

配布データと応募用ファイルの作成方法について

本コンペティションで配布されるデータと応募用ファイルの作成方法について説明する。

配布データ

配布するデータは以下の通り。

- `README.pdf`
- `train_0.zip~train_5.zip`
- `train_master.csv`
- `test_0.zip~test_2.zip`
- `sample_submit.csv`

README

このファイル、最初に参照するドキュメント。`README.pdf`として与えられる。

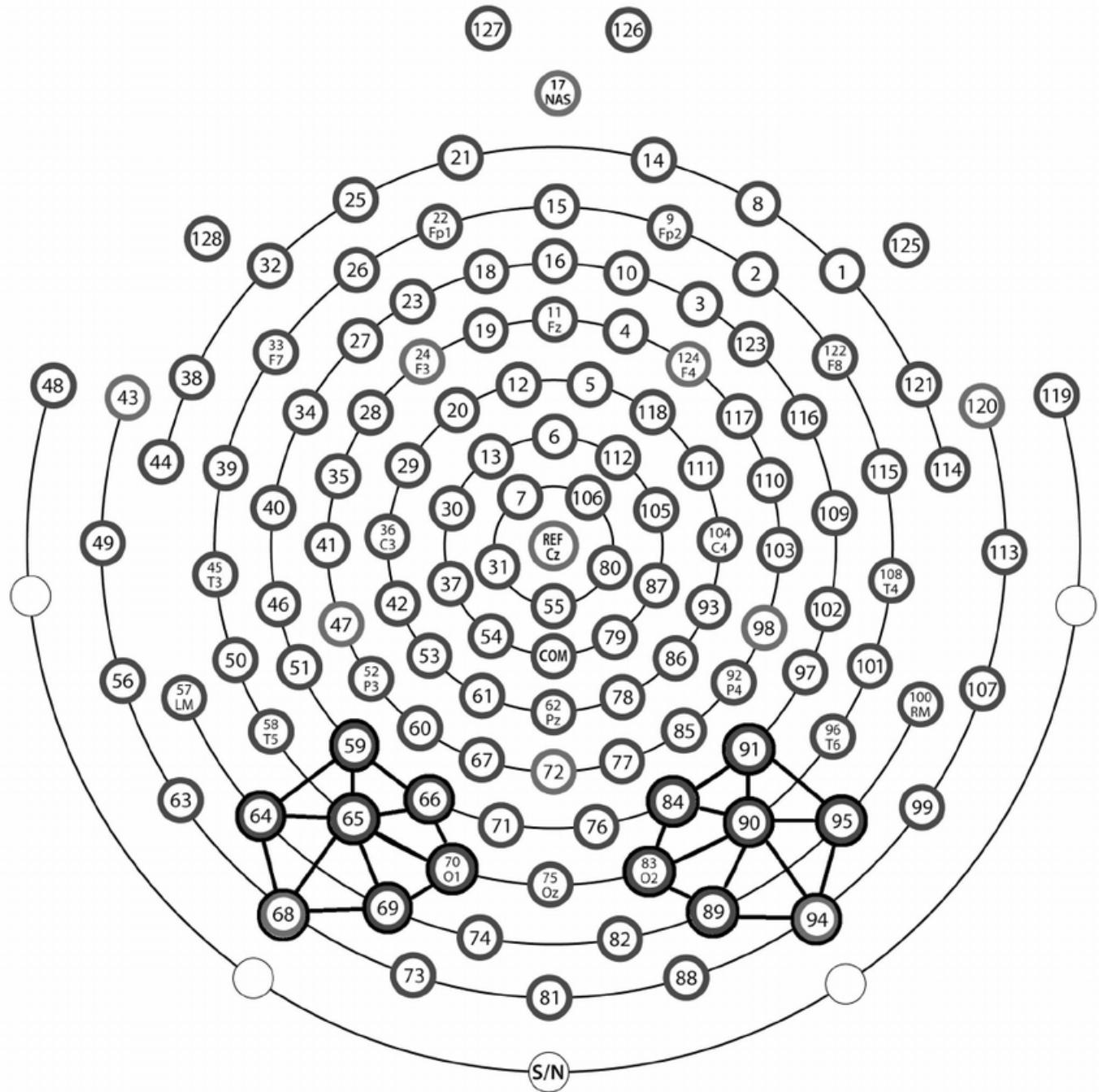
学習用データ

モデルを学習するための系列データ。`train_0.zip~train_5.zip`として配布されている。解凍すると以下のようないいきなりで展開される。

```
train
├── train_000.csv
└── ...
└── train_2247.csv
```

全部で2248個のcsvファイルが与えられる。各csvファイルはヘッダーなしで、行は時刻を表し、列はチャンネルを表す。チャンネル数129、float型のデータで単位は`[μV]`である。サンプリングレートは100`[Hz]`。

各チャンネルイメージは以下の通り。



学習用アノテーションデータ

学習用データに対応するアノテーションデータ. [train_master.csv](#)として配布されている. ヘッダーありで, 定義は以下の通り.

カラム名	説明
id	サンプルID, データのファイル名(拡張子なし)に対応する.
label	対応するデータのラベル. 0(安静), 1(運動)のいずれか.
user	被験者ID.

カラム名	説明
block	実験を時間軸で切った単位(ブロック). 一ブロックの中には、安静状態を5秒、運動をイメージする時間が5秒ある。この10秒1セットを1試行と呼ぶ。これが20回繰り返されて、一つのブロックになっている。

評価用データ

精度評価の対象となる系列データ. [test_0.zip~test_2.zip](#)として配布されている。解凍すると以下のようなディレクトリで展開される。

```
test
└── test_000.csv
    ...
└── test_757.csv
```

全部で758個のcsvファイルが与えられる。各csvファイルのフォーマットは学習用データと同じ。

応募用サンプルファイル

投稿用のサンプルファイル. [sample_submit.csv](#)として配布されている。このフォーマットに従って応募用ファイルを作成する。ヘッダーなしのcsvファイル。

カラム番号	説明
0	サンプルID、評価用データのファイル名(拡張子なし)に対応する。
1	対応するデータのラベル名。0(安静)、1(運動)のいずれか。

応募用ファイルの作成方法

上記の[応募用サンプルファイル](#)のフォーマットに従って作成する。ヘッダーなしのcsvファイルで、すべての評価用ID([test_000~test_757](#))を含めること。

注意事項

応募用ファイルを投稿する前に、正しいフォーマットになっているかを確認すること。特に下記の点に注意。

- ヘッダーなしになっているか
- IDとラベルが正しく対応しているか
- IDに過不足はないか
- ファイルの拡張子が[.csv](#)になっているか