

EPIC Kitchens 100 Dataset

EPIC-Kitchens-100 is the largest dataset in first-person (egocentric) vision; itself an extension of the EPIC-Kitchens-55 dataset (formally known as EPIC-Kitchens-2018).

Authors

Dima Damen (1) Hazel Doughty (1) Giovanni Maria Farinella (2) Antonino Furnari (2) Evangelos Kazakos (1) Jian Ma (1) Davide Moltisanti (1) Jonathan Munro (1) Toby Perrett (1) Will Price (1) Michael Wray (1)

- (1 University of Bristol)
- (2 University of Catania)

Contact: uob-epic-kitchens2018@bristol.ac.uk

Citing

When using the dataset, kindly reference:

TBA

Dataset Details

The EPIC-Kitchens-100 dataset is an extension of the EPIC-Kitchens-55 dataset. Videos are distinguished as follows:

- PXX_YY.MP4 videos originate from EPIC-Kitchens-55.
- PXX_1YY.MP4 videos originate from the extension collected for EPIC-Kitchens-100 (thus represent new videos).

The dataset currently has 6 active benchmarks:

- Action Recognition
- Weakly Supervised Action Recognition
- Action Detection
- Action Anticipation
- Unsupervised Domain Adaptation
- Action Retrieval

We provide csv files for the train/val/test sets of each benchmark detailed below for ease of use.

Ground truth is provided for action segments as action/verb/noun labels along with the start and end times of the segment.

Quick Start

Here you can download the annotation files for all of the challenges. For more information on each challenge, please see the paper [here](#). Download scripts are provided for the videos, RGB Frames and Flow frames [here](#).

Action Recognition Challenge

1. Download the videos/RGB/Flow frames.
2. Download the Action Recognition train/val/test files.

Weakly Supervised Action Recognition Challenge

1. Download the videos/RGB/Flow frames.
2. This challenge uses the Action Recognition files, download the train/val/test files.

Action Detection Challenge

1. Download the videos/RGB/Flow frames.
2. This challenge uses the Action Recognition files, download the train/val/test files.

Action Anticipation Challenge

1. Download the videos/RGB/Flow frames.
2. This challenge uses the Action Recognition files, download the train/val/test files.

Unsupervised Domain Adaptation Challenge

1. Download the videos/RGB/Flow frames.
2. Download the Unsupervised Domain Adaptation source train/target train/target test files.

Action Retrieval Challenge

1. Download the videos/RGB/Flow frames.
2. Download the Action Retrieval train/test files.
3. Extract video features (for both the train and test set) using an off-the-shelf model trained on **EPIC-Kitchens-55** (example model).
4. Extract word2vec features for the captions from both the train and test set (example model).
5. Enjoy the EPIC-Kitchens-100 dataset in your favourite video retrieval model, see the paper for details on the models we used for this baseline.

Important Files

For ease of use, download scripts are provided to download the videos and RGB/Flow frames. (see file downloads for more details). We direct the reader to RDSF for the full release of videos and RGB/Flow frames. We provide html and pdf alternatives to this README which are auto-generated.

- README.md (this file)
- README.pdf
- README.html
- license.txt
- EPIC_100_train.csv (info) (Pickle)
- EPIC_100_validation.csv (info) (Pickle)
- EPIC_100_test_timestamps.csv (info) (Pickle)
- EPIC_100_noun_classes.csv (info)
- EPIC_100_verb_classes.csv (info)

Additional Files

- UDA_annotations/EPIC_100_uda_source_train (info) (Pickle)
- UDA_annotations/EPIC_100_uda_target_train_timestamps (info) (Pickle)
- UDA_annotations/EPIC_100_uda_target_test_timestamps (info) (Pickle)
- retrieval_annotations/EPIC_100_retrieval_train (info) (Pickle)
- retrieval_annotations/EPIC_100_retrieval_test (info) (Pickle)

File Structure

EPIC_100_train.csv

This CSV file contains the action annotations for the training set and contains 15 columns:

