



TorchIO

🔍 Search

Getting started  
Data structures  
Patch-based pipelines  
Transforms  
**Medical image datasets**  
Additional interfaces  
Examples gallery  
GitHub repository ↗  
Paper ↗



## Medical image datasets

TorchIO offers tools to easily download publicly available datasets from different institutions and modalities.

The interface is similar to `torchvision.datasets`.

If you use any of them, please visit the corresponding website (linked in each description) and make sure you comply with any data usage agreement and you acknowledge the corresponding authors' publications.

If you would like to add a dataset here, please open a discussion on the GitHub repository:

💬 Discuss

ⓘ

### CONTENTS

IXI  
IXITiny  
EPISURG  
EPISURG  
RSNAMICCAI  
RSNAMICCAI  
MNI  
ICBM2009CNonlinearSymmetric  
Colin27  
Pediatric  
Sheep  
BITE3  
Visible Human Project  
VisibleMale  
VisibleFemale  
ITK-SNAP  
BrainTumor  
T1T2  
AorticValve  
3D Slicer  
slicer  
FFG  
MedMNIST

### IXI

The [Information eXtraction from Images \(IXI\)](#) dataset contains "nearly 600 MR images from normal, healthy subjects", including "T1, T2 and PD-weighted images, MRA images and Diffusion-weighted images (15 directions)".

📝 Note

This data is made available under the Creative Commons CC BY-SA 3.0 license. If you use it please acknowledge the source of the IXI data, e.g. [the IXI website](#).

IXI

```
class torchio.datasets.ixi.IXI(root: Union[str, os.PathLike], transform:  
    Optional[torchio.transforms.Transform] = None, download: bool = False,  
    modalities: Sequence[str] = ('T1', 'T2'), **kwargs) [source]
```

Bases: `torchio.data.dataset.SubjectsDataset`

Full IXI dataset.

PARAMETERS

- **root** – Root directory to which the dataset will be downloaded.
- **transform** – An instance of `Transform`.
- **download** – If set to `True`, will download the data into `root`.
- **modalities** – List of modalities to be downloaded. They must be in `('T1', 'T2', 'PD', 'MRA', 'DTI')`.

⚠ Warning

The size of this dataset is multiple GB. If you set `download` to `True`, it will take some time to be downloaded if it is not already present.

Example:

```
>>> import torchio as tio  
>>> transforms = [  
...     tio.ToCanonical(), # to RAS  
...     tio.Resample((1, 1, 1)), # to 1 mm iso  
... ]  
>>> ixii_dataset = tio.datasets.IXI(  
...     'path/to/ixi_root/',  
...     modalities=('T1', 'T2'),  
...     transform=tio.Compose(transforms),  
...     download=True,  
... )  
>>> print('Number of subjects in dataset:', len(ixii_dataset)) # 577  
>>> sample_subject = ixii_dataset[0]  
>>> print('Keys in subject:', tuple(sample_subject.keys())) # ('T1', 'T2')  
>>> print('Shape of T1 data:', sample_subject['T1'].shape) # [1, 180, 268, 268]  
>>> print('Shape of T2 data:', sample_subject['T2'].shape) # [1, 241, 257, 188]
```

### IXITiny

```
class torchio.datasets.ixi.IXITiny(root: Union[str, os.PathLike], transform:  
    Optional[torchio.transforms.Transform] = None, download: bool = False,  
    **kwargs) [source]
```

Bases: `torchio.data.dataset.SubjectsDataset`

This is the dataset used in the main [notebook](#). It is a tiny version of IXI, containing 566  $T_1$ -weighted brain MR images and their corresponding brain segmentations, all with size  $83 \times 44 \times 55$ .

It can be used as a medical image MNIST.

PARAMETERS

- **root** – Root directory to which the dataset will be downloaded.
- **transform** – An instance of `Transform`.
- **download** – If set to `True`, will download the data into `root`.

