Assignment 12

Name: Soumya Sudhakar Gite
Referral ID: SIRSS2276

FACE DETECTION IN LIVE VIDEO FEED

- What Is Face Detection?

  Face detection is a computer technology that determines the location and size of a human face in digital images. Given an image, the goal of facial recognition is to determine whether there are any faces and return the bounding box of each detected face (see object detection).

  Face detection is the necessary first step for all facial analysis algorithms, including face alignment, face recognition, face verification, and face parsing. Also, facial recognition is used in multiple areas such as content-based image retrieval, video coding, video conferencing, crowd surveillance, and intelligent human-computer interfaces.

  The detecting of human faces is a difficult computer vision problem. Mainly because the human face is a dynamic object and has a high degree of variability in its appearance. In recent years, facial recognition techniques have achieved significant progress.

  However, high-performance face detection remains a challenging problem, especially when there are many tiny faces.

  There are two types of approaches to detect facial parts,

  a. Feature-based
  b. Image-based approaches.

a. Feature-based approach

- Technique: Feature-based methods try to find invariant features of faces for detection. The underlying idea is based on the observations that human vision can effortlessly detect faces in different poses and lighting conditions, so there must be properties or features which are consistent despite those variabilities.
- Examples: Edge detectors commonly extract facial features such as eyes, nose, mouth, eyebrows, skin color, and hairline. Based on the extracted features, statistical models were built to describe their relationships and verify a face's presence in an image.
- Advantages: Easy to implement, the traditional approach

- Disadvantages**:** A major problem of feature-based algorithms is that the image features can be severely corrupted due to illumination, noise, and occlusion. Also, feature boundaries can be weakened for faces, and shadows can cause strong edges, which together render perceptual grouping algorithms useless.

b. Image-based approach

- Technique: Image-based methods try to learn templates from examples in images. Hence, appearance-based methods rely on machine learning and statistical analysis techniques to find the relevant characteristics of "face" and "no-face" images. The learned characteristics are in the form of distribution models or discriminant functions that is applied for face detection tasks.

- Examples: Image-based approaches include neural networks (CNN), support vector machines (SVM).
- Advantages: Good performance, higher efficiency
- Disadvantages: Difficult to implement. Dimensionality reduction is usually required for the sake of computation efficiency and detection efficacy. This means reducing the dimensionality of the feature space with consideration by obtaining a set of principal features, retaining meaningful properties of the original data.

The face detection dataset WIDER FACE has a high degree of variability in scale, pose, occlusion, expression, appearance, and illumination.

- Face Detection Methods: Multiple face detection techniques have been introduced.

  i. Knowledge-Based: The knowledge-based method depends on the set of rules, and it is based on human knowledge to detect the faces. Ex- A face must have a nose, eyes, and mouth within certain distances and positions with each other. The big problem with these methods is the difficulty in building an appropriate set of rules. There could be many false positive if the rules were too general or too detailed. This approach alone is insufficient and unable to find many faces in multiple images.

  ii. Feature-Based: The feature-based method is to locate faces by extracting structural features of the face. It is first trained as a classifier and then used to differentiate between facial and non-facial regions. The idea is to overcome the limits of our instinctive knowledge of faces. This approach divided into several steps and even photos with many faces they report a success rate of 94%.

iii. Template Matching: Template Matching method uses pre-defined or parameterized face templates to locate or detect the faces by the correlation between the templates and input images. Ex- a human face can be divided into eyes, face contour, nose, and mouth. Also, a face model can be built by edges just by using edge detection method. This approach is simple to implement, but it is inadequate for face detection. However, deformable templates have been proposed to deal with these problems.

- The Main Challenges

Challenges in face detection are the reasons which reduce the accuracy and detection rate of facial recognition. These challenges are complex background, too many faces in images, odd expressions, illuminations, less resolution, face occlusion, skin color, distance, and orientation, etc.

- **Unusual expression.** Human faces in an image may show unexpected or odd facial expressions.
- **Illuminations.** Some image parts may have very high or low illumination or shadows.
- **Skin types.** Detecting faces of different face colors is challenging for detection and requires a wider diversity of training images.
- **Distance.** If the distance to the camera is too high, the object size (face size) may be too small.
- **Orientation.** The face orientation and angle to the camera impact the rate of face detection.
- **Complex background.** A high number of objects in a scene reduces the accuracy and rate of detection.
- **Many faces in one image.** An image with a high number of human faces is very challenging for an accurate detection rate.
- **Face occlusion.** Faces may be partially hidden by objects such as glasses, scarves, hands, hairs, hats, and other objects, which impacts the detection rate.
- **Low resolution.** Low-resolution images or image noise impacts the detection rate negatively.

Face Detection Applications

- **Crowd surveillance.** Face detection is used to detect crowds in frequented public or private areas.
- **Human-computer interaction.** Multiple human-computer interactionbased systems use facial recognition to detect the presence of humans.
- **Photography.** Some recent digital cameras use face detection for autofocus. Mobile apps use facial recognition to detect regions of interest in slideshows.
- **Facial feature extraction.** Facial features like nose, eyes, mouth, skin color, and more can be extracted from images.

- **Gender classification.** Applications are built to detect gender information with face detection methods.
- **Face recognition.** A face recognition system is designed to identify and verify a person from a digital image or video frame.
- **Marketing.** Face detection is becoming more and more important for marketing, analyzing customer behavior, or targeted advertising.
- **Attendance.** Facial recognition is used to detect the attendance of humans.
- It is often combined with biometric detection for access management.

Datasets used for Face Recognition

- **Annotated Faces in the Wild Dataset (**AFW**).** The AFW dataset is built using Flickr images. It includes 205 images with 473 labeled faces. For each face, image annotations include a rectangular bounding box, 6 landmarks, and the pose angles.
- **PASCAL Face Dataset (**PASCAL FACE**).** This dataset is used for facial recognition and face recognition; it is a subset of the PASCAL VOC and contains 1‘335 labeled faces in 851 images with large face appearance and pose variations.
- **MIT Face Dataset:** The MIT-CBCL face recognition database contains a training set (2’429 faces, 4’548 non-faces) and a test set (472 faces, 23’573 non-faces).
- **Face Detection Data Set and Benchmark:** The dataset contains 5‘171 faces annotated in 2‘845 images with a wide range of difficulties, such as occlusions, difficult poses, and low image resolutions. These images are used to train with large appearance changes, heavy occlusions, and severe blur degradations that are prevalent in detecting
- **CMU Multi-PIE Database (**PIE**).** The CMU Multi-PIE Face Database contains 41’368 images of 68 people, each person under 13 different poses,
    - 43 different illumination conditions, and 4 different expressions.

- **Surveillance Cameras Face: SC** face is a database of static images of human faces. The images were taken in an uncontrolled indoor environment using five video surveillance cameras of various qualities.
    - The dataset contains 4’160 static images (visible and infrared spectrum) of 130 subjects.
- **WIDER FACE dataset: The** face detection benchmark dataset includes 32’203 images and 393’703 labeled faces with a high degree of variability in scale, pose, and occlusion, making face detection extremely challenging. Also, the WIDER FACE dataset is organized based on 61 event classes.

Conclusion

Face recognition is an emerging technology that can provide many benefits. Face recognition can save resources and time, and even generate new income streams, for companies that implement it right.