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-kitchens@bristol.ac.uk

Citing

When using the dataset, kindly reference:

```
@ARTICLE{Damen2020RESCALING,
   title={Rescaling Egocentric Vision},
   author={Damen, Dima and Doughty, Hazel and Farinella, Giovanni Maria  and and Furnari, Antonino
          and Ma, Jian and Kazakos, Evangelos and Moltisanti, Davide and Munro, Jonathan
          and Perrett, Toby and Price, Will and Wray, Michael},
          journal   = {CoRR},
          volume    = {abs/2006.13256},
          year      = {2020},
          ee        = {http://arxiv.org/abs/2006.13256},
}
```

Index

- Dataset Details
- Quick start Guides
- Important Files
- File Structure
- Additional Information
- License

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, see Important Files for more information.

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

back to top

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. A download script is provided for the videos, RGB Frames and Flow frames here.

Action Recognition Challenge

1. Download the videos/RGB/Flow frames here with the following command:

```
python epic_downloader.py --videos --rgb-frames --flow-frames
```

2. Download the Action Recognition train/val/test files.
3. Enjoy the EPIC-KITCHENS-100 dataset in your favourite action recognition model, see the paper for details on the models we used for this baseline. Models trained on EPIC-KITCHENS-55 can be found here as a starting point.

Weakly Supervised Action Recognition Challenge

1. Download the videos/RGB/Flow frames here with the following command:

```
python epic_downloader.py --videos --rgb-frames --flow-frames
```

2. This challenge uses the Action Recognition files, download the train/val/test files.
3. The weakly supervised challenge uses the narration timestamp, not the the start/end times of the action. Therefore a simple baseline would be to modify an action recognition model to use the surrounding 5s worth of frames. See the paper for details on the models we used for this baseline.

Action Detection Challenge

1. Download the videos/RGB/Flow frames here with the following command:

```
python epic_downloader.py --videos --rgb-frames --flow-frames
```

2. This challenge uses the Action Recognition files, download the train/val/test files.
3. Train an action proposal network on the EPIC-KITCHENS-100 train set, for example this model. This model predicts action-agnostic segments which still need to be classified.
4. Use your favourite action recognition model to classify the proposals (example models).

Action Anticipation Challenge

1. Download the videos/RGB/Flow frames here with the following command:

```
python epic_downloader.py --videos --rgb-frames --flow-frames
```

2. This challenge uses the Action Recognition files, download the train/val/test files.
3. A simple baseline for this task is to train an action recognition model (example models here) on the 5 seconds that precede an action with a 1 second gap. For example, an action that starts at 20.00s in a video would see frames between 14.00s and 19.00s.

Unsupervised Domain Adaptation Challenge

The unsupervised domain adaptation challenge tests how models can cope with similar data collected 2 years later on the task of action recognition.

1. Download the videos/RGB/Flow frames here with the following command:

```
python epic_downloader.py --videos --rgb-frames --flow-frames --domain-adaptation
```

2. Download the Unsupervised Domain Adaptation source train/target train/target test files.
3. Extract video features (for all three splits) using an off-the-shelf model trained on **EPIC-KITCHENS-55** (example model).
4. A simple baseline is using a domain discriminator (prediciting whether a video came from the source, EPIC-KITCHENS-55, or the target, EPIC-KITCHENS-100) to align the two domains. See the paper for details on the models we used for this baseline.

Action Retrieval Challenge

1. Download the videos/RGB/Flow frames here with the following command:

```
python epic_downloader.py --videos --rgb-frames --flow-frames --action-retrieval
```

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 models).
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.

back to top

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.csv` (info) (Pickle)
- `UDA_annotations/EPIC_100_uda_source_test_timestamps.csv` (info) (Pickle)
- `UDA_annotations/EPIC_100_uda_target_train_timestamps.csv` (info) (Pickle)
- `UDA_annotations/EPIC_100_uda_target_test_timestamps.csv` (info) (Pickle)
- `retrieval_annotations/EPIC_100_retrieval_train.csv` (info) (Pickle)
- `retrieval_annotations/EPIC_100_retrieval_test.csv` (info) (Pickle)
- `EPIC_100_train_missing_timestamps_narrations.csv` (info)
- `EPIC_100_validation_missing_timestamps_narrations.csv` (info)

back to top

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
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:

4

Column Name	Type	Example	Description
`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`.
`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
`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.

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

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.

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
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.

Column Name	Type	Example	Description
`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_source_test_timestamps.csv

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

Column Name	Type	Example	Description
`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	`1`	Start frame of the action.
`stop_frame`	int	`113`	End frame of the action.

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
`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
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
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.

Column Name	Type	Example	Description
`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
`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.
`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_train_missing_timestamps_narrations.csv

This CSV file contains the narration IDs of all EPIC-KITCHENS-55 videos in the training set which do not have a narration timestamp (see here for more details). This file has one column:

Column Name	Type	Example	Description
`narration_id`	string	`P01_09_660`	Unique ID for the segment as a string with participant ID and video ID.

Back to Important Files

EPIC_100_validation_missing_timestamps_narrations.csv

This CSV file contains the narration IDs of all EPIC-KITCHENS-55 videos in the validation set which do not have a narration timestamp (see here for more details). This file has one column:

Column Name	Type	Example	Description
narration_id	string	P02_12_293	Unique ID for the segment as a string with participant ID and video ID.

Back to Important Files

back to top

Additional Information

File Downloads

Due to the size of the dataset we provide a script for downloading parts of the dataset which can be found here. If you wish to download the extension only (i.e. you have already downloaded EPIC-KITCHENS-55) the following command can be run:

```
python epic_downloader.py --extension-only
```

If you wish to download the whole dataset, the following command can be run:

```
python epic_downloader.py
```

See the README for more information.

Differences to EPIC-Kitchen-55

Updated Annotations

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.

Missing Narration Timestamps

Due to the differences in the annotation pipeline between EPIC-KITCHENS-100 and EPIC-KITCHENS-55, it was impossible to assign the narration timestamp to every action. Because of this, there are actions within EPIC_100_train.csv and EPIC_100_validation.csv which do not have timestamp narrations and are thus marked with NaN within the dataframes.

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 in python using the following commands:

```
>>> import pandas as pd
>>> train = pd.read_pickle('EPIC_100_train.pkl')
>>> train._metadata
{'commit_hash': 'ce7a0fb', 'version_number': '1.0.0'
```

showing that this version of the `EPIC_100_train.pkl` came from commit hash ce7a0fb and version number 1.0.0.

back to top

License

All files in this dataset are copyright by us and published under the Creative Commons Attribution-NonCommerial 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.

back to top

Changelog

Please see the release history for the changelog.

back to top