## EPISURG

### EPISURG

```
class torchio.datasets.episurg.EPISURG(root: Union[str, os.PathLike], transform:  
    Optional[torchio.transforms.Transform] = None, download: bool = False,  
    **kwargs) [source]
```

Bases: [torchio.data.dataset.SubjectsDataset](#)

EPISURG is a clinical dataset of  $T_1$ -weighted MRI from 430 epileptic patients who underwent resective brain surgery at the National Hospital of Neurology and Neurosurgery (Queen Square, London, United Kingdom) between 1990 and 2018.

The dataset comprises 430 postoperative MRI. The corresponding preoperative MRI is present for 268 subjects.

Three human raters segmented the resection cavity on partially overlapping subsets of EPISURG.

If you use this dataset for your research, you agree with the *Data use agreement* presented at the EPISURG entry on the [UCL Research Data Repository](#) and you must cite the corresponding publications.

#### PARAMETERS

- **root** – Root directory to which the dataset will be downloaded.
- **transform** – An instance of [Transform](#).
- **download** – If set to `True`, will download the data into `root`.

#### ⚠ Warning

The size of this dataset is multiple GB. If you set `download` to `True`, it will take some time to be downloaded if it is not already present.

[get\\_labeled\(\)](#) → [torchio.data.dataset.SubjectsDataset](#)

[source]

Get dataset from subjects with manual annotations.

[get\\_paired\(\)](#) → [torchio.data.dataset.SubjectsDataset](#)

[source]

Get dataset from subjects with pre- and post-op MRI.

[get\\_unlabeled\(\)](#) → [torchio.data.dataset.SubjectsDataset](#)

[source]

Get dataset from subjects without manual annotations.

## RSNAMICCAI

#### RSNAMICCAI

```
class torchio.datasets.rsnamiccai.RSNAMICCAI(root_dir: Union[str, os.PathLike],  
                                              train: bool = True, ignore_empty: bool = True, modalities: Sequence[str] =  
                                              ('T1w', 'T1wCE', 'T2w', 'FLAIR'), **kwargs)
```

[source]

Bases: [torchio.data.dataset.SubjectsDataset](#)

RSNA-MICCAI Brain Tumor Radiogenomic Classification challenge dataset.

This is a helper class for the dataset used in the [RSNA-MICCAI Brain Tumor Radiogenomic Classification challenge](#) hosted on [kaggle](#). The dataset must be downloaded before instantiating this class (as opposed to, e.g., [torchio.datasets.IXI](#)).

This [kaggle kernel](#) includes a usage example including preprocessing of all the scans.

If you reference or use the dataset in any form, include the following citation:

U.Baid, et al., "The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification", arXiv:2107.02314, 2021.

#### PARAMETERS

- **root\_dir** – Directory containing the dataset (`train` directory, `test` directory, etc.).
- **train** – If `True`, the `train` set will be used. Otherwise the `test` set will be used.
- **ignore\_empty** – If `True`, the three subjects flagged as "presenting issues" (empty images) by the challenge organizers will be ignored. The subject IDs are `00109`, `00123` and `00709`.

#### EXAMPLE

```
>>> import torchio as tio  
>>> from subprocess import call  
>>> call('kaggle competitions download -c rsna-miccai-brain-tumor-radiogenomic-classification')  
>>> root_dir = 'rsna-miccai-brain-tumor-radiogenomic-classification'  
>>> train_set = tio.datasets.RSNAMICCAI(root_dir, train=True)  
>>> test_set = tio.datasets.RSNAMICCAI(root_dir, train=False)  
>>> len(train_set), len(test_set)  
(582, 87)
```

## MNI

#### ICBM2009CNonlinearSymmetric

```
class torchio.datasets.mni.ICBM2009CNonlinearSymmetric(load_4d_tissues: bool = True)
```

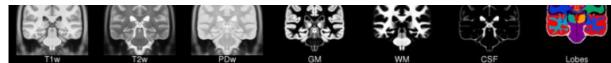
[source]

Bases: [torchio.datasets.mni.mni.SubjectMNI](#)

ICBM template.

