

Introduction

D-LORD is a large-scale dataset that aims at strengthening the research on low-resolution disguised face detection. It leverages surveillance videos where the subjects may or may not be wearing disguise accessories, such as hats, beards and mustaches, sunglasses, and face masks. In addition, D-LORD contains HR mugshots of the subjects. We benchmark the performance of various face detection, face recognition, face restoration, and frame selection algorithms on the proposed D-LORD.

Objectives

D-LORD is a large-scale dataset consisting of HR mugshot images and LR videos of 2,100 subjects. D-LORD also addresses the challenging problem of facial disguises. The subjects appear to be donning various disguise artifacts such as face masks, hats, sunglasses, monkey caps, and wigs. D-LORD is the first large-scale real-world dataset that addresses the complex problem of low-resolution face recognition under disguise. D-LORD will be released to the research community to further strengthen state-of-the-art in face recognition.

Funding Agency

D-LORD is created by DYSL-AI Bangalore and IIT Jodhpur on behalf of Defense Research and Development Organization, India.

Grant no. DYSL-AI/CARS/20-21/02

Composition

D-LORD consists of 14,098 videos with over 1.2 million frames of 2,100 subjects. These videos are captured using various surveillance cameras under different lighting conditions. The subjects in these videos appear with and without wearing disguise artifacts such as face masks, hats, sunglasses, monkey caps, and wigs. Along with videos, D-LORD also provides HR images of the related subjects. HR images of the subjects are mugshots captured in a constrained environment, while LR images are the frames selected from surveillance videos. In all these videos, the subject walks toward the surveillance camera. The props occlude the facial region and thus can be viewed as disguise accessories where an individual tries to obfuscate the identity.

Labeling and Nomenclature

D-LORD also provides information about bounding boxes for the facial region. Bounding boxes for each HR image and each video frame are manually annotated. There are 5 HR mugshot images in different poses and disguise artifacts for each subject. Each HR image is named as "pose_disguise_subjectID.jpg", where "pose" can be "st" (straight), "hl" (horizontal left), "hr" (horizontal right), "vd" (vertical down), "vu" (vertical up), disguise can be "np" (no prop), "mk" (mask), "sg" (sunglasses), "wg" (wig), "mc" (monkey cap), "hs" (hats and scarves).

Directory Structure

For releasing the dataset, we divide the videos and keep them in separate directories based on different scenes, i.e., s1, s2, and s3. For better convenience, we create different sub-directories based on disguise categories and keep them inside scene directories. Inside

these sub-directories, there are video frames annotated with their information. Each frame is named as "SubjectID_camera_scene_disguise_frameNo.jpg", where "camera" is "rgbn_08" in our case, "scene" can be "s1/s2/s3" and "frameNo" ranges from 1 to n. Each directory that contains mugshot images or video frames is also accompanied by a ".json" file consisting of the annotation information.

Intended Use

The proposed D-LORD database is useful for evaluating face recognition algorithms for surveillance-based applications, where probes are affected by low-resolution and facial occlusions. D-LORD offers 5 HR mugshot images for each subject (straight, vertical up, vertical down, horizontal left, and horizontal right) and LR videos under various settings. The variety and the size of D-LORD make it the most extensive database available for research on face recognition. To the best of our knowledge, low-resolution face recognition under disguise has not been addressed in the past, and based on the results observed in the paper, it can be observed that significant research is required to address this challenge. Specifically, we can utilize the database for:

RQ 1: evaluating the performance of face detection in low-resolution face images and disguised images.

RQ 2: matching LR probe images against HR gallery images.

RQ 3: understanding how disguise affects the low recognition performance, and

RQ 4: reconstructing HR images for LR disguised images.

RQ 5: detecting the disguised (occluded) region in low-resolution images.

Age Distribution

18-27 years: 37.33%,

28-37 years: 49.33%,

38-47 years: 12.50%,

48-60 years: 0.84%.

Gender Distribution

Train: 60% Male, 40% Female,

Test: 45.33% Male, 54.66% Female.

Height and Weight

Average height: 165.75cm;

Average weight: 66.23 KG.

Scenes

S1) No sunlight on camera/person (normal brightly lit environment),

S2) Outdoor daytime sun facing,

S3) Outdoor daytime opposite sunlight (or backlighting).

Camera Specifications

Sensor size > 1/2.7",

Resolution >= 4MP,

Dynamic range > 100dB,

Minimum illumination for capturing color image =0.2lux,
Frame rate =5fps.

Camera positioning for scenes 1, 2, and 3 = RGB camera with a narrow view at heights 8ft. and 12ft. is installed on a pole directly in front of the person walking. Another RGB camera with a wider view is installed at the height of 8ft.

Licensing Information and Maintenance

The dataset is password protected, and the password will be shared once the data license agreement is signed by the requester. The dataset will be released under the license CC BY-NC-SA.

Limitations

As the proposed dataset is composed of only Indian subjects, it can be used to strengthen the face recognition model against racial bias when augmented with existing datasets containing images of subjects pertaining to different origins/nationalities. It is our assertion that the proposed database can be a useful benchmark and a valuable asset to the research community.