

# Functional Description: Performance Evaluation Tools for Zone Segmentation and Classification (PETS)

## Introduction

This document provides a functional overview of the LAMP PETS software (Performance Evaluation Tools) for the evaluation of zone segmentation and classification algorithms. It is intended to be used with the output of the GEDI ground truthing tool, and for results in the GEDI format. The document describes in simple terms to approach to matching, highlights of GEDI and how to use the software.

This software is version 1.0. If you have any questions or find bugs in the software, please contact anyone in the lab or email [doermann@umd.edu](mailto:doermann@umd.edu)

## Background of Approach

The software runs in three modes: 1) Zone detection requiring only spatial overlap, 2) Zone matching (default) requiring detection of the zone and consistent labeling; and 3) Classification – requiring only consistent labeling, while assuming zoneIDs establish correspondence.

Let  $G$  be the set of ground truth zone  $G=\{g_1, g_2, \dots, g_n\}$ ,  $R$  be the set of result zone  $R=\{r_1, r_2, \dots, r_n\}$ ,  $L(s)$  be the entity of the zone  $s$ ,  $P(s)$  be the foreground pixels of zone  $s$  and  $T(s)$  be a function that counts the elements of zone  $s$ . Using this definition, we can define the matching score metric as

$$MatchScore(i, j) = \frac{T(P(r_i) \cap P(g_j))}{T(P(r_i) \cup P(g_j))}$$

Using this metric, we construct a matching score table (MTS) between result zones and ground truth zones. Once we construct a MTS we define four types of zone matching by type of overlapping such as 'one-to-one', 'one-to-many', 'many-to-one' and 'many-to-many'.

### – One-to-one

This case is simplest case of overlap. Only one result zone

overlaps with one ground truth zone. Figure 1 shows an example of a one-to-one overlap and the matching score table. In this case,  $R1$  is defined as 'MATCHED' if  $L(R1)=L(G1)$  and the matching score is greater than a threshold, otherwise  $R1$  is defined as 'DETECTED' if  $L(R1) \neq L(G1)$ . The default threshold is 80%.

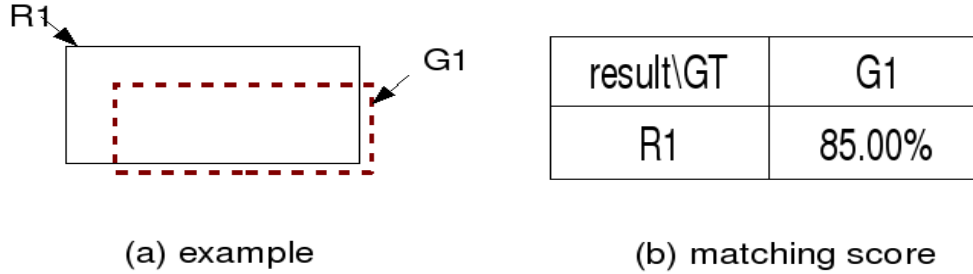


Figure 1. example of one-one overlapping

#### - One-to-many

If one result overlaps multiple ground truth zones, we need to define which ground truth zones are matched to the result zone. Figure 2 shows an example of a one-to-many overlap. In this case, if  $L(R1)=L(G1)=L(G2)$ , we define  $R1$  as 'MATCHED' to  $G1$  and  $G2$  as 'MISSED'. IF  $L(R1)=L(G2) \neq L(G1)$ ,  $R1$  is 'MATCHED' to  $G2$  and  $G1$  is 'MISSED'. IF  $L(R1) \neq L(G1) \neq L(G2)$ ,  $R1$  is 'DETECTED' by  $G1$  and  $G2$  is 'MISSED'.

