DESIGN DOCUMENT

Modular Machine Learning
Google Summer of Code

Sumit Kumar Vohra (Student)
Dr. Dimiter Prodanov (Mentor)

Year 2016

Contents

1 Introduction

1.1 What is Active Segmentation?

Active Weka Segmentation is Modular Machine Learning and Classification Toolbox for ImageJ. Active Weka Segmentation is inspired from TWS(Trainable Weka Segmentation) which is a framework for interactive image segmentation. Therefore, Active Weka Segmentation inherits the most of the advantage of the TWS framework. We provides a high level architecture to incorporate any kind of filter, learning and evalaution method. Its main features are:

- **Generic:** Active Segmenation provide a posiblity of including new Filter, learning framework as well adapt or modify the code according to the developer's need.

- **User Friendly:** It provides a user friendly interface to mark region of interest and to visualize the result of classification.The use of Json file format provides a common ground to integrate the framework with other systems and store intermediate results.

- **Open Source:** The source code of Active Segmentation is available under the GNU General Public License (GPL).

2 Workflow

- **Apply Filters :** The starting point for Active Segmentation are filter's. The user can apply wide range of filters on Image. The user can enable or disable filter according to problem. Users can set wide range of parameters related to filter and can store the settings for future reference.

- **Feature extraction :** The Feature extraction opens with binary classes. The user can add, update to delete classes according to their problem but 2 classes are mandatory. The ouput of first phase(Apply Filters) is used to generate vector of each pixel. The user can add their own algoritm of feature feature extraction.

- **Learning :**

- **Evaluation :**
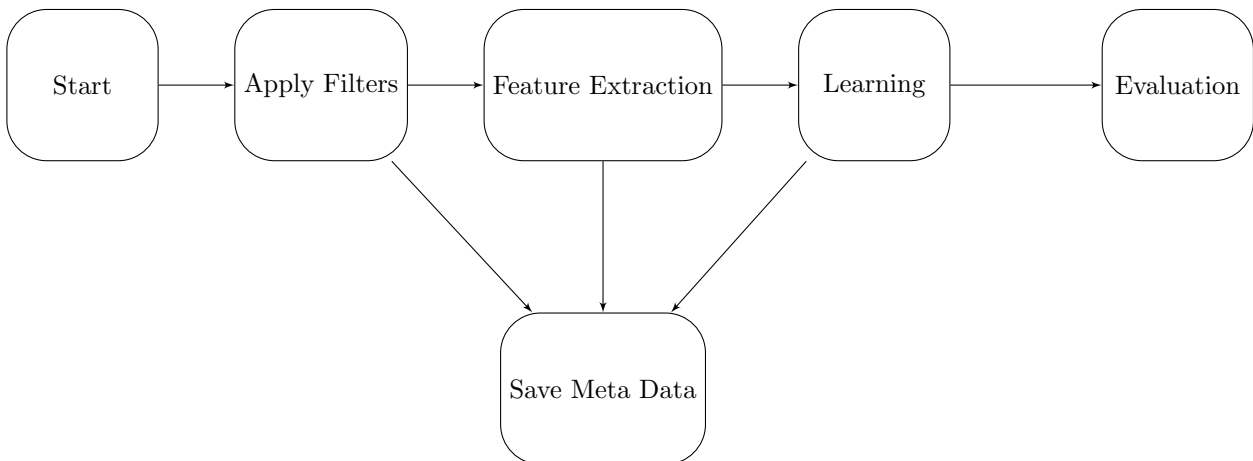
3 Description

3.1 Filter Form

3.2 Feature Form

3.3 Learning Form

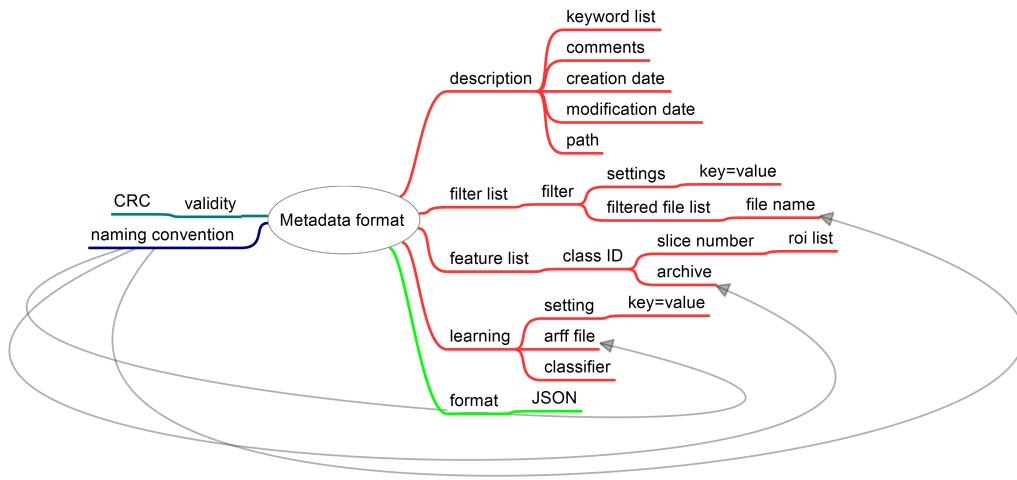## 3.4 Evaluation Form

4 Meta Data

4.1 Mind Map

Figure 1: The Mind Map for Meta Data.

- **Filter List:** consists of meta data information related to all the filters which are currently available in active segmentation. Each element in the list will give information about individual Filter. The key field in Filter is `FilterName` thay will be unique for every filter. Filter will also consist of settings in key- value pair format.

```
1    {
2    "FilterList": [
3    {"Filter":"FilterName1","Setting1":"true","Setting2":"2
        ","Setting3":"false", "FileList": "FilterName1.tif"}
        ,
4    {"Filter":"FilterName2","Setting1":"true","Setting2":"2
        ",
5    "FileList": "FilterName2.tif"}
6    ]
7    }
```

- **Feature List:** consist of list of classId which represents the number of classes for the current segmentation task. We have used the `FeatureList` keyword to access the list of feature. Each classId will consist of `RoiItem`, `Class` and `ZipFile`. Each `RoiItem` will consist of name of region of interest related to each slice. The `ZipFile` will contain all the region of interest related to particular class.

```
1    {
2    "FeatureList": [
3    {"RoiItem": [
```

3

```
 4                          ["0001-0414-0779","0001-0602-0552"],
 5                          ["0002-0425-0407"],
 6                          ["0003-0234-0479"]
 7                          ],
 8                          "zipfile":"ROISET0.zip",
 9                          "class":1
10                          }
11                          }
```

- **Learning:** It consist of meta-data information related to arff file and classifier. To keep the things simple, we are keeping weka format instead of defining our own format for arff and classifier. We are storing only file names of arff and classifier.

```
 1                          {
 2                          "Learning": [
 3                          {"Classifier": "classifier.model",
 4                           "ArffFile": "fileName.arff"
 5                           "Settings":["learningType" :"active"]
 6                          }
 7
 8                          }
```

- **Description:** contains basic information related to Meta data file. It includes date of creation, modified date, keyword List. The keyword List associates classes with the names used by the user.

```
 1                  {
 2                          "Description": [
 3                          {"KeywordList": ["class1" :"BCK","class2" : "
                                Cell"],
 4                           "Comments":"",
 5                           "CreatedDate":"01/07/2014"
 6                           "ModifyDate": "02/07/2014"
 7                           "Path":"C://Program Files//ImageJ//plugins//
                                activeSegmentation//"
 8                  }
 9
10                  }
```