More information can be found in the [website](#).





#### PARAMETERS

`load_4d_tissues` – If `True`, the tissue probability maps will be loaded together into a 4D image. Otherwise, they will be loaded into independent images.

#### EXAMPLE

```
>>> import torchio as tio
>>> icbm = tio.datasets.ICBM2009CNonlinearSymmetric()
>>> icbm
ICBM2009CNonlinearSymmetric(Keys: ('t1', 'eyes', 'face', 'brain', 't2', 'pd', 'tissues'))
>>> icbm = tio.datasets.ICBM2009CNonlinearSymmetric(load_4d_tissues=False)
>>> icbm
ICBM2009CNonlinearSymmetric(Keys: ('t1', 'eyes', 'face', 'brain', 't2', 'pd', 'gm', 'wm',
```

#### Colin27

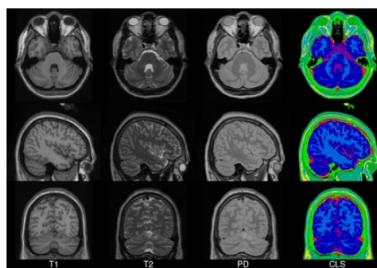
`class torchio.datasets.mni.Colin27(version=1998)`

[\[source\]](#)

Bases: `torchio.datasets.mni.mni.SubjectMNI`

Colin27 MNI template.

More information can be found in the website of the [1998](#) and [2008](#) versions.



#### PARAMETERS

`version` – Template year. It can be [1998](#) or [2008](#).

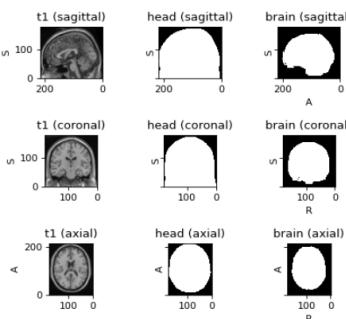
#### ⚠ Warning

The resolution of the [2008](#) version is quite high. The subject instance will contain four images of size  $362 \times 434 \times 362$ , therefore applying a transform to it might take longer than expected.

#### EXAMPLE

```
>>> import torchio as tio
>>> colin_1998 = tio.datasets.Colin27(version=1998)
>>> colin_1998
Colin27(Keys: ('t1', 'head', 'brain'); images: 3)
>>> colin_1998.load()
>>> colin_1998.t1
ScalarImage(shape: (1, 181, 217, 181); spacing: (1.00, 1.00, 1.00); orientation: RAS+; me
>>>
>>> colin_2008 = tio.datasets.Colin27(version=2008)
>>> colin_2008
Colin27(Keys: ('t1', 't2', 'pd', 'cls'); images: 4)
>>> colin_2008.load()
>>> colin_2008.t1
ScalarImage(shape: (1, 362, 434, 362); spacing: (0.50, 0.50, 0.50); orientation: RAS+; me
```

[\(Source code, png, hires.png, pdf\)](#)



#### Pediatric

`class torchio.datasets.mni.Pediatric(years, symmetric=False)`

[\[source\]](#)

Bases: `torchio.datasets.mni.mni.SubjectMNI`

MNI pediatric atlases.

See the [MNI website](#) for more information.

Pediatric MNI template

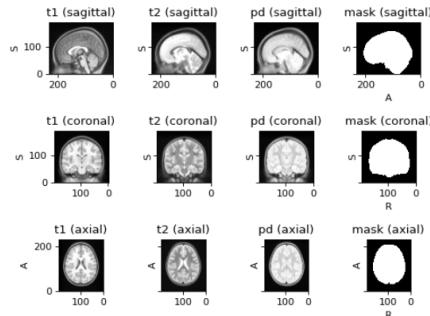
#### PARAMETERS

- `years` – Tuple of 2 ages. Possible values are:  $(4.5, 18.5)$ ,  $(4.5, 8.5)$ ,  $(7, 11)$ ,  $(7.5,$

13.5), (10, 14) and (13, 18.5).