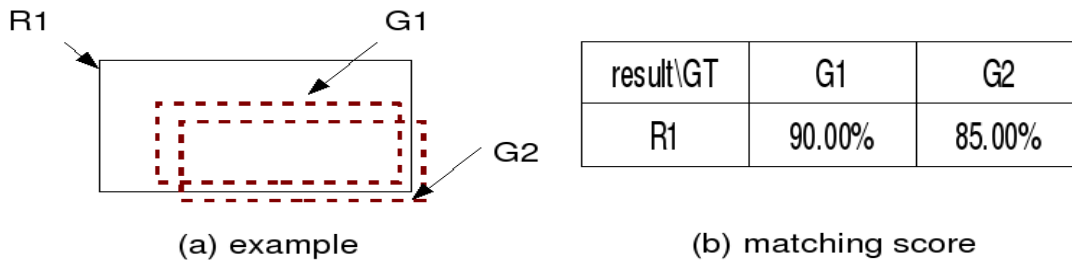
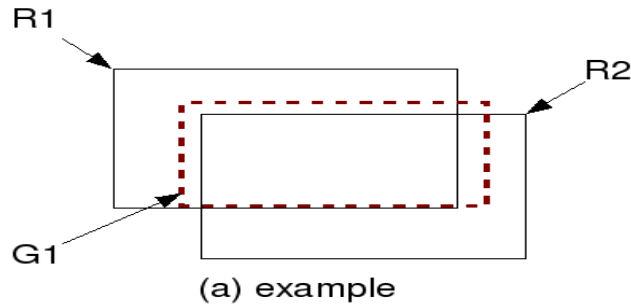


Figure 2. example of one-many overlapping

#### - many-to-one

If multiple result zones overlap one ground truth zone, we need to define which result zones are matched to the ground truth zone. Figure 3 shows an example of a many-to-one overlapping case. In this case, if  $L(R1)=L(R2)=L(G1)$ , we define  $R1$  as 'MATCHED' to  $G1$  and  $R2$  as 'FALSE ALARM'. If  $L(R1) \neq L(R2)=L(G1)$ ,  $R2$  is 'MATCHED' by  $G1$  and  $R1$  is 'FALSE ALARM'. If

$L(R1) \neq L(R2) \neq L(G1)$ ,  $R1$  is 'DETECTED' to  $G1$  and  $R2$  is a 'FALSE ALARM'.



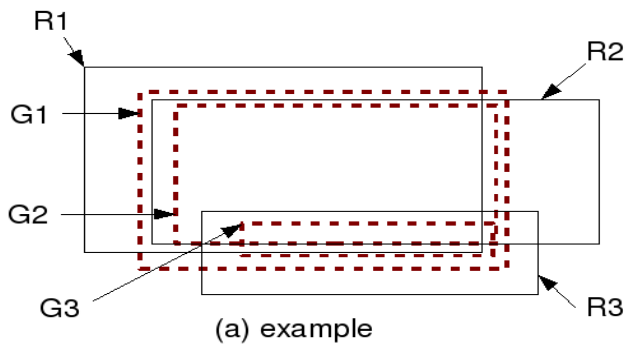
result\GT	G1
R1	95.00%
R2	90.00%

(b) matching score

Figure 3. example of many-one overlapping

#### - many-to-many

This case is the most complicated type of overlap. Multiple result zones overlap multiple ground truth zones. Figure 4 shows an example of a many-to-many overlap. In this case, we find the maximum number of 'MATCHED' cases in the first step, and then find the maximum number of 'DETECTED' cases among the remainder from the first step.



result\GT	G1	G2	G3
R1	90.00%	85.00%	50.00%
R2	80.00%	82.00%	45.00%
R3	30.00%	0.00%	85.00%

