Open Images Extended - MIAP

Link to dataset
Link to paper

This dataset was created for fairness research and fairness evaluations in person detection. This dataset contains 100,000 images sampled from Open Images V6 with additional annotations added. Annotations include the image coordinates of bounding boxes for each visible person. Each box is annotated with attributes for perceived gender presentation and age range presentation. It can be used in conjunction with Open Images V6.

Publishers

PUBLISHER(S)	**INDUSTRY TYPE**	**PUBLISHER CONTACT**
Google LLC	Corporate - Tech	Open Images Extended
	AUTHOR CONTACT	**AUTHOR & AFFILIATIONS**
	open-images-extended+miap@google.com	● Candice Schumann, Google, 2021 ● Susanna Ricco, Google, 2021 ● Utsav Prabhu, Google, 2021 ● Vittorio Ferrari, Google, 2021 ● Caroline Pantofaru, Google, 2021

Licenses & Access

LICENSE TYPE(S)	**LICENSE PERMISSIONS (*CC-BY-4.0*)**	
CC-BY-4.0	You are free to share and adapt. Attribution required. You cannot apply any additional restrictions.	
ACCESS	**ACCESS COST**	**ACCESS PREREQUISITES**
Open Access	Open Access	Read the note on perceived gender presentation and perceived age presentation and acceptable use.
ACCESS DOCUMENTATION Available	**DIRECT LINKS TO DATASET** Dataset website **LINKS TO DATASET DOCUMENTATION** Research paper published at AIES 2021.	**ACCESS DETAILS** Dataset includes bounding box annotations only. Images are accessed separately. Users should cite:

```
@inproceedings{miap_aies,
  title = {A Step Toward More Inclusive
People Annotations for Fairness},
  author = {Candice Schumann and Susanna
Ricco and Utsav Prabhu and Vittorio Ferrari
and Caroline Rebecca Pantofaru},
  booktitle = {Proceedings of the AAAI/ACM
Conference on AI, Ethics, and Society
(AIES)},
  year = {2021}
}
```

Dataset Snapshot

DATA TYPE Static Data	**NATURE OF CONTENT** Bounding boxes of people with perceived gender presentation attributes (*predominantly feminine, predominantly masculine, unknown*) and age range presentation attributes (*young, middle, older, unknown*).	**KNOWN CORRELATIONS** • Gender presentation numbers are skewed towards predominantly perceived as masculine and unknown. • Age range presentation range numbers are skewed towards *middle*. • Perceived gender presentation is *unknown* for all bounding boxes with age range attribute annotated *young*.

PRIMARY DATA FORMAT(S) Annotations for image data	**BREAKDOWN - BY INSTANCE** Total Instances 100,000 Training 70,000 Validation 7,410 Testing 22,590 Total boxes 454,331 Human Annotated Labels — All labels manually annotated	**NOTES** All annotated images included at least one person bounding box in Open Images v6. 30,474 of the 100k images contain a MIAP-annotated bounding box with no corresponding annotation in Open Images. Almost 100,000 of the bounding boxes have no corresponding annotation in Open Images. Attributes were annotated for all boxes.

PRIMARY DATA SUBJECT(S) Person boxes	**EXAMPLE OF ACTUAL DATA POINT**	**HOW TO INTERPRET A DATAPOINT** Each datapoint includes a bounding box denoted by XMin, XMax, YMin, and YMax in normalized image coordinates. The next five attributes (IsOccluded through IsInsideOf) follow the definitions from Open Images V6. The last two values for each datapoint correspond to the gender presentation attribute and an age range presentation attribute, respectively. Each annotation is linked to an Open Images key pointing to an image that can be found in CVDF.

EXAMPLE OF ACTUAL DATA POINT

Field	Value
ImageID	164b0e6d1fcf8e61
LabelName	/m/01g317
Confidence	1
XMin	0.897112
XMax	0.987365
YMin	0.615523
YMax	0.895307
IsOccluded	0
IsTruncated	1
IsGroupOf	0
IsDepictionOf	0
IsInsideOf	0
GenderPresentation	Predominantly Masculine
AgePresentation	Middle

Motivations & Use

DATASET PURPOSE(S)
Training

Testing

Validation

Research

KEY DOMAIN APPLICATION(S)
Machine Learning, Object Recognition, Machine Learning Fairness

PRIMARY MOTIVATION(S)
Provide more complete ground-truth for bounding boxes around people. Provide a standard fairness evaluation set for the broader fairness community.

PROBLEM SPACE
This dataset was created for fairness research and fairness evaluation with respect to person detection.