- **symmetric** – If `True`, the left-right symmetric templates will be used. Else, the asymmetric (natural) templates will be used.

([Source code](#), [png](#), [hires.png](#), [pdf](#))

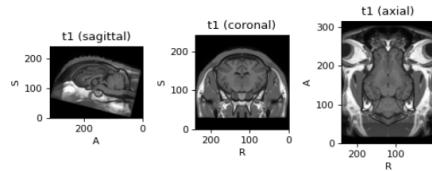


### Sheep

```
class torchio.datasets.mni.Sheep  
Bases: torchio.datasets.mni.mni.SubjectMNI
```

[[source](#)]

([Source code](#), [png](#), [hires.png](#), [pdf](#))



### BITE3

```
class torchio.datasets.bite.BITE3(root: Union[str, os.PathLike], transform:  
    Optional[torchio.transforms.Transform] = None, download: bool = False,  
    **kwargs)  
Bases: torchio.datasets.bite.BITE
```

[[source](#)]

Pre- and post-resection MR images in BITE.

The goal of BITE is to share *in vivo* medical images of patients with brain tumors to facilitate the development and validation of new image processing algorithms.

Please check the [BITE website](#) for more information and acknowledgments instructions.

#### PARAMETERS

- **root** – Root directory to which the dataset will be downloaded.
- **transform** – An instance of `Transform`.
- **download** – If set to `True`, will download the data into `root`.

## Visible Human Project

The [Visible Human Project](#) is an effort to create a detailed data set of cross-sectional photographs of the human body, in order to facilitate anatomy visualization applications. It is used as a tool for the progression of medical findings, in which these findings link anatomy to its audiences. A male and a female cadaver were cut into thin slices which were then photographed and digitized (from [Wikipedia](#)).

### VisibleMale

```
class torchio.datasets.visible_human.VisibleMale(part: str)  
Bases: torchio.datasets.visible_human.VisibleHuman
```

[[source](#)]

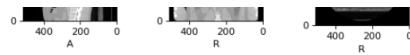
Visible Male CT Datasets.

#### PARAMETERS

- **part** – Can be 'Head', 'Hip', 'Pelvis' or 'Shoulder'.

([Source code](#), [png](#), [hires.png](#), [pdf](#))





### VisibleFemale

```
class torchio.datasets.visible_human.VisibleFemale(part: str) [source]
```

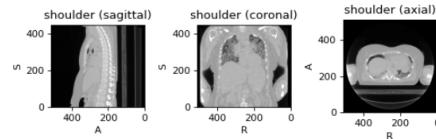
Bases: torchio.datasets.visible\_human.VisibleHuman

Visible Female CT Datasets.

#### PARAMETERS

**part** – Can be 'Ankle', 'Head', 'Hip', 'Knee', 'Pelvis' OR 'Shoulder' .

([Source code](#), [png](#), [hires.png](#), [pdf](#))



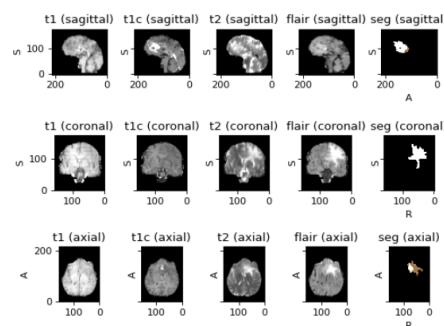
## ITK-SNAP

### BrainTumor

```
class torchio.datasets.itk_snap.BrainTumor [source]
```

Bases: torchio.datasets.itk\_snap.itk\_snap.SubjectITKSNAP

([Source code](#), [png](#), [hires.png](#), [pdf](#))