(b) matching score

Figure 4. example of many-many overlapping

## Input GEDI Results and Ground Truth

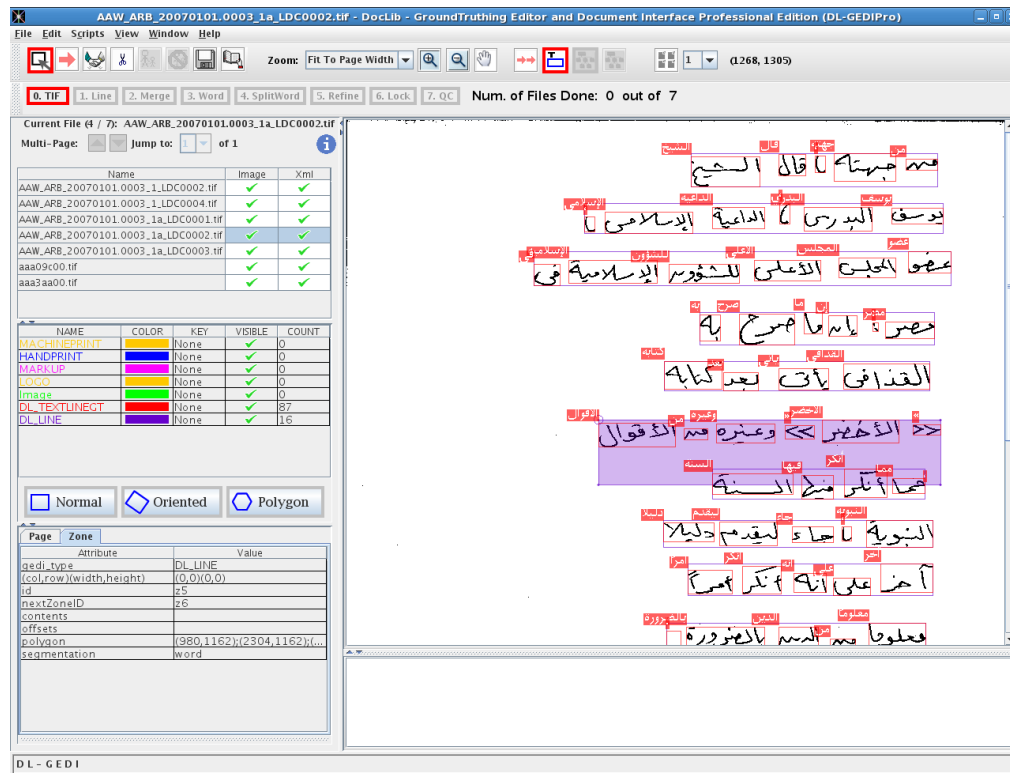


Figure 5. Screen shot of GEDI

The software uses XML which follows the GEDI data format specification for input of results and ground truth. GEDI is the Ground truthing Editor and Document Interface developed by the LAMP laboratory of University of Maryland Institute for Advanced Computer Studies.

### - GEDI tool

There are a limited number of functions that one should become familiar with. Additional details are provided in documentation with the software.

1. The upper left corner contains a list of files to be labeled. A new directory is loaded by selecting the file menu and choosing Load File/Directory. Click on the Name of the file to show both the image and XML metadata in the preview pane (to the right)

2. Below the file list is a set of possible zone types. Single clicking on the name of a zone type puts the system into draw mode. Double clicking will put the system into multi-draw mode (note the name becomes darker) and you can ground truth multiple boxes or polygons without reselecting a zone type.
3. At type bottom of the zone type panel, there is a choice of Normal, Oriented box or Polygon.



- a. A normal box is an upright rectangle. Drawing should start in the upper left corner of the text with the left mouse button down and dragging to the lower right. Sides and corners can be moved after selecting the polygon.
  - b. An oriented box is drawn by dragging a line from the upper left corner of the text, to the upper right corner, and dragging down to the bottom of the text box. Sides and corners can be moved after selecting the polygon.
  - c. A polygon is draw by single clicking points along the pologyn and double clicking to end. Points can be edited by selecting them or holding control down to add a point along an existing side.
4. The selection mode is indicated by the first icon on the



toolbar. When this is selected, one can select and delete existing zones.

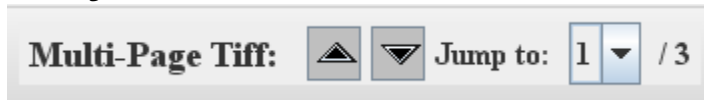
5. Control-Z and Control-Y implement undo, redo respectively.
6. A zoom function on the toolbar lets one zoom in on the document by percentage or in incremental steps



7. The type of zone is displayed in the middle of the box. Zone type labels can be turned off via Edit/Preferences. The size of the zone type label can be increased or decreased with control plus/minus.