INTENDED USE CASE(S)
Dataset is intended for:
ML Model Evaluation for the following
- Person detection
- Fairness evaluation

ML Model Training for the following:
- Person detection
- Object detection

Extended Use

SAFETY OF USE
Conditional use (some unsafe applications)

SAFE USE TYPE
Person detection
Fairness evaluations
Fairness research

INTENDED USE CASES
Person detection: Without specifying gender or age presentations
Fairness evaluations: Over gender and age presentations
Fairness research: Without building gender presentation or age classifiers

UNSAFE USE TYPE
Gender or age classification

This dataset should **not** be used to create gender or age classifiers.

CONJUNCTIONAL USE
Safe to use with other datasets

KNOWN SAFE DATASETS
These data can be combined with Open Images V6.

KNOWN CONJUNCTIONAL PRACTICES
Analyzing bounding box annotations not annotated under the Open Images V6 procedure.

Maintenance, Versions and Status

STATUS
Actively Maintained

CURRENT VERSION 1.0

FIRST RELEASE 05/2021

STATUS DESCRIPTION
Updates will be pushed to the dataset website.

FIRST EDITION
Annotations completed late 2019 - early 2020.

Data Collection Methods

DATA COLLECTION
Derived

DATA SOURCE
Open Images V6

DATA COLLECTED
100k randomly sampled images containing at least one person box (man, woman, boy, girl, or person).

SELECTION CRITERIA
100k randomly sampled images containing at least one person box (labeled as man, woman, boy, girl, person).

EXCLUDED DATA
No excluded data

Labelling Methods

LABELING METHOD(S) Human labels	LABEL TYPES AND SOURCES Bounding boxes: Human annotators Perceived age range and gender presentation: Human annotators	LABEL DESCRIPTION Bounding boxes were created around *all* people in an image and perceived age ranges as well as perceived gender presentation were labeled.
LABEL TYPE: Bounding boxes	**LABEL TASK(S)** ● Create the bounding box around all people ● Label object attributes **LABELLER DESCRIPTION(S)** ● Compensated workers based out of India	**LABEL DESCRIPTION** A rectangular bounding box around each person in an image. **LABELING TASK OR PROCEDURE** Annotators were asked to place boxes around all people in an image. If there were 5 or more people grouped together a single box was used and a *group of* attribute was associated with that box. Annotators were asked if the person inside of the box was *truncated, occluded,* or *inside of* something. They were also asked if the person inside of the box was a *depiction of* a person (such as a painting or figurine).
LABEL TYPE: Perceived gender presentation and age range	**LABEL TASK(S)** ● Label the perceived gender presentation ● Label the perceived age range **LABELLER DESCRIPTION(S)** ● Compensated workers based out of India	**LABEL DESCRIPTION** Perceived gender presentation: *predominantly feminine, predominantly masculine, unknown* Perceived age range: *young, middle, older, unknown* Note that gender presentation for people marked as *young* is always set to *unknown*. **LABELING TASK OR PROCEDURE** Annotators were asked to select either *predominantly feminine*, *predominantly masculine*, or *unknown* to describe the human-perceived gender presentation of an individual based on the visual cues in the image. Annotators were also asked to select either *young*, *middle*, *older*, or *unknown* to describe the perceived age range of an individual based on their appearance in the image. Annotators were instructed to prefer the older of two categories in situations where there was enough information to form an impression but were unsure of a boundary case. For example, someone who appears old enough to possibly belong to *middle* should be assigned that attribute label.

Fairness Indicators

SENSITIVE HUMAN ATTRIBUTES:
Age, Gender

SUBGROUP INTENTIONALITY
Perceived age ranges: intended
Perceived gender presentation: intended

INTENTIONALITY OF SUBGROUP
This data collection effort was primarily introduced to help fairness research and evaluations.

SUBGROUP TYPE:
Perceived Age Ranges

REPRESENTED DISTRIBUTION

Young	6.3%
Middle	51.4%
Older	2.0%
Unknown	40.2%

SOURCE OF SUBGROUP
Annotators were given examples of different age ranges and asked to label each person in an image with an age range. If annotators were unsure of the age range, they were asked to select *Unknown*.

EXPECTATIONS, RISK, AND CAVEATS
This label does *not* represent the actual age of the individuals in the images. It rather represents the *perceived* age range of the individuals as determined by the human annotators.

TRADEOFFS
Although these labels do not represent the true age ranges of individuals in images, they are still valuable because they allow researchers to assess the performance of models across age ranges, which can ultimately lead to less biased models that work well for all users.

