FB 19| Geography
Department of Environmental Informatics.

Philipps Universität Marburg

Dirk Zeuss
Spaska Forteva
Madhuvanthi Venkatesh

Course Name: Species Distribution Modelling *(SDM).*
Software: *Digitize IT.*
Module: *Object Detection.*

Introduction:

Detecting the elements from the scanned input images is called Object Detection. Since, our goal is to produce georeferenced raster maps from the text books, the desirable output of object detection will be to detect and extract the maps from the textbook. Object Detection is done through template matching algorithm.

Two sets of images namely, input images and template images are the image inputs to the template matching algorithm.

Input Images:

Input Images are the scanned pages from the textbook.

Template Images:

Template Images are the cropped images of desired maps or objects to be detected. The template images are the sample maps to be detected and they can be manually cropped through the shiny application (Front end application of Digitize IT software).

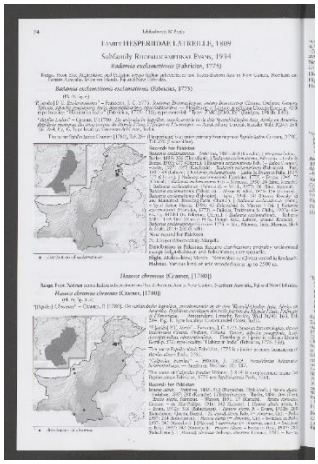


Fig 1: *Page 39 from the Scanned Input, "Butterflies of Pakistan".*

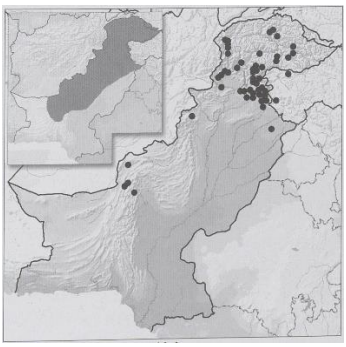
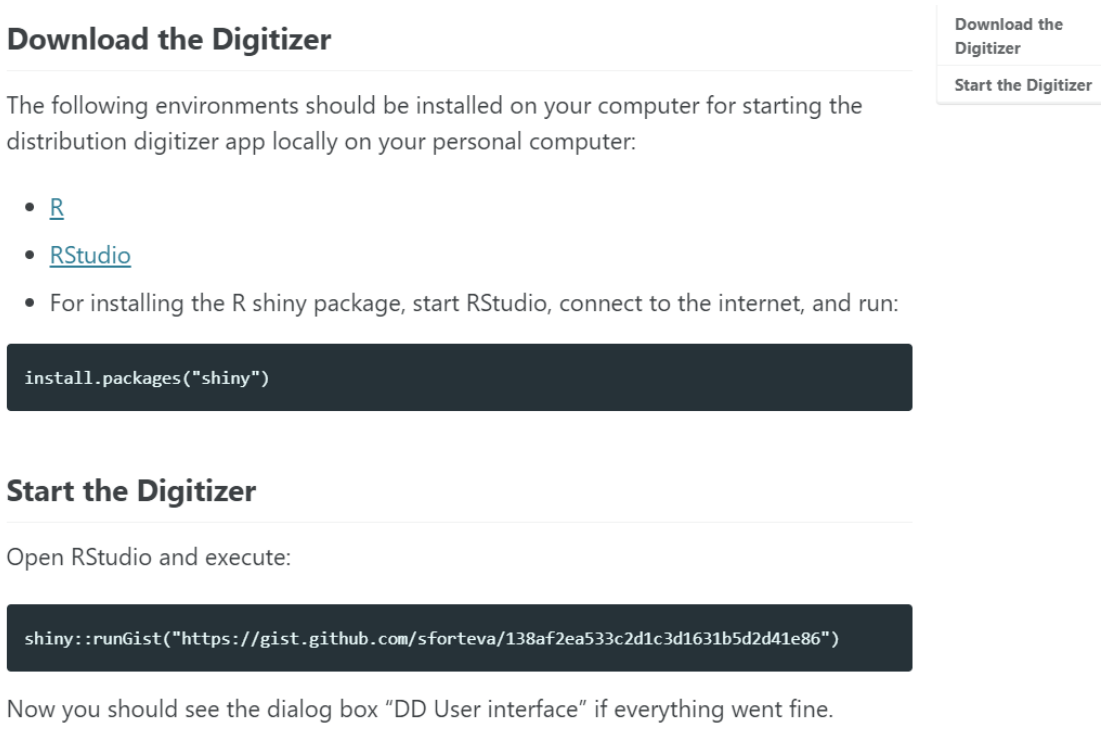


Fig 2: *Map from the Page 51 in the Scanned Input, "Butterflies of Pakistan".*

Note: *Template images from other pages can also be used to extract the maps.*

Extracting template images through the application window:

Step 1: Run the DD shiny application by copying and pasting the link in the R Studio. This is available in the website "--------". Follow the tutorial step by step to run ths shiny application.