8. Some documents are multipage. Multipage documents will show a navigator on the toolbar:



All pages should be labeled.

9. Zones are automatically saved when exiting or changing pages
- The Gedi software is available upon request from researchers at the LAMP laboratory or by contacting David Doermann([doermann@umd.edu](mailto:doermann@umd.edu))

#### - **Input GEDI xml format**

There are three xml tags which are essential to represent the document element. All of these are generated by default with GEDI.

##### 1. DL\_DOCUMENT

This tag represents the document image itself and this is the parent tag of the DL\_PAGE tag. This tag should have several attributes including 'NrOfPages', 'docTag' and 'src'.

- NrOfPages : number of pages which this image document consists of.
- docTag : usually 'xml'
- src : name of document image

##### 2. DL\_PAGE

This tag represents the page of document image and this is the parent tag of the DL\_ZONE tag. It is possible that there are several DL\_PAGE tags, if the document image is multiple page image such as 'TIFF' or 'GIF'. This tag should have several attribute such as 'gedi\_type', 'pageID', 'src', 'height' and 'width'.

- gedi\_type : usually DL\_PAGE
- pageID : identity of the page
- src : name of document image
- width : horizontal length of the page the image
- height : vertical length of the page of the image

##### 3. DL\_ZONE

This tag represents the zone in the page of image. This tag should have 5 essential attributes and can have user defined attribute. The essential attributes are 'gedi\_type', 'id',

'col', 'row', 'width' and 'height'.

- o gedi\_type : label of the zone such as Table, Text, Stamp and so on.
- o id : identity of the zone
- o col : upper right column point of the zone
- o row : upper right row point of the zone
- o width : horizontal length of the zone
- o height : vertical length of the zone
- o orientationD : rotation degree of the oriented box. It uses radian degree
- o polygon : points that consist of (col,row) for each point
- o lineID : identity of the one line of text
- o RLEIMAGE : Run-length code of the zone

```
<DL_DOCUMENT src="aaa3aa00.tif" docTag="xml" NrOfPages="2">
  <DL_PAGE gedi_type="DL_PAGE" src="aaa3aa00.tif" pageID="1" width="2544" height="3296">
    <DL_ZONE gedi_type="HANDPRINT" id="1" polygon="(153,282);(1155,145);(1163,390);(182,527);(149,274)" contents=""> </DL_ZONE>
    <DL_ZONE gedi_type="MACHINEPRINT" id="3" col="1712" row="2747" width="137" height="75" contents=""> </DL_ZONE>
    <DL_ZONE gedi_type="MARKUP" id="4" col="1579" row="1525" width="283" height="108"> </DL_ZONE>
    <DL_ZONE gedi_type="MARKUP" id="5" col="842" row="1566" width="370" height="110" contents=""> </DL_ZONE>
    <DL_ZONE gedi_type="MACHINEPRINT" id="6" col="1868" row="2700" width="172" height="92" orientationD="0.0" contents=""> </DL_ZONE>
    <DL_ZONE gedi_type="HANDPRINT" id="10" col="463" row="1178" width="738" height="100" orientationD="28.196" contents=""> </DL_ZONE>
    <DL_ZONE gedi_type="MARKUP" id="11" polygon="(1317,1338);(1496,1129);(1702,1259);(1605,1499);(1371,1453);(1317,1338)" contents=""> </DL_ZONE>
  </DL_PAGE>
  <DL_PAGE gedi_type="DL_PAGE" src="aaa3aa00.tif" pageID="2" width="2544" height="3296">
    <DL_ZONE gedi_type="HANDPRINT" id="7" col="108" row="486" width="1878" height="536" contents=""> </DL_ZONE>
    <DL_ZONE gedi_type="MACHINEPRINT" id="9" col="1652" row="2724" width="416" height="99" orientationD="1.102" contents=""> </DL_ZONE>
    <DL_ZONE gedi_type="MARKUP" id="12" col="448" row="1496" width="537" height="96" orientationD="18.772" contents=""> </DL_ZONE>
  </DL_PAGE>
</DL_DOCUMENT>
```