SUBGROUP TYPE:
Perceived Gender Presentation

REPRESENTED POPULATION

Predominantly Feminine	22.2%
Predominantly Masculine	38.3%
Unknown	39.5%

EXPECTATIONS, RISK, AND CAVEATS
Note that gender is *not* binary, and an individual's gender identity may not match their gender presentation. It is *not* possible to label gender identity from images. Additionally, norms around gender expression vary across cultures and have changed over time. No single aspect of a person's appearance "defines" their gender expression. For example, a person may still present as predominantly masculine while wearing jewelry. Another may present as predominantly feminine while having short hair. The intention of these labels is to capture gender presentation as assessed by a third party based on visual cues alone, rather than an individual's self-identified gender.

SOURCE OF SUBGROUP
Annotators were given diverse examples of different gender presentations and asked to label each person in an image with a perceived gender presentation. If annotators were unsure about a gender presentation they were asked to select *Unknown*.

TRADEOFFS
These labels are still valuable because they allow researchers to assess the performance of models across gender presentation, which can ultimately lead to less biased models that work well for all users. While these annotations will sometimes be misaligned with each individual's self-identified gender, in aggregate the annotations are useful to give us a simplified overall sense of how model performance may differ for people who present gender differently.
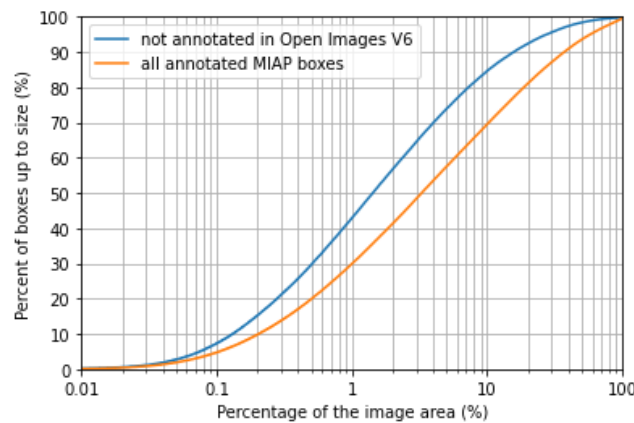
Bounding Box Sizes

SIZE DISTRIBUTIONS

Box size distributions

BOX SIZES BY ATTRIBUTES
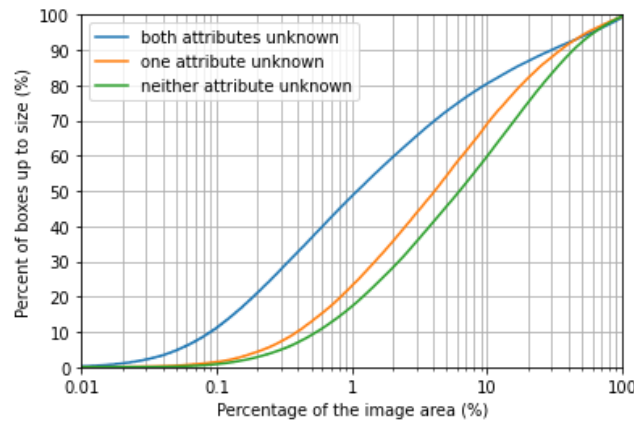
BOX SIZES FOR PREVIOUSLY MISSING ANNOTATIONS

EXAMPLES OF BOX SIZES

REASONS FOR DIFFERENCES

Many boxes are annotated with either *unknown* perceived gender presentation or perceived age range. These bounding boxes are typically smaller, corresponding to people that are either farther away or occluded in some way.

- 48.5% of boxes with both attributes annotated as *unknown* are smaller than 1% of the total image area.
- Just 17.2% of boxes with both perceived age range and perceived gender presentation annotated as a value other than *unknown* are smaller than 1% of the total image area.
- 40.1% of boxes without an *unknown* annotation are larger than 10% of the image area.

Almost 100,000 of the bounding boxes in MIAP do not have a corresponding bounding box in the Open Images V6 annotations. These boxes tend to be smaller than the average across all boxes. However:

- 57% are larger than 1% of the image.
- 26% are larger than 5% of the image.
- 15% are larger than 10% of the image.

The white boxes shown to the left correspond to 1%, 5%, 10%, and 25% of the black square, respectively.

Methods

ML APPLICATION(S)

Object Detection

Fairness

SUMMARY - OBJECT DETECTION

A person object detector can be trained using the Object Detection API in Tensorflow.

KNOWN CAVEATS - METHOD 1

If this dataset is used in conjunction with the original Open Images dataset, negative examples of people should only be pulled from images with an explicit negative *person* image level label.

The dataset does not contain any examples not annotated as containing at least one person by the original Open Images annotation procedure.

SUMMARY - FAIRNESS EVALUATION

Fairness evaluations can be run over the splits of gender presentation and age presentation.

KNOWN CAVEATS - METHOD 2

There still exists a gender presentation skew towards unknown and predominantly masculine, as well as an age presentation range skew towards middle.