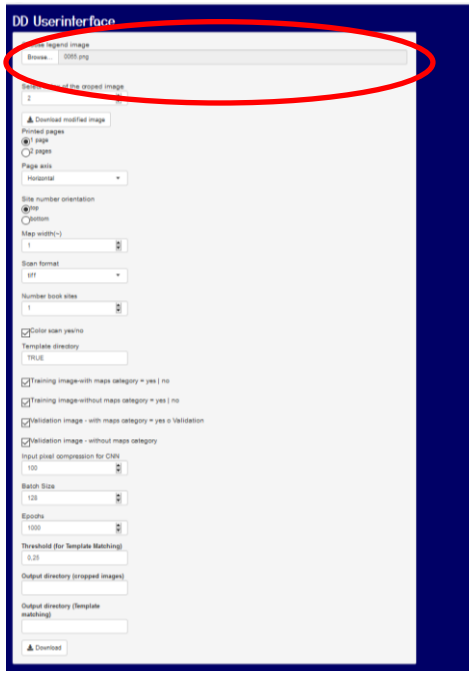


FB 19| Geography
Department of Environmental Informatics.

Philipps Universität Marburg

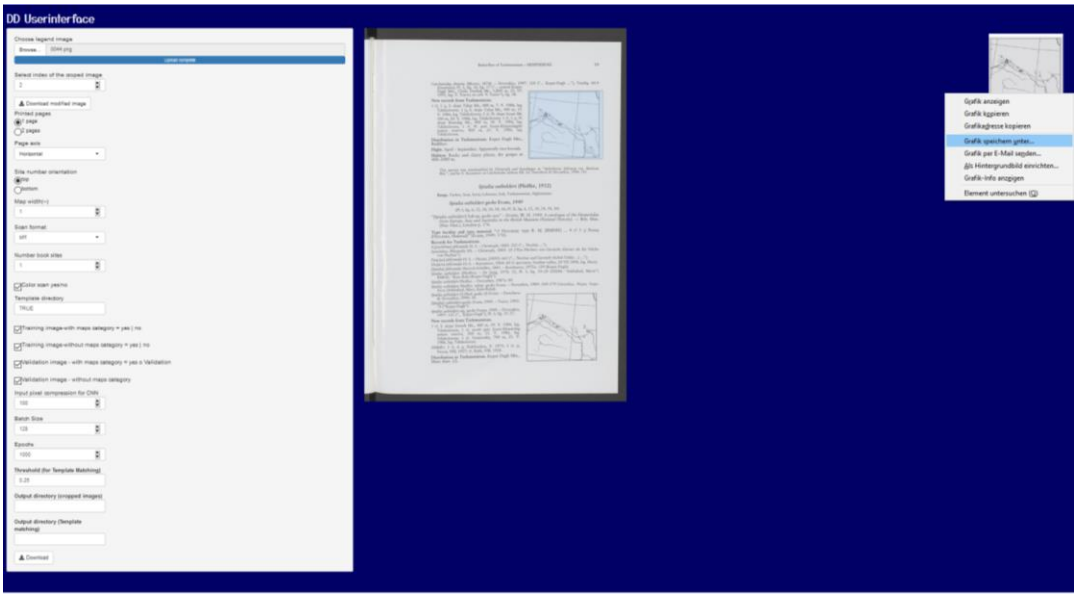
Dirk Zeuss
Spaska Forteva
Madhuvanthi Venkatesh

Step 2: Upload the images in PNG format and choose the options as shown in the image.



Note: *The files are in .zip format, extract them before uploading into the application.*

Step 3: Once the image gets uploaded, drag the mouse pointer around the map and right click on the dragged region to save the template image.



Dirk Zeuss
Spaska Forteva
Madhuvanthi Venkatesh

Template Matching Algorithm:

The template matching algorithm performs cross correlation between the input image and the template image. Cross correlation function calculates the similarities between the input image and the template image. If there is any similarity between the input and the template image, then the cross correlation output is greater than zero. In case, if there is no match between the input image and the template image, the output of cross correlation will be zero (in ideal case).

Since only a few set of template images are used to extract the raster maps from the scanned input pages, the output of cross correlation for the matched template and input image will be slightly higher than zero in some of the pages. This will be the controlling parameter by the user to extract the raster maps and this parameter is called as threshold.

The maximum value for the accurate map extraction (without texts or other elements in the page) would be 1. So the threshold value should always be lesser than 1. Threshold is the minimum value (of output of cross correlation) at which all the maps from the given input Scanned pages can be extracted.

