

Course Name: Species Distribution Modelling (SDM).

Software: Digitize IT.

Module: Object Detection.

Software Tutorial



Tasks (12th May)

- Extraction of Template Images.
- Find the real size in cm^2 of maps on your test images by trying several resolutions in dots per inch (DPI). **Not, Number of Pixels!**
- Find the maximum value for the threshold of template matching.

Output

Table 2: 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

(w, h)

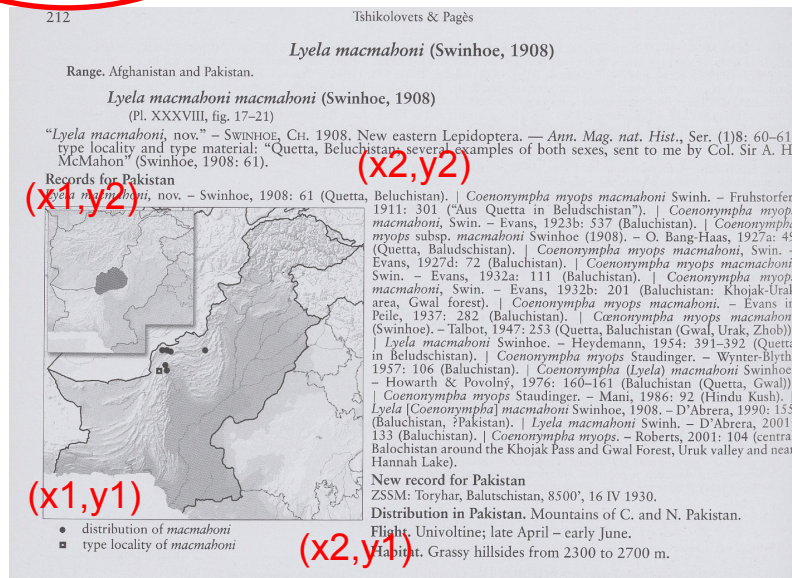


Figure 12: Page 212 [2]

Output

```
rows = [[tiffname, w, h, pt[1] + w, pt[0] + h, size, threshold, (time.time() - start_time)]]
```

Filename	x1	y1	x2	y2	size	threshold	time
/content/drive/MyDrive/Book 14/0124.tif	214	254	392	637	2.19176981	0.2	0.7414739132
Filename	x1	y1	x2	y2	size	threshold	time
/content/drive/MyDrive/Book 14/0123.tif	214	254	340	319	2.19176981	0.2	1.596010208

x1 = pt[1]
y1 = pt[0]
x2 = x1 + w
y2 = x2 + h

Size of the map in the textbook = 51.2 (8*6.4) cm².

Size = w * h * (2.54 / no of pixels) * (2.54 / no of pixels)

No of Pixels ~ 250 (261-262).

What's next?.

Let's execute the program together.

- Step 1: Data.

<https://rstudio.cloud/>

https://github.com/environmentalinformatics-marburg/distribution_digitizer_students.git

- Step 2: R-Cloud.

```
install.packages("reticulate")
```

```
library(reticulate)
```

```
os <- import("os")
```

```
library(reticulate)
```

```
use_python("/usr/local/bin/python")
```

- Step 3: Packages.

```
py_install(packages = "opencv-python", pip = TRUE)
```

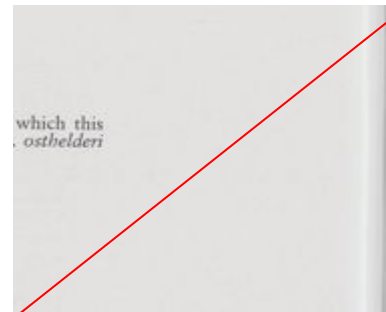
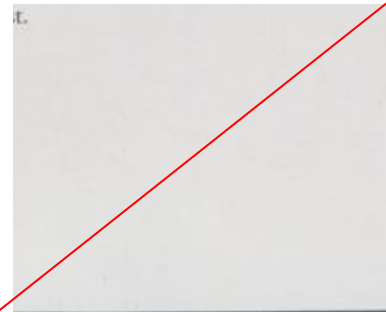
```
py_install(packages = "pillow", pip = FALSE)
```

- Step 4: Execution .

```
source_python("template_matching_png.py")
```

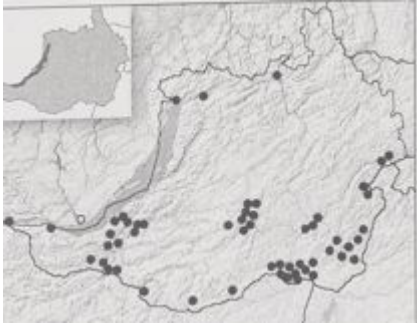
Outputs (Threshold = 0.25)

- 2020_suprascan_00030 The Butterflies of Turkmenistan



Outputs (Threshold = 0.25)

- **2020_suprascan_00034 The Butterflies of Transbaikal Siberia.**



- **2020_suprascan_00050 The Butterflies of Afghanistan.**



Manually filter and choose the outputs for georeferencing!.

Importance of Records

- Helps to identify the missing files. **For example** : I can find out whether the file was executed or not in case if I don't get any output!.
- The threshold and order of execution.
- Time for executing the files (how fast the program is ?!).
- The coordinates can be used for training Convolutional Neural Network(CNN) or any other neural networks for training the data.

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