| Column Name | Type | Example | Description |
|---------------------|----------------------------|--------------|---|
| uid | int | 0 | Unique ID for the segment as an int. |
| narration_id | string | P01_01_0 | Unique ID for the segment as a string with participant ID and video ID. |
| participant_id | int | P01 | ID of the participant (unique per participant). |
| video_id | string | P01_01 | ID of the video where the segment originated from (unique per video). |
| narration_timestamp | string | 00:00:01.089 | Timestamp of when the original narration was recorded in HH:mm:ss.SSS. |
| start_timestamp | string | 00:00:00.14 | Start time in HH:mm:ss.SS of the action segment. |
| stop_timestamp | string | 00:00:03.37 | End time in HH:mm:ss.SS of the action segment. |
| start_frame | int | 8 | Start frame of the action. |
| stop_frame | int | 202 | End frame of the action. |
| narration | string | open door | Transcribed description of the English narration provided by the participant. |
| verb | string | open | Parsed verb from the narration. |
| verb_class | int | 3 | Numeric ID of the verb's class. |
| noun | string | door | First parsed noun from the narration. |
| noun_class | int | 3 | Numeric ID of the first noun's class. |
| all_nouns | list of string (1 or more) | [door] | List of all parsed nouns within the narration. |
| all_noun_classes | list of int (1 or more) | [3] | Numeric ID of all of the parsed noun's classes. |

[Back to Important Files](#)

EPIC_100_validation.csv

This CSV file contains the action annotations for the validation set and contains 15 columns:

| Column Name | Type | Example | Description |
|---------------------|--------|--------------|---|
| uid | int | 4972 | Unique ID for the segment as an int. |
| narration_id | string | P01_01_11 | Unique ID for the segment as a string with participant ID and video ID. |
| participant_id | int | P01 | ID of the participant (unique per participant). |
| video_id | string | P01_11 | ID of the video where the segment originated from (unique per video). |
| narration_timestamp | string | 00:00:00.560 | Timestamp of when the original narration was recorded in HH:mm:ss.SSS. |

| Column Name | Type | Example | Description |
|------------------|----------------------------|-------------|---|
| start_timestamp | string | 00:00:00.00 | Start time in HH:mm:ss.SS of the action segment. |
| stop_timestamp | string | 00:00:01.89 | End time in HH:mm:ss.SS of the action segment. |
| start_frame | int | 1 | Start frame of the action. |
| stop_frame | int | 113 | End frame of the action. |
| narration | string | take plate | Transcribed description of the English narration provided by the participant. |
| verb | string | take | Parsed verb from the narration. |
| verb_class | int | 0 | Numeric ID of the verb's class. |
| noun | string | plate | First parsed noun from the narration. |
| noun_class | int | 2 | Numeric ID of the first noun's class. |
| all_nouns | list of string (1 or more) | [plate] | List of all parsed nouns within the narration. |
| all_noun_classes | list of int (1 or more) | [2] | Numeric ID of all of the parsed noun's classes. |

[Back to Important Files](#)

EPIC_100_test_timestamps.csv

This CSV file contains the action annotations for the testing set and contains 9 columns:

| Column Name | Type | Example | Description |
|---------------------|--------|--------------|---|
| uid | int | 1924 | Unique ID for the segment as an int. |
| narration_id | string | P01_101_0 | Unique ID for the segment as a string with participant ID and video ID. |
| participant_id | int | P01 | ID of the participant (unique per participant). |
| video_id | string | P01_101 | ID of the video where the segment originated from (unique per video). |
| narration_timestamp | string | 00:00:02.851 | Timestamp of when the original narration was recorded in HH:mm:ss.SSS. |
| start_timestamp | string | 00:00:02.86 | Start time in HH:mm:ss.SSS of the action segment. |
| stop_timestamp | string | 00:00:03.87 | End time in HH:mm:ss.SSS of the action segment. |
| start_frame | int | 143 | Start frame of the action. |
| stop_frame | int | 193 | End frame of the action. |

[Back to Important Files](#)

EPIC_100_noun_classes.csv

This CSV file contains information on the 300 noun classes and contains 4 columns.

| Column Name | Type | Example | Description |
|-------------|----------------------------|------------------------|---|
| id | int | 222 | Unique ID for the noun class. |
| key | string | label | Key used for the noun class (all keys are a member of their own class). |
| instances | list of string (1 or more) | "['label', 'sticker']" | All nouns within the class, including the key. |

| Column Name | Type | Example | Description |
|-----------------|--------|------------------|---|
| category | string | materials | Name of the higher-level noun category that this noun class belongs to. |

[Back to Important Files](#)

EPIC_100_verb_classes.csv

This CSV file contains information on the 97 verb classes and contains 4 columns.

| Column Name | Type | Example | Description |
|------------------|----------------------------|---------------------|---|
| id | int | 79 | Unique ID for the verb class. |
| key | string | let-go | Key used for the verb class (all keys are a member of their own class). |
| instances | list of string (1 or more) | "['let', 'let-go']" | All verbs within the class, including the key. |
| category | string | leave | Name of the higher-level verb category that this verb class belongs to. |

