TFDS now supports the Croissant 🥐 format (https://mlcommons.org/croissant)! Read the documentation (https://www.tensorflow.org/datasets/format_specific_dataset_builders#croissantbuilder) to know more.

plant_village

- **Description**:

The PlantVillage dataset consists of 54303 healthy and unhealthy leaf images divided into 38 categories by species and disease.

Note: The original dataset is not available from the original source (plantvillage.org), therefore we get the unaugmented dataset from a paper that used that dataset and republished it. Moreover, we dropped images with Background_without_leaves label, because these were not present in the original dataset.

Original paper URL: https://arxiv.org/abs/1511.08060 (https://arxiv.org/abs/1511.08060) Dataset URL: https://data.mendeley.com/datasets/tywbtsjrjv/1 (https://data.mendeley.com/datasets/tywbtsjrjv/1)

- **Additional Documentation**:
  Explore on Papers With Code ↗ (https://paperswithcode.com/dataset/plantvillage)

- **Homepage**: https://arxiv.org/abs/1511.08060 (https://arxiv.org/abs/1511.08060)

- **Source code**: `tfds.datasets.plant_village.Builder` (https://github.com/tensorflow/datasets/tree/master/tensorflow_datasets/datasets/plant_village/plant_village_dataset_builder.py)

- **Versions**:

  - `1.0.2` (default): No release notes.

- **Download size**: `827.82 MiB`

- **Dataset size**: `815.37 MiB`

- **Auto-cached** (documentation (https://www.tensorflow.org/datasets/performances#auto-caching)): No

- **Splits**:

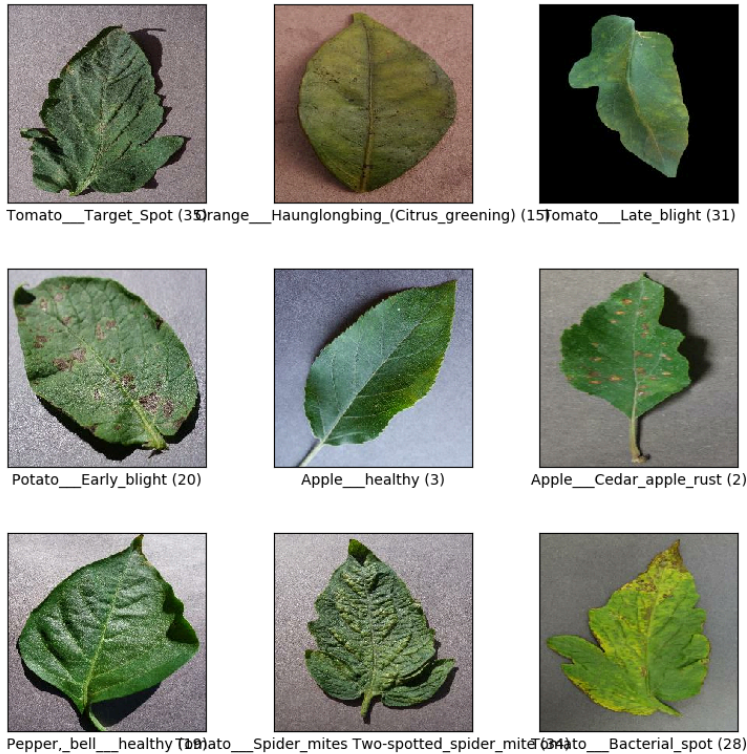| Split	Examples
`'train'`	54,303

- **Feature structure**:

```
FeaturesDict({
    'image': Image(shape=(None, None, 3), dtype=uint8),
    'image/filename': Text(shape=(), dtype=string),
    'label': ClassLabel(shape=(), dtype=int64, num_classes=38),
})
```

- **Feature documentation**:

Feature	Class	Shape	Dtype	Description
	FeaturesDict			
image	Image	(None, None, 3)	uint8	
image/filename	Text		string	
label	ClassLabel		int64	

- **Supervised keys** (See as_supervised doc
  (https://www.tensorflow.org/datasets/api_docs/python/tfds/load#args)): (`'image'`, `'label'`)

- **Figure** (tfds.show_examples
  (https://www.tensorflow.org/datasets/api_docs/python/tfds/visualization/show_examples)):

Tomato___Target_Spot (35)   Orange___Haunglongbing_(Citrus_greening) (15)   Tomato___Late_blight (31)

Potato___Early_blight (20)   Apple___healthy (3)   Apple___Cedar_apple_rust (2)

Pepper,_bell___healthy (19)   Tomato___Spider_mites Two-spotted_spider_mite (34)   Tomato___Bacterial_spot (28)

- **Examples** (tfds.as_dataframe
  (https://www.tensorflow.org/datasets/api_docs/python/tfds/as_dataframe)):

  Display examples...

- **Citation**:

```
@article{DBLP:journals/corr/HughesS15,
  author    = {David P. Hughes and
               Marcel Salath{\'{e} } },
  title     = {An open access repository of images on plant health to enable the
               development of mobile disease diagnostics through machine
               learning and crowdsourcing},
  journal   = {CoRR},
  volume    = {abs/1511.08060},
  year      = {2015},
  url       = {http://arxiv.org/abs/1511.08060},
  archivePrefix = {arXiv},
```

```
  eprint    = {1511.08060},
  timestamp = {Mon, 13 Aug 2018 16:48:21 +0200},
  biburl    = {https://dblp.org/rec/bib/journals/corr/HughesS15},
  bibsource = {dblp computer science bibliography, https://dblp.org}
}
```