Figure 6. Example of GEDI xml

A sample GEDI config file is included with the software distribution.

```
<?xml version="1.0" encoding="UTF-8"?>
<!--GEDI was developed at Language and Media Processing Laboratory, University of Maryland.-->

<GEDI xmlns="http://lamp.cfar.umd.edu/media/projects/GEDI/" GEDI_version="2.1.06" GEDI_date="07/24/2008">
  <configuration>
    <type_settings>
      <type_setting_entry gedi_type="DL_PAGE"> </type_setting_entry>
      <type_setting_entry gedi_type="DL_LINE" color="#0000cc" visible="true" key="None">
        <type_attribute name="offsets" default="" possible_values="" ArbitraryVal="false"> </type_attribute>
        <type_attribute name="contents" default="" possible_values="" ArbitraryVal="false"> </type_attribute>
        <type_attribute name="segmentation" default="" possible_values="" ArbitraryVal="false"> </type_attribute>
      </type_setting_entry>
      <type_setting_entry gedi_type="DL_TEXTLINEGT" color="#ff0033" visible="true" key="None">
        <type_attribute name="offsets" default="" possible_values="" ArbitraryVal="false"> </type_attribute>
        <type_attribute name="contents" default="" possible_values="" ArbitraryVal="false"> </type_attribute>
        <type_attribute name="segmentID" default="" possible_values="" ArbitraryVal="false"> </type_attribute>
        <type_attribute name="segmentation" default="" possible_values="" ArbitraryVal="false"> </type_attribute>
        <type_attribute name="lineID" default="" possible_values="" ArbitraryVal="false"> </type_attribute>
      </type_setting_entry>
    </type_settings>
  </configuration>
</GEDI>
```

Figure 7. Example of GEDI config xml

## Evaluation Software

This software is a command line based program which evaluate the performance of zone segmentation and zone classification algorithms. The program is written in C++ and has been tested extensively the UINX/LINUX platform and Windows platform.

For the performance evaluation of zone segmentation, this program needs three files as input: an image file, a ground truth file which follows the GEDI format and a result file which also follows the GEDI format. In the GEDI format file, Document, Page and Zone are described as XML file format. A zone is defined geometrically as a rectangle, a rectangle with orientation or a polygon. A rectangle box has 4 attributes col and row, width and height. A rectangle box with orientation has one additional attribute which is the angle of orientation. A polygon box is a set of points.

Every zone has a 'gedi\_type' attribute that represent the label of zone. In the program, every zone in the result file is matched to every zone in the ground truth file.

If more than one zone has same line id represented by 'lineID', they will be merged together and treated as one zone if the 'lid' flag is set.

This software uses a matching score that is calculated using ON pixels. An ON or activated pixel is a foreground pixel which is black. To compare every zone from the result file and ground truth file, a matching score table is constructed, and the matching scores which are larger than a threshold are considered. There are four types of result as following :

- MATCHED : A result zone detecting a ground truth zone with same label
- DETECTED : A result zone detecting a ground truth zone with different label
- FALSEALARM : A result zone detecting no ground truth zone
- MISSED : A ground truth zone which is not detected by any result zone

The result is exported as text type file and xml type file for visualization.

The software distribution consists of following directories which contain :



- bin : executive program files
- src : source codes of the software
- data : data files used by the software
- doc : documentation files of the software
- examples : examples of ground truth and result which one can test the software

## Usage

Name :

PETS - Performance Evaluation Tools for zone segmentation and classification

Synopsis :

Unix/Linux platform command : PETS

Window platform command : PETS.exe

```
command -r {<FILE>|<DIR>} -g{<FILE>|<DIR>} -i {<FILE>|<DIR>}
        [-o <FILE>] [-v <DIR>] [-m <FILE>] [-t <NUM>] [-detail]
        [-lid] [-rle] [-segonly|-zoneclass]
        [-az <FILE>|-naz <FILE>]
```

Options :

-r {<FILE>|<DIR>}  
: Location of Results File(s)

-g {<FILE>|<DIR>}  
: Location of Ground Truth File(s)

-i {<FILE>|<DIR>}  
: Location of Image File(s). Default location is the location of ground truth

-o <FILE>  
: Name of File for Evaluation Results. Default is 'PETS-Eval.txt'.

