DecordFace: A Framework for Corrupted and Degraded Face Recognition

April 14, 2025

Motivation

For what purpose was the dataset created? Was there a specific task in mind? Was there a specific gap that needed to be filled? Please provide a description.

The Degraded and Corrupted Face Recognition Evaluation (DecordFace) framework is designed to aid the development of robust face recognition systems through thorough evaluation of a variety of image corruptions. Although benchmark datasets exist for evaluating computer vision models on various downstream tasks like classification and object detection in the presence of common corruptions, there has been no such benchmark to measure the test efficacy of FR models. This has led to scattered research on the corruption performance of various FR models under different settings, making corruption-based performance comparison between FR models difficult. DecordFace addresses this gap in literature.

Who created this dataset (e.g., which team, research group) and on behalf of which entity (e.g., company, institution, organization)?

The DecordFace framework utilizes existing face recognition datasets. All the face images in the paper are either taken or generated from face images present in existing datasets. This research is conducted in IIT Jodhpur at the Image Analysis and Biometrics (IAB) Lab.

Who funded the creation of the dataset? If there is an associated grant, please provide the name of the grantor and the grant name and number. The dataset creation was done using the lab servers.

Composition

What do the instances that comprise the dataset represent (e.g., documents, photos, people, countries)? Are there multiple types of instances (e.g., movies, users, and ratings; people and interactions between them; nodes and edges)? Please provide a description. The framework is created by applying the corruptions to the test images of five standard FR datasets, namely AgeDB, CALFW, CFP-FP, CPLFW, and IJB-C, to cre-

ate their corresponding corrupted variants. Each image is a face image. The framework consists of 16 image corruptions, namely Gaussian noise, impulse noise, shot noise, speckle noise, defocus blur, Gaussian blur, glass blur, motion blur, zoom blur, brightness, contrast, saturate, jpeg compression, pixelate, elastic transform, and spatter. Each corruption is applied at five increasing levels of severity to each image.

How many instances are there in total (of each type, if appropriate)? Taken together, the DecordFace framework comprises over 126 million image verification pairs formed using the face images from the datasets. The AgeDB, CALFW, and CPLFW variants consist of 480k verification pairs each. The CFP-FP variant consists of 560k verification pairs, and IJB-C consists of 125M verification pairs.

Does the dataset contain all possible instances or is it a sample (not necessarily random) of instances from a larger set? If the dataset is a sample, then what is the larger set? Is the sample representative of the larger set (e.g., geographic coverage)? If so, please describe how this representativeness was validated/verified. If it is not representative of the larger set, please describe why not (e.g., to cover a more diverse range of instances, because instances were withheld or unavailable).
The verification pairs used in DecordFace are the verification pairs from the standard FR datasets.

What data does each instance consist of? "Raw" data (e.g., unprocessed text or images) or features? In either case, please provide a description.
Each verification pair is either a genuine pair (both face images belong to the same person) or an impostor pair (the face images belong to different people).

Is there a label or target associated with each instance? If so, please provide a description.
The genuine pair has a target value of *1*, while the impostor pair has a target value of *0*.

Is any information missing from individual instances? If so, please provide a description, explaining why this information is missing (e.g., because it was unavailable). This does not include intentionally removed information, but might include, e.g., redacted text.
None.

Are relationships between individual instances made explicit (e.g., users' movie ratings, social network links)? If so, please describe how these relationships are made explicit.
The genuine and impostor pairs for evaluation are fixed and annotated as such.

Are there recommended data splits (e.g., training, development/validation, testing)? If so, please provide a description of these splits, explaining the rationale behind them.
This framework is only for evaluation.

Is the dataset self-contained, or does it link to or otherwise rely on external resources (e.g., websites, tweets, other datasets)? If it links to or relies on external resources, a) are there guarantees that they will exist, and remain constant, over time;

b) are there official archival versions of the complete dataset (i.e., including the external resources as they existed at the time the dataset was created); c) are there any restrictions (e.g., licenses, fees) associated with any of the external resources that might apply to a future user? Please provide descriptions of all external resources and any restrictions associated with them, as well as links or other access points, as appropriate.

The DecordFace framework utilizes existing datasets for designing degraded variants. (a) These datasets may become unavailable if decommissioned by the original authors. (b) Yes, at the time of creation, there were official archival versions of the datasets. (c) The datasets can be obtained either through the official websites or by request through a licensing form from the authors. The original LFW dataset can be obtained from http://vis-www.cs.umass.edu/lfw/. The original CALFW dataset can be obtained from http://whdeng.cn/CALFW/index.html?reload=true.
The original CPLFW dataset can be obtained from http://www.whdeng.cn/CPLFW/index.html?reload=true.
Our lab obtained IJB-C through the official NIST form, but its distribution has since been discontinued. An archival version of the IJB suite is available at https://github.com/deepinsight/insightface/tree/master/recognition/_evaluation_/ijb.

Does the dataset contain data that might be considered confidential (e.g., data that is protected by legal privilege or by doctor-patient confidentiality, data that includes the content of individuals non-public communications)? If so, please provide a description.

The datasets comprise of face images. However, they are publicly available.

Does the dataset contain data that, if viewed directly, might be offensive, insulting, threatening, or might otherwise cause anxiety? If so, please describe why.

