Yaniv Taigman, Ming Yang, Marc'Aurelio Ranzato, and Lior Wolf, aims to improve face recognition accuracy by addressing two crucial steps in the conventional pipeline: <u>alignment and representation</u>.

Traditionally, face recognition involves four stages: detect, align, represent, and classify. The authors focus on the alignment and representation steps, proposing a approach that utilizes explicit <u>3D face modeling</u> for accurate alignment and a nine-layer deep neural network for representation.

- Key Points:

1. **Alignment Step Enhancement:**

   The authors employ 3D face modeling to apply a piecewise transformation, enhancing the alignment step in face recognition. This improves the accuracy of aligning facial features, crucial for subsequent stages.

2. **Representation Step Improvement:**

   A nine-layer deep neural network is utilized to derive face representations. This network comprises more than 120 million parameters and includes locally connected layers.

3. **Performance:**

   The proposed method achieves a significant improvement in accuracy, reaching 97.35% on the Labeled Faces in the Wild (LFW) dataset. This represents a reduction in error of more than 27% compared to the current state

- Datasets:

  We evaluate the proposed DeepFace system, by learning the face representation on a very large-scale labelled face dataset collected online

  1. **SFC (Social Face Classification) Dataset:**

     The proposed face representation is learned from a large collection of photos from Facebook, referred to as the Social Face Classification (SFC) dataset

     The SFC dataset includes 4.4 million labelled faces from 4,030 people each with 800 to 1200 faces, where the most recent 5% of face images of each identity are left out for testing.

  2. **LFW (Labelled Faces in the Wild) Dataset:**

     The representations are then applied to the Labelled Faces in the Wild database (LFW), which is the de facto benchmark dataset for face verification in unconstrained environments.

     The LFW dataset consists of 13,323 web photos of 5,749 celebrities which are divided into 6,000 face pairs in 10 splits.

  3. **YTF (YouTube Faces) Dataset:**

     YouTube Faces (YTF) dataset, which is modelled similarly to the LFW but focuses on video clips.

     The YTF dataset collects 3,425 YouTube videos of 1,595 subjects (a subset of the celebrities in the LFW). These videos are divided into 5,000 video pairs and 10 splits and used to evaluate the video-level face verification.

This remarkable achievement brings DeepFace close to human-level performance in unconstrained face recognition. DeepFace, a face recognition system, has gained widespread recognition due to its outstanding performance in verifying faces across various datasets like Labeled Faces in the Wild (LFW) and YouTube Faces (YTF).

DeepFace excels in its robustness to factors like varying lighting, poses, and expressions. Its efficiency in both implementation and computation further enhances its applicability in real-world scenarios.