**Regarding the Data Annotations**

In this dataset of the project, we release the data with their corresponding labels, where we will give you three alternatives of labels, i.e., binary labels of real or fake, labels for specific manipulations, and semantics-oriented labels.

In the fake data within each file split, the name of its file represents a specific manipulation. For example, FakeManipulation-1 represents a form of fake manipulation while FakeManipulation-2 represents another fake manipulation. Regarding the semantics-oriented labels for the fake data, the details are listed below:

|  |  |
| --- | --- |
| Manipulation | Semantic Attribute(s) |
| FakeManipulation-1 | Identity |
| FakeManipulation-2 | Identity |
| FakeManipulation-3 | Expression, Identity |
| FakeManipulation-4 | Expression, Identity |
| FakeManipulation-5 | Expression, Head Pose |

The data in the files started with Real is of the real category.

You may encounter some of the following pictures. This is normal and can be regarded as data augmentations or interference from data propagation in online social networks.

   

Though we have split the data into training and validation sets, you can also reorganize them as you wish during training.