-v {<FILE>|<DIR>}  
: directory where Xml output of GEDI format will be saved

-lid  
: Zones which have same 'lineID' attribute in Ground truth will be merged to one zone

-rle  
: run-length code will be add to visualization output

-detail  
: enable detailed output for each zone

-t <NUM>

: set the threshold by user for determining a zone match based on pixel counts. Default is 80(%).

-m <FILE>  
: result zones which are in a ground truth zone will be merged if it's types are in the <FILE>. First line of the FILE should have numeric data which is used as threshold for zone merging.

-segonly  
: Evaluation will perform detection by not consider zone labels for matching.

-zoneclass  
: Evaluation will rely on ZoneIDs for correspondence, considering only zone labels for results

-az <FILE>  
: Zones which its types are in the <FILE> will be treated in the program, otherwise deleted from the result.

-naz <FILE>  
: Zones which its types are in the <FILE> will be deleted from the result.

## Output of the program

There are two types of output from the program. One is the text file which has the detail result of evaluation process for every single zone and the summarized result for all zones.

In the individual result, symbols, 'O', '-' and 'X', are used to represent the result of each zone for easy reading. 'O' means that the result zone detected one ground truth zone with matching score over the threshold and the type of two zones matched. '-' is same as 'O' in terms of detection, but the type of two zones is not matched. 'X' means that the result zone is a false alarm. The matching score is displayed if the result zone detects the one of ground truth zone with matching score over threshold. There are also overall result for each page at the end of the individual result.

```
=====
Result of Individual File
=====

[O] : Matched, [-] : Detected, [X] : False Alarm

AAW_ARB_20070101.0003_1_LDC0002.tif
=====

Page ID : 1
-----
[O]      M0,      zone,      z0,      zone, 99.79%
[O]      M1,      zone,      z1,      zone, 99.89%
[O]      M2,      zone,      z2,      zone, 99.60%
[O]      M3,      zone,      z3,      zone, 99.68%
[X]      M4,      zone
[X]      M5,      zone
[O]      M6,      zone,      z9,      zone, 81.76%
[O]      M7,      zone,      z10,     zone, 99.93%
[O]      M8,      zone,      z11,     zone, 99.92%
[O]      M9,      zone,      z12,     zone, 99.90%
[O]      M10,     zone,      z13,     zone, 99.87%
[O]      M11,     zone,      z14,     zone, 99.71%
[O]      M12,     zone,      z15,     zone, 99.88%
[X]      5,      zone
[X]      6,      zone
[X]      13,     zone
[OVERALL] 11/0/5/16, 68.75%
```

*Figure 8. Example of individual result of a page*

In the summary of result section, There are some information on zones and confusion matrix for the evaluation result which is very commonly used tool for visualization of classification result. Finally, summarized result for each type of zones are shown as type of table. Rate of precision, recall, F-score, missing and false

alarm is representing the result of the segmentation algorithm and the classification algorithm.

User can switch on or off the individual result by using '-detail' option and can specify the name of the result file using '-o <FILE>' option. Default file name is "PETS-Eval.txt" otherwise '-o <FILE>' option is not used.

```

=====
Summary of Results
=====

- Number of Files : 7
- Accuracy of Zone Detecting : 34.76%

01. Information on Zones
=====
Label   Class of Zone           Num. in Result   Num. in G-Truth   Accuracy
-----
1       zone                    210              97                34.76%

02. Confusion Matrix
=====
Result\GT   unmatched           1
-----
unmatched   0( 0.0%)*  24(24.7%)
            1 137(65.2%)  73(34.8%)*

03. Result Table
=====
Label   GT   Result   Correct   Precsion   Recall   F-Score   Missing   FalseAlarm
-----
1       97    210      73        34.76%     75.26%    47.56%    24.74%     65.24%

```

Figure 9. Example of summary of results

The other output of the program is a GEDI file which is very useful when user wants to check the results of algorithm visually. There are four types of zones in the visual output: 'MATCHED', 'DETECTED', 'FALSEALARM' and 'MISSED'. Each zone has more information which is useful for understanding the result of algorithm include 'GTID', 'GTClass', 'RESID', 'RESClass', 'MZID' and 'Score'.