[Back to Important Files](#)

EPIC_100_uda_source_train.csv

This CSV file contains the action annotations for the **source training set** used for **Unsupervised Domain Adaptation** and contains 15 columns:

| Column Name | Type | Example | Description |
|----------------------------|----------------------------|------------------|---|
| uid | int | 0 | Unique ID for the segment as an int. |
| narration_id | string | P01_01_0 | Unique ID for the segment as a string with participant ID and video ID. |
| participant_id | int | P01 | ID of the participant (unique per participant). |
| video_id | string | P01_01 | ID of the video where the segment originated from (unique per video). |
| narration_timestamp | string | 00:00:01.089 | Timestamp of when the original narration was recorded in HH:mm:ss.SSS. |
| start_timestamp | string | 00:00:00.14 | Start time in HH:mm:ss.SS of the action segment. |
| stop_timestamp | string | 00:00:03.37 | End time in HH:mm:ss.SS of the action segment. |
| start_frame | int | 8 | Start frame of the action. |
| stop_frame | int | 202 | End frame of the action. |
| narration | string | open door | Transcribed description of the English narration provided by the participant. |
| verb | string | open | Parsed verb from the narration. |
| verb_class | int | 3 | Numeric ID of the verb's class. |
| noun | string | door | First parsed noun from the narration. |
| noun_class | int | 3 | Numeric ID of the first noun's class. |
| all_nouns | list of string (1 or more) | [door] | List of all parsed nouns within the narration. |
| all_noun_classes | list of int (1 or more) | [3] | Numeric ID of all of the parsed noun's classes. |

Note that this file contains only videos from EPIC-Kitchens-55 which is used as the source domain.

See here for more details on the unsupervised domain adaptation challenge.

[Back to Important Files](#)

EPIC_100_uda_target_train_timestamps.csv

This CSV file contains the action annotations for the **target training set** used for **Unsupervised Domain Adaptation** and contains 9 columns:

| Column Name | Type | Example | Description |
|---------------------|--------|--------------|---|
| uid | int | 2571 | Unique ID for the segment as an int. |
| narration_id | string | P01_102_0 | Unique ID for the segment as a string with participant ID and video ID. |
| participant_id | int | P01 | ID of the participant (unique per participant). |
| video_id | string | P01_102 | ID of the video where the segment originated from (unique per video). |
| narration_timestamp | string | 00:00:01.100 | Timestamp of when the original narration was recorded in HH:mm:ss.SSS. |
| start_timestamp | string | 00:00:00.54 | Start time in HH:mm:ss.SS of the action segment. |
| stop_timestamp | string | 00:00:02.23 | End time in HH:mm:ss.SS of the action segment. |
| start_frame | int | 27 | Start frame of the action. |
| stop_frame | int | 111 | End frame of the action. |

Note that this file contains only videos from EPIC-Kitchens-100 which is used as the target domain.

See here for more details on the unsupervised domain adaptation challenge.

[Back to Important Files](#)

EPIC_100_uda_target_test_timestamps.csv

This CSV file contains the action annotations for the **target testing set** used for **Unsupervised Domain Adaptation** and contains 9 columns:

| Column Name | Type | Example | Description |
|---------------------|--------|--------------|---|
| uid | int | 1924 | Unique ID for the segment as an int. |
| narration_id | string | P01_101_0 | Unique ID for the segment as a string with participant ID and video ID. |
| participant_id | int | P01 | ID of the participant (unique per participant). |
| video_id | string | P01_101 | ID of the video where the segment originated from (unique per video). |
| narration_timestamp | string | 00:00:02.851 | Timestamp of when the original narration was recorded in HH:mm:ss.SSS. |
| start_timestamp | string | 00:00:02.86 | Start time in HH:mm:ss.SS of the action segment. |
| stop_timestamp | string | 00:00:03.87 | End time in HH:mm:ss.SS of the action segment. |
| start_frame | int | 143 | Start frame of the action. |
| stop_frame | int | 193 | End frame of the action. |

Note that this file contains only videos from EPIC-Kitchens-100 which is used as the target domain.

See here for more details on the unsupervised domain adaptation challenge.