Example: *For the extraction of a raster map for the given input image, the threshold value used was 0.3.*

Controlling Template Matching Algorithm:

The template matching algorithm can be controlled by changing,

- The set of template images (either increasing or decreasing them).
- The threshold value.

Output of Template Matching Algorithm:

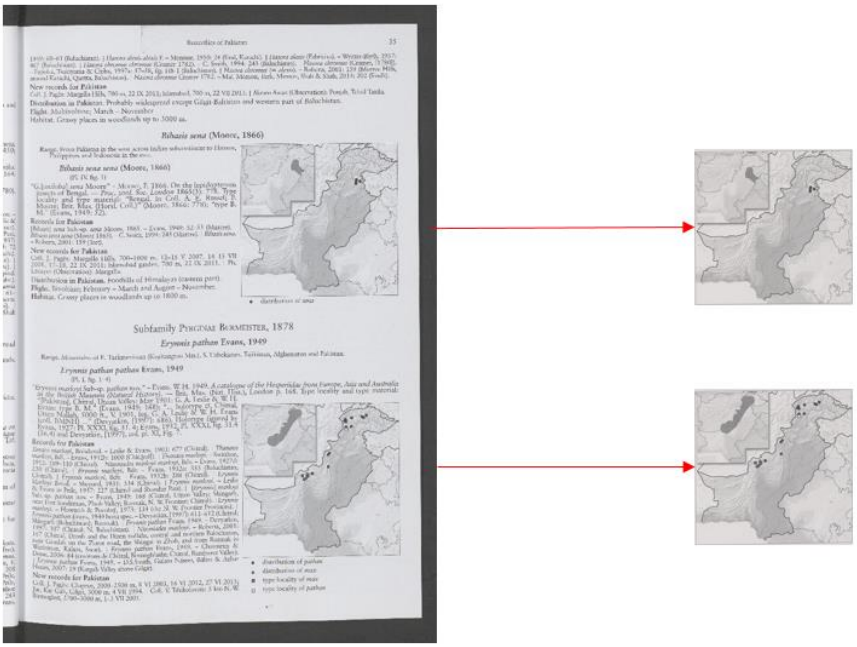


Fig 3: *Output of Template Matching Algorithm.*

Coordinates and the other outputs of template matching:

Table 1 shows the coordinates and the other csv outputs of the template matching program. The spreadsheet output contains the file name, coordinates (x1, y1, x2, y2), size, threshold and time. File name represents the name and the directory of the input image file. The coordinates (x1, y1, x2, y2) represent the coordinates of the rectangular maps detected in the input images. Time is the consecutive time taken for the template matching between the input image and the template image. For example, if the time for the execution of the first image is ‘x’ seconds, then the time for execution of the second image is written as ‘x + t’ seconds, where,‘t’ is the consecutive time in seconds taken and this consecutive time represents the multiple template matching within a single map of the input image (time taken for searching the right template image and detecting a single map within the image).

Table 1: *Coordinates and size output of template matching.*

Filename	x1	y1	x2	y2	size	threshold	time
/content/drive/My Drive/testpakistan/0217.tif	1197	1221	4431	1502	58.9328257	0.25	18.3143082
Filename	**x1**	**y1**	**x2**	**y2**	**size**	**threshold**	**time**
/content/drive/My Drive/testpakistan/0217.tif	1190	1217	4427	1500	58.3962542	0.25	22.9821382

Pixels:

The word pixel is derived from the combination of two words namely, picture element. Pix = picture, el = element. Pixels are smaller units or elements that comprise a digital image.

Resolution:

The number of pixels in an image is called as resolution. The higher the number of pixels in an image, the higher the resolution it would constitute, the higher the quality of image would be.

Calculation of size through the resolution:

The size is the conversion of the area of rectangle in the image which is in terms of number of pixels to the area of the rectangle in cm^2.

This is done through the formula:

Size = w * h * (2.54 / no of pixels) * (2.54 / no of pixels)…… [1]

Where,

w = width of the image (in number of pixels).

h = height of the image (in number of pixels).

Example:

Size of the map in the textbook = *56.25* cm^2.

Size (from the output) *= 58.396* cm^2.

Note: *The values of pixels and size of the maps might differ for each and every book.*

The size from the output of the program might not match exactly with the actual size, this is because the width and height are just the approximations from the number of pixels. Hence, when there are minute approximations in the height and width (in the number of pixels), there would be a slight deviation when the area (w x h) is calculated.

Tasks:

1) Find the real size in cm^2 of the input images by trying several resolutions in dot per inch (DPI).
2) Find maximum value of threshold for template matching.