- GTID : ID of the ground truth zone
- GTClass : Type of the ground truth zone
- RESID : ID of the result zone
- RESClass : Type of the result zone
- MZID : ID of the merged zone, merged zones have same MZID
- Score : Matching score of the result zone

User can check these information at the attribute window of GEDI program by click a zone in the screen.

The program will generate the visualization output when '-v <DIR>' option is used and the visualization output file name has 'evl' in the middle of it original file name.

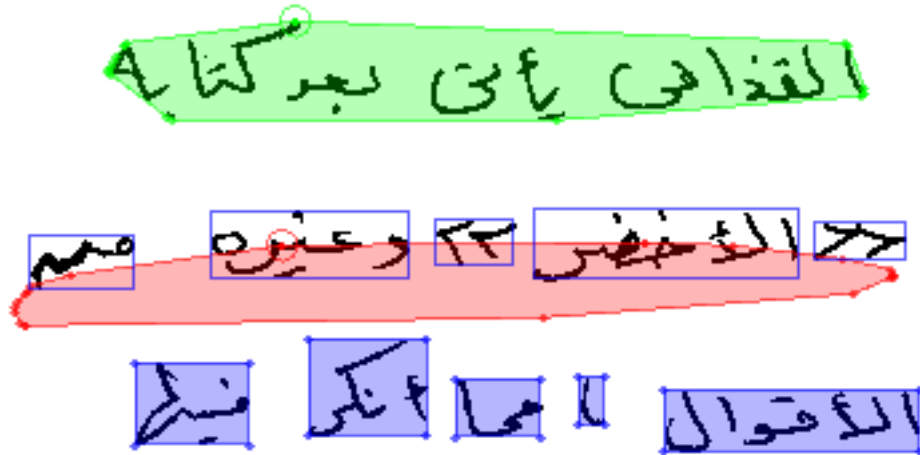


Figure 10: Zone examples of visualized output

Page Zone	
Attribute	Value
gedi_type	DETECTED
(col,row)(width,height)	(0,0)(0,0)
id	12
DL_RLEIMAGE	(0,351,353);(1,350,354);(...
GTCClass	DL_LINE
GTID	z3
MZID	11
RESCClass	DL_TEXTLINEGT
RESID	12
contents	
polygon	(1843,700);(1842,700);(...
score	99.69

Figure 11: Attribute window of GEDI


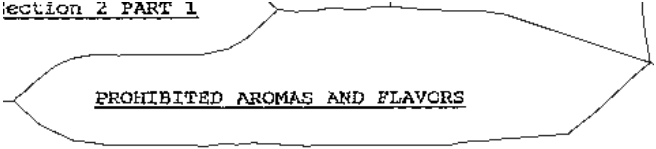
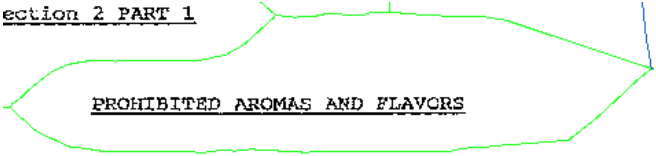
# Examples

- Cases by types of result

1. Matched zone

Using the pixel overlap metric, a result zone is specified as a Matched zone if the matching score from the result and ground truth is over a threshold and the types of two zone are identical.