No.

Preprocessing/cleaning/labeling

Was any preprocessing/cleaning/labeling of the data done (e.g., discretization or bucketing, tokenization, part-of-speech tagging, SIFT feature extraction, removal of instances, processing of missing values)? If so, please provide a description. If not, you may skip the remainder of the questions in this section.

The images from the original datasets was preprocessed to be corrupted for 16 image corruptions at five severities each.

Was the "raw" data saved in addition to the preprocessed/cleaned/labeled data (e.g., to support unanticipated future uses)? If so, please provide a link or other access point to the "raw" data.

Yes. The 'raw' data is available as part of the existing datasets.

Is the software used to preprocess/clean/label the instances available? If so, please provide a link or other access point.

The script utilized for corrupting the images from the original datasets is available at https://github.com/surbhim18/DecordFace.

Evaluation Metrics

Explain any novel evaluation metrics, if any.
The framework provides two novel evaluation metrics. The $mVCE$ (mean Verification Corruption Error) and the $mCEI$ (mean Corruption Embedding Invariance) score. The $mVCE$ utilizes the standard $TPR@FPR$ metric in face verification. On the other hand, the $mCEI$ metric utilizes high-level features extracted from the models for evaluation.

Is the code for implementation of metrics available? If yes, please provide the link. The code for evaluation is available on GitHub at https://github.com/surbhim18/DecordFace.

Uses

Has the dataset been used for any tasks already? If so, please provide a description.
The data created for DecordFace has been utilized for the evaluation and study of different face recognition models to analyze their robustness.

What (other) tasks could the dataset be used for?
The framework and data can be utilized for designing face recognition models that are robust in the presence of image corruptions.

Is there anything about the composition of the dataset or the way it was collected and preprocessed/cleaned/labeled that might impact future uses? For example, is there anything that a future user might need to know to avoid uses that could result in unfair treatment of individuals or groups (e.g., stereotyping, quality of service issues) or other undesirable harms (e.g., financial harms, legal risks) If so, please provide a description. Is there anything a future user could do to mitigate these undesirable harms?
A fairness analysis across different gender and ethnicity subgroups is recommended to understand model behavior and develop fair algorithms.

Are there tasks for which the dataset should not be used? If so, please provide a description.
The data and verification pairs used in evaluation should not be directly utilized for training models. The framework for corruptions and data created through that may be utilized for designing algorithms.

Distribution

Will the dataset be distributed to third parties outside of the entity (e.g., company, institution, organization) on behalf of which the dataset was created? If so, please provide a description.
The tools to create the dataset are publicly available on GitHub at https://github.com/surbhim18/DecordFace.

How will the dataset will be distributed (e.g., tarball on website, API, GitHub) Does the dataset have a digital object identifier (DOI)?
The data framework and tools are available on GitHub at https://github.com/surbhim18/DecordFace.

Will the dataset be distributed under a copyright or other intellectual property (IP) license, and/or under applicable terms of use (ToU)? If so, please describe this license and/or

ToU, and provide a link or other access point to, or otherwise reproduce, any relevant licensing terms or ToU, as well as any fees associated with these restrictions.
The data is available freely under an academic license.

Have any third parties imposed IP-based or other restrictions on the data associated with the instances? If so, please describe these restrictions, and provide a link or other access point to, or otherwise reproduce, any relevant licensing terms, as well as any fees associated with these restrictions.
No.

Do any export controls or other regulatory restrictions apply to the dataset or to individual instances? If so, please describe these restrictions, and provide a link or other access point to, or otherwise reproduce, any supporting documentation.
No.

Maintenance

Who will be supporting/hosting/maintaining the dataset?
The tools to create the data are hosted on GitHub at https://github.com/surbhim18/DecordFace.

How can the owner/curator/manager of the dataset be contacted (e.g., email address)?
The dataset manager can be contacted mittal.5@iitj.ac.in.

Is there an erratum? If so, please provide a link or other access point.
No.

Will the dataset be updated (e.g., to correct labeling errors, add new instances, delete instances)? If so, please describe how often, by whom, and how updates will be communicated to users (e.g., mailing list, GitHub)?
No.

If the dataset relates to people, are there applicable limits on the retention of the data associated with the instances (e.g., were individuals in question told that their data would be retained for a fixed period of time and then deleted)? If so, please describe these limits and explain how they will be enforced.
Since the data is based on existing datasets, the retention of data is dependent on the original dataset curators. We retain copies of the corrupted sets, which may be released upon request.

Will older versions of the dataset continue to be supported/hosted/maintained? If so, please describe how. If not, please describe how its obsolescence will be communicated to users.
The older versions of the dataset will continue to be maintained and should be accessible through the GitHub URL: https://github.com/surbhim18/DecordFace.

If others want to extend/augment/build on/contribute to the dataset, is there a mechanism for them to do so? If so, please provide a description. Will these contributions be validated/verified? If so, please describe how. If not, why not? Is there a process for communicating/distributing these contributions to other users? If so, please provide a description.
Other researchers are welcome to utilize the framework to extend the data.

These contributions will be added to the GitHub repository through the Readme. These researchers can contact the data manager at mittal.5@iitj.ac.in.