Distribution Digitizer – Installation Guide (2025)

Your Name

Contents

1 Distribution Digitizer – Installation Guide (2025)

This document provides the complete installation instructions for the **Distribution Digitizer** under **Windows 10/11**, including:

- R
- RStudio
- Python
- Tesseract OCR
- GDAL / OpenCV
- Required R & Python packages
- Shiny app startup instructions

This version is updated for the current project structure.

---

2 Citation

If you use this program, please cite:

Venkatesh M, Forteva S, Zeuss D (2021)
Distribution digitizer: Software for digitizing species distributions from analogue maps.
Version 0.0.1
https://github.com/environmentalinformatics-marburg/distribution_digitizer_students

---

3 1. Install Required Software

Please install the following before running the app:

3.1 R (recommended: 4.4.x or higher)

https://cran.r-project.org/

3.2 RStudio Desktop

https://posit.co/download/rstudio-desktop/

3.3 Git (optional)

https://git-scm.com/download/win

3.4 Python 3.11.x

https://www.python.org/downloads/

Install Python as Administrator
Enable: "Add Python to PATH"

3.5 Tesseract OCR

https://github.com/UB-Mannheim/tesseract/wiki

Install English + German language packs.

---

4 2. Download or Clone the Project

4.1 Option A – Using Git

```
git clone https://github.com/YourUsername/distribution_digitizer.git
```

4.2 Option B – Download ZIP

- Download ZIP
- Extract
- Rename folder to:

```
distribution_digitizer
```

---

5 3. Required Project Structure

Your project must look like this:

```
distribution_digitizer/

   app/
       server/
       ui/
       functions/
       app.R

   config/
   data/
       pages/
       templates/
       books/

   www/

   start_digitizer.bat
   app_start.R
```

---

6 4. Install Required R Packages

Run RStudio **as Administrator**.

```r
install.packages(c(
  "shiny",
  "shinydashboard",
  "shinyFiles",
  "shinyalert",
  "shinyjs",
  "DT",
  "magick",
  "grid",
  "reticulate",
  "tesseract",
  "leaflet",
  "raster",
  "sf",
  "jsonlite",
  "stringr",
  "dplyr",
```

```
  "tidyr",
  "remotes"
))
```

Optional Dropbox support:

```
remotes::install_github("karthik/rdrop2")
```

---

7 5. Install Python Packages

Open **CMD (not admin)**.

Check Python:

```
python --version
```

Install Python modules:

```
pip install opencv-python pillow pandas GDAL imutils rasterio geopandas numpy pytesseract shape
```

---

8 6. Test R–Python Connection

In R:

```
library(reticulate)
py_config()
```

Should show Python 3.11 and installed modules.

---

9 7. Start the Application

9.1 Option A – Using start script (recommended)

Double-click:

```
start_digitizer.bat
```

9.2 Option B – Manual start in RStudio

```r
setwd("D:/distribution_digitizer")
options(shiny.port = 8888, shiny.host = "127.0.0.1")
shiny::runApp("app")
```

---

10 Troubleshooting

10.0.1 GDAL not found

```
pip install GDAL==3.10.1
```

10.0.2 Tesseract path missing

```r
Sys.setenv(TESSDATA_PREFIX = "C:/Program Files/Tesseract-OCR/tessdata")
```

10.0.3 DT missing

```
install.packages("DT")
```

---

11 8. Optional: Web-Demo / Docker Deployment

A Dockerfile + lightweight demo version (small uploads only) can be deployed to:

- Render.com
- HuggingFace Spaces
- GitHub Actions

This is for publication and demonstration, not large files.

---

12 More Information

Project homepage:
https://environmentalinformatics-marburg.github.io/distribution_digitizer_webpage

---

End of README.