Table 1: Example of matched zone

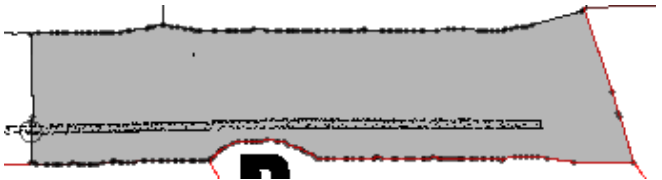
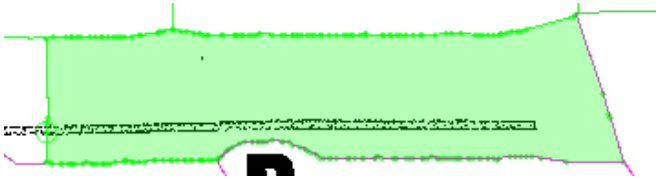
Category	Example
Ground Truth	
Result	
Output	

2. False Alarm zone

False Alarm zone means that there is no matching ground truth zone for a result zone. It is a wrong segmented zone of the segmentation algorithm.

Table 2: Example of False Alarm zone

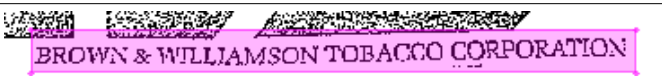
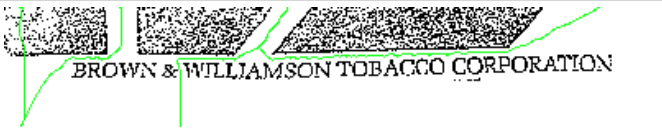
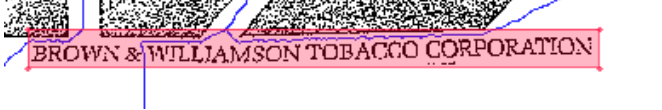
Category	Example
Ground truth	

Result	
Output	

### 3. Missed zone

Missed zone means that there is no matching result zone for a ground truth zone.

*Table 3: Example of Missed zone*


Category	Example
Ground truth	
Result	
Output	

## - Cases by operations on zone



### 1. Result zone merging

This program provides a functionality that two or more result zones are merged when the result zones are covered by one ground truth zone and the type of result zones are in the list specified by the option '-m <FILE>'.

*Table 4: Example of result zone merging*

Category	Example
Ground truth	



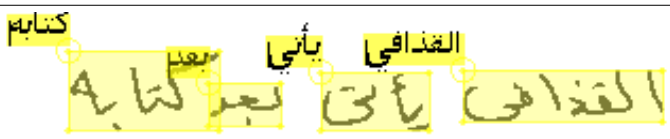


Result	
Output	

## 2. Ground truth zone merging

Ground truth zone can be merged if the zones have identical value in 'lineID' attribute and '-lid' option is used. This merging process is done before the matching process, so the matching score is calculated based on the ground truth zone that is merged.

The goal of this option is to provide evaluation of the special case where lines of text are represented by sets of polygons.

*Table 5: Example of Ground truth zone merging*

Category	Example
Ground truth	
Result	
Output	

## - Zone type selection

### 1. Zones to be considered

If '-az <FILE>' option is used, only zones which its type is in the list specified in the <FILE> will be considered as a zone, others are disregarded in the matching process and will not be even displayed in the visualized output.


## 2. Zones to be not considered

'-naz <FILE>' option is the opposite to '-az <FILE>' options. The program will disregard the zone from the matching process if the type of zone is in the list specified in the <FILE> and will not displayed in the visualized output.

### - Complex zone representation

GEDI provides three types of zoning method but if the user wants more detailed representation in some cases like exclusive intersection zone, user can use Run Length Encoding by 'RLEIMAGE' attribute. If a zone has 'RLEIMAGE' attribute, only pixels represented in the encoding will be treated in the matching process. Run Length Code is consist of three values in the form of '<x, y, r>'. 'x' is the offset of column from the zone origin point, 'y' is the offset of row from the zone origin point, and 'r' is ( x + run + 1 ). GEDI has option to display the run length code. Visualization output file will have Run Length Code for each zone if '-rle' option is used.

*Table 6: Example of Run Length Code representation*

	Example
Zone w/o RLE code	
Zone w/ RLE code	