[Back to Important Files](#)

EPIC_100_retrieval_train.csv

This CSV file contains the action annotations for the **action retrieval** training set and contains 15 columns:

| Column Name | Type | Example | Description |
|---------------------|----------------------------|--------------|---|
| uid | int | 0 | Unique ID for the segment as an int. |
| narration_id | string | P01_01_0 | Unique ID for the segment as a string with participant ID and video ID. |
| participant_id | int | P01 | ID of the participant (unique per participant). |
| video_id | string | P01_01 | ID of the video where the segment originated from (unique per video). |
| narration_timestamp | string | 00:00:01.089 | Timestamp of when the original narration was recorded in HH:mm:ss.SSS. |
| start_timestamp | string | 00:00:00.14 | Start time in HH:mm:ss.SS of the action segment. |
| stop_timestamp | string | 00:00:03.37 | End time in HH:mm:ss.SS of the action segment. |
| start_frame | int | 8 | Start frame of the action. |
| stop_frame | int | 202 | End frame of the action. |
| narration | string | open door | Transcribed description of the English narration provided by the participant. |
| verb | string | open | Parsed verb from the narration. |
| verb_class | int | 3 | Numeric ID of the verb's class. |
| noun | string | door | First parsed noun from the narration. |
| noun_class | int | 3 | Numeric ID of the first noun's class. |
| all_nouns | list of string (1 or more) | [door] | List of all parsed nouns within the narration. |
| all_noun_classes | list of int (1 or more) | [3] | Numeric ID of all of the parsed noun's classes. |

[Back to Important Files](#)

EPIC_100_retrieval_test.csv

This CSV file contains the action annotations for the **action retrieval** testing set and contains 15 columns:

| Column Name | Type | Example | Description |
|---------------------|--------|--------------|---|
| uid | int | 0 | Unique ID for the segment as an int. |
| narration_id | string | P01_11_0 | Unique ID for the segment as a string with participant ID and video ID. |
| participant_id | int | P01 | ID of the participant (unique per participant). |
| video_id | string | P01_11 | ID of the video where the segment originated from (unique per video). |
| narration_timestamp | string | 00:00:00.560 | Timestamp of when the original narration was recorded in HH:mm:ss.SSS. |
| start_timestamp | string | 00:00:00.00 | Start time in HH:mm:ss.SS of the action segment. |
| stop_timestamp | string | 00:00:01.89 | End time in HH:mm:ss.SS of the action segment. |
| start_frame | int | 8 | Start frame of the action. |
| stop_frame | int | 113 | End frame of the action. |
| narration | string | take plate | Transcribed description of the English narration provided by the participant. |

| Column Name | Type | Example | Description |
|-------------------------------|----------------------------|----------------------|---|
| <code>verb</code> | string | <code>take</code> | Parsed verb from the narration. |
| <code>verb_class</code> | int | 0 | Numeric ID of the verb's class. |
| <code>noun</code> | string | <code>plate</code> | First parsed noun from the narration. |
| <code>noun_class</code> | int | 2 | Numeric ID of the first noun's class. |
| <code>all_nouns</code> | list of string (1 or more) | <code>[plate]</code> | List of all parsed nouns within the narration. |
| <code>all_noun_classes</code> | list of int (1 or more) | <code>[2]</code> | Numeric ID of all of the parsed noun's classes. |

[Back to Important Files](#)

Additional Information

File Downloads

Due to the size of the dataset we provide scripts for downloading parts of the dataset:

- videos (GB)
- frames (GB)
 - rgb-frames (GB)
 - flow-frames (GB)

Note: These scripts will work for Linux and Mac. For Windows users a bash installation should work.

These scripts replicate the folder structure of the dataset release on RDSF, found [here](#).

If you wish to download part of the dataset, instructions can be found [here](#).

Differences to EPIC-Kitchen-100

Whilst videos from EPIC-Kitchens-55 are used within EPIC-Kitchens-100 some of the annotations have been modified to improve the quality of the annotations. Additionally, with EPIC-Kitchens-100, the verb/noun classes have been updated to cover the annotations from the new videos. Because of this, the annotations from EPIC-Kitchens-55 cannot be used for EPIC-Kitchens-100.

Pickle Files

We also provide pickle files for all of the main train/val/test csvs for ease of use. These files require python 3.5+ and pandas 1.0.0+ to read. The pickle files are automatically tagged with the commit hash and version for version control purposes which can be found as follows in python:

```
>>> import pandas as pd
>>> train = pd.read_pickle('EPIC_100_train.pkl')
>>> train._metadata
{'commit_hash': 'ce7a0fb', 'version_number': '0.1'}
```

showing that this version of the `EPIC_100_train.pkl` came from commit hash `ce7a0fb` and version number `0.1`.

License

All files in this dataset are copyright by us and published under the Creative Commons Attribution-NonCommercial 4.0 International License, found [here](#). This means that you must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use. You may not use the material for commercial purposes.

Changelog

Please see the release history for the changelog.