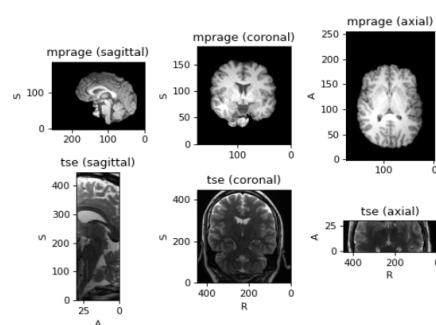


### T1T2

```
class torchio.datasets.itk_snap.T1T2 [source]
```

Bases: torchio.datasets.itk\_snap.itk\_snap.SubjectITKSNAP

([Source code](#), [png](#), [hires.png](#), [pdf](#))

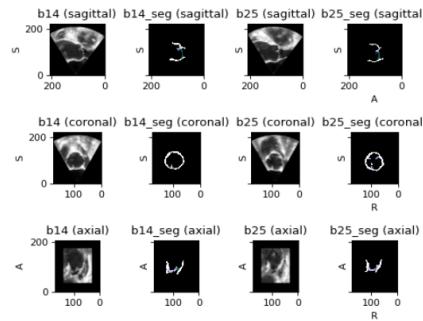


### AorticValve

```
class torchio.datasets.itk_snap.AorticValve [source]
```

Bases: torchio.datasets.itk\_snap.itk\_snap.SubjectITKSNAP

([Source code](#), [png](#), [hires.png](#), [pdf](#))



## 3D Slicer

[Slicer](#)

```
class torchio.datasets.slicer.Slicer(name='MRHead') [source]
```

Bases: [torchio.data.subject.Subject](#)

Sample data provided by 3D Slicer.

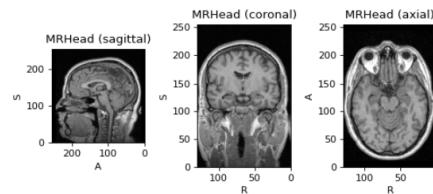
See [the Slicer wiki](#) for more information.

For information about licensing and permissions, check the [Sample Data module](#).

### PARAMETERS

**name** – One of the keys in [torchio.datasets.slicer.URLS\\_DICT](#).

([Source code](#), [png](#), [hires.png](#), [pdf](#))



## FPG

```
class torchio.datasets.fpg.FPG(load_all: bool = False) [source]
```

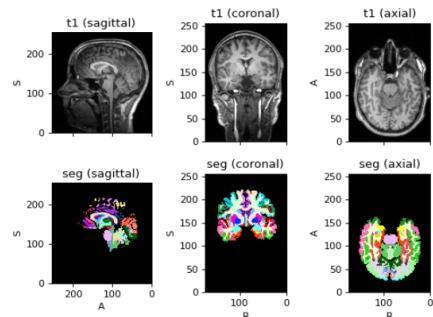
Bases: [torchio.data.subject.Subject](#)

3T  $T_1$ -weighted brain MRI and corresponding parcellation.

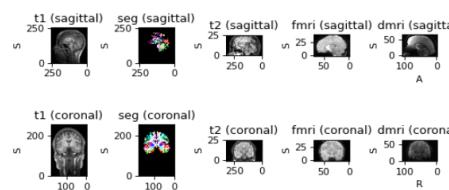
### PARAMETERS

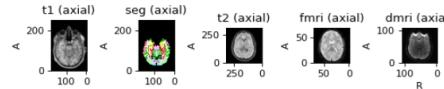
**load\_all** – If `True`, three more images will be loaded: a  $T_2$ -weighted MRI, a diffusion MRI and a functional MRI.

([Source code](#), [png](#), [hires.png](#), [pdf](#))



([Source code](#), [png](#), [hires.png](#), [pdf](#))





## MedMNIST

```
class torchio.datasets.medmnist.OrganMNIST3D(split, **kwargs)
```

[source]

3D MedMNIST v2 datasets.

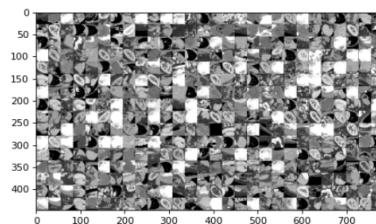
Datasets from [MedMNIST v2: A Large-Scale Lightweight Benchmark for 2D and 3D Biomedical Image Classification](#).

