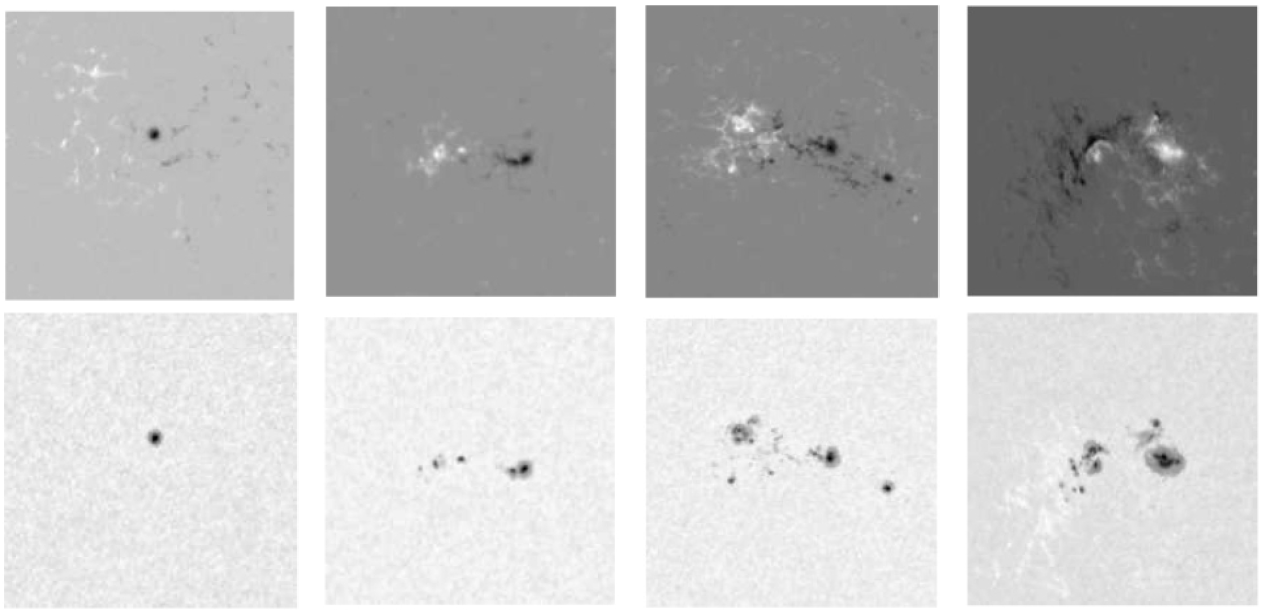
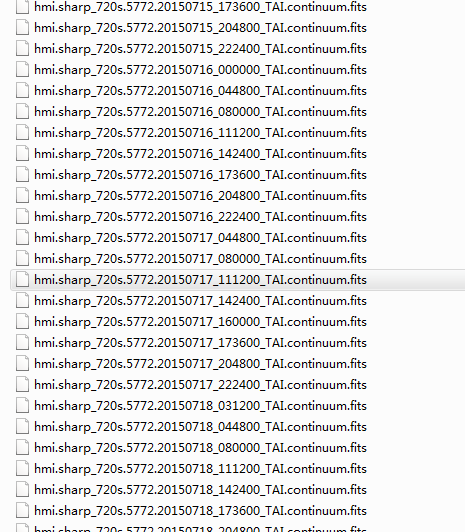
**比赛题目**

太阳活动区中黑子群的分类在太阳风暴预报中一直发挥着重要作用。然而，对该特征的识别还主要是通过人工判断的方式进行提取。针对当前太阳黑子群分类工作中面临的问题，本次比赛利用积累的海量太阳黑子磁场观测图像及其对应的磁类型标签，基于人工智能方法建立对黑子磁类型的自动分类模型。  
  
如上分别为alpha(左1)、beta(左2)和更复杂的beta-x(左3、左4)磁类型的太阳黑子群磁场观测图像（上）和白光观测图像（下）

**数据说明**

1. 观测图像文件在trainset.zip，太阳黑子群磁图或白光图分别对应目录magnetogram\或continuum\，数据格式为fits(用astropy.io.fits包读取)；
2. 观测图像对应三种标签：alpha、beta、betax，每一个磁图或白光图数据文件均对应唯一标签，并存储于相应标签名称的子目录下；
3. 文件夹magnetogram或continuum中的数据是存在一一对应关系的，是同一时间对同一日面区域（即黑子群所在位置）的不同方法的观测图。通过文件名称能够发现这种对应关系。如：  
   在文件夹magnetogram\alpha\中存在黑子群磁图文件：  
   hmi.sharp\_720s.5772.20150717\_111200\_TAI.magnetogram.fits  
   那么，在文件夹continuum\alpha夹中也存在黑子群相应的白光图文件：  
   hmi.sharp\_720s.5772.20150717\_111200\_TAI.continuum.fits  
   文件名中的5772表示观测对象，20150717\_111200表示观测时间。文件夹magnetogram或continuum中的一一对应的数据文件是在相同观测时间对相同观测对象通过不同观测方式得到的观测结果。  
   测试集fits图像数据统一存于test\_input.zip，并隐藏观测对象和观测时间信息，改为样本编号：  
   hmi.sharp\_720s.xxxx.yyyymmdd\_000001\_TAI.continuum.fits
4. 数据文件的时间精度是96分钟，见下图所示的5772号黑子群的观测数据文件，表示每间隔96分钟对统一黑子群对象的观测文件。由于黑子群演化的缓慢性，同一个黑子群在一定时间内常会保持同一种磁类型，如下图中的这些相邻数据对应的标签均为alpha类型，此时黑子群5772的同类标签观测文件相似度过高，不宜将以下样本同时采用为训练样本和验证样本，不建议利用随机方式切分训练集与验证集。  
   
5. 提醒：各类样本的数目是不均衡的，TrainSet.zip中，alpha、beta、betax三类样本的文件数目分别4709、7353、2407。

## 提交说明

上传测试结果的格式要求：

1. 文件名为‘C\_队伍名称.txt’，无需表头；
2. 以两列进行输出，空格作间隔，第1列对应测试数据的六位样本编号，第2列为数据输出的标签类，标签1、2、3，分别表示磁类型alpha, beta, betax。
3. 文件名格式和内容格式不符合格式、测试样本条数不全的测试结果视作无效结果。测试样本共1172条。

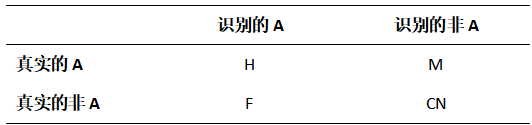
## 评估指标

评分排名考虑的评估指数及优先级排序为：

beta类的F1 score > betax类的F1 score > alpha类的F1 score

当多个队伍的优先指数相同时，比较次优先的评估参数进行排序。

以下为各项评估参数的计算方法：

计算单类样本的评估指数时，将三分类问题处理为二分类问题，以A类为例：  


1. 召回率（Recall）  
   Recall = H/(H+M)
2. 精确率（Precision）  
   Precision = H/(H+F)
3. F1 score  
   2/F1 = 1/Precision + 1/Recall  
   F1 = 2×Precision×Recall /(Precision+Recall)