Please check the [MedMNIST website](#) for more information, including the license.

PARAMETERS

**split** – Dataset split. Should be `'train'`, `'val'` or `'test'`.

([Source code](#), [png](#), [hires.png](#), [pdf](#))



```
class torchio.datasets.medmnist.NoduleMNIST3D(split, **kwargs)
```

[source]

3D MedMNIST v2 datasets.

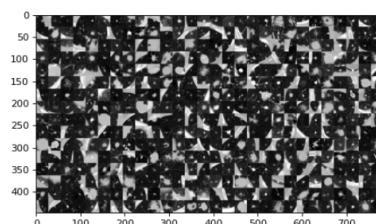
Datasets from [MedMNIST v2: A Large-Scale Lightweight Benchmark for 2D and 3D Biomedical Image Classification](#).

Please check the [MedMNIST website](#) for more information, including the license.

PARAMETERS

**split** – Dataset split. Should be `'train'`, `'val'` or `'test'`.

([Source code](#), [png](#), [hires.png](#), [pdf](#))



```
class torchio.datasets.medmnist.AdrenalMNIST3D(split, **kwargs)
```

[source]

3D MedMNIST v2 datasets.

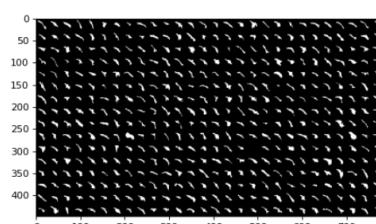
Datasets from [MedMNIST v2: A Large-Scale Lightweight Benchmark for 2D and 3D Biomedical Image Classification](#).

Please check the [MedMNIST website](#) for more information, including the license.

PARAMETERS

**split** – Dataset split. Should be `'train'`, `'val'` or `'test'`.

([Source code](#), [png](#), [hires.png](#), [pdf](#))



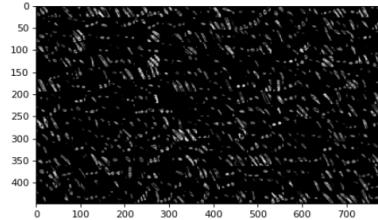
```
class torchio.datasets.medmnist.FractureMNIST3D(split, **kwargs) [source]
3D MedMNIST v2 datasets.

Datasets from MedMNIST v2: A Large-Scale Lightweight Benchmark for 2D and 3D Biomedical
Image Classification.

Please check the MedMNIST website for more information, including the license.

PARAMETERS
split – Dataset split. Should be 'train', 'val' or 'test'.
```

([Source code](#), [png](#), [hires.png](#), [pdf](#))



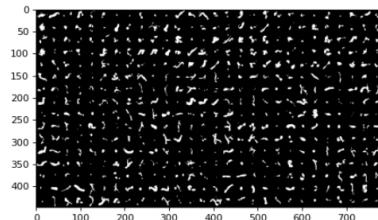
```
class torchio.datasets.medmnist.VesselMNIST3D(split, **kwargs) [source]
3D MedMNIST v2 datasets.

Datasets from MedMNIST v2: A Large-Scale Lightweight Benchmark for 2D and 3D Biomedical
Image Classification.

Please check the MedMNIST website for more information, including the license.

PARAMETERS
split – Dataset split. Should be 'train', 'val' or 'test'.
```

([Source code](#), [png](#), [hires.png](#), [pdf](#))



```
class torchio.datasets.medmnist.SynapseMNIST3D(split, **kwargs) [source]
3D MedMNIST v2 datasets.

Datasets from MedMNIST v2: A Large-Scale Lightweight Benchmark for 2D and 3D Biomedical
Image Classification.

Please check the MedMNIST website for more information, including the license.

PARAMETERS
split – Dataset split. Should be 'train', 'val' or 'test'.
```

([Source code](#), [png](#), [hires.png](#), [pdf](#))

