

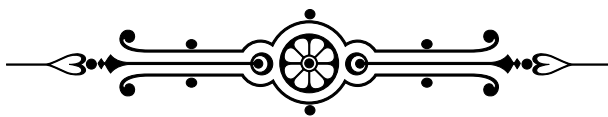
图片查重程序

2024-06-27

LiChuang Huang



@ 立效研究院



Contents

1	摘要	1
2	安装	2
2.1	Python 3.9.0	2
2.2	findsimilar	3
2.2.1	(Option 1) 运行 bat 文件安装 (未测试)	3
2.2.2	(Option 2) 通过命令安装	3
3	使用示例	3



List of Figures

1	Panel	2
2	Unnamed chunk 11	5
3	Unnamed chunk 12	6



List of Tables

1 摘要

Findsimilar 可递归搜索文件夹下所有图片 (没有明确的数量上限, 但最少 10 张图片), 根据设定的阈值 (threshold 参数), 寻找相似图片, 最后将结果以网页报告的形式输出。

步骤 1: 选择搜索目录 (需要搜索的路径)。步骤 2: 选择输出目录 (生成的报告文件和其他分析数据存放)。

步骤 3: 点击 “Run”。步骤 4: 点击 “Similarity Gallery”, 在网页浏览器中显示报告。

注意: 支持的格式: .png, .jpg, .jpeg, .gif, .giff, .tif, .tiff, .heic, .heif, .bmp, .webp, .jif

注: 以下面板示例为 Linux 系统的界面, Windows 下略有不同。



Figure 1 (下方图) 为图 Panel 概览。

(对应文件为 **Figure+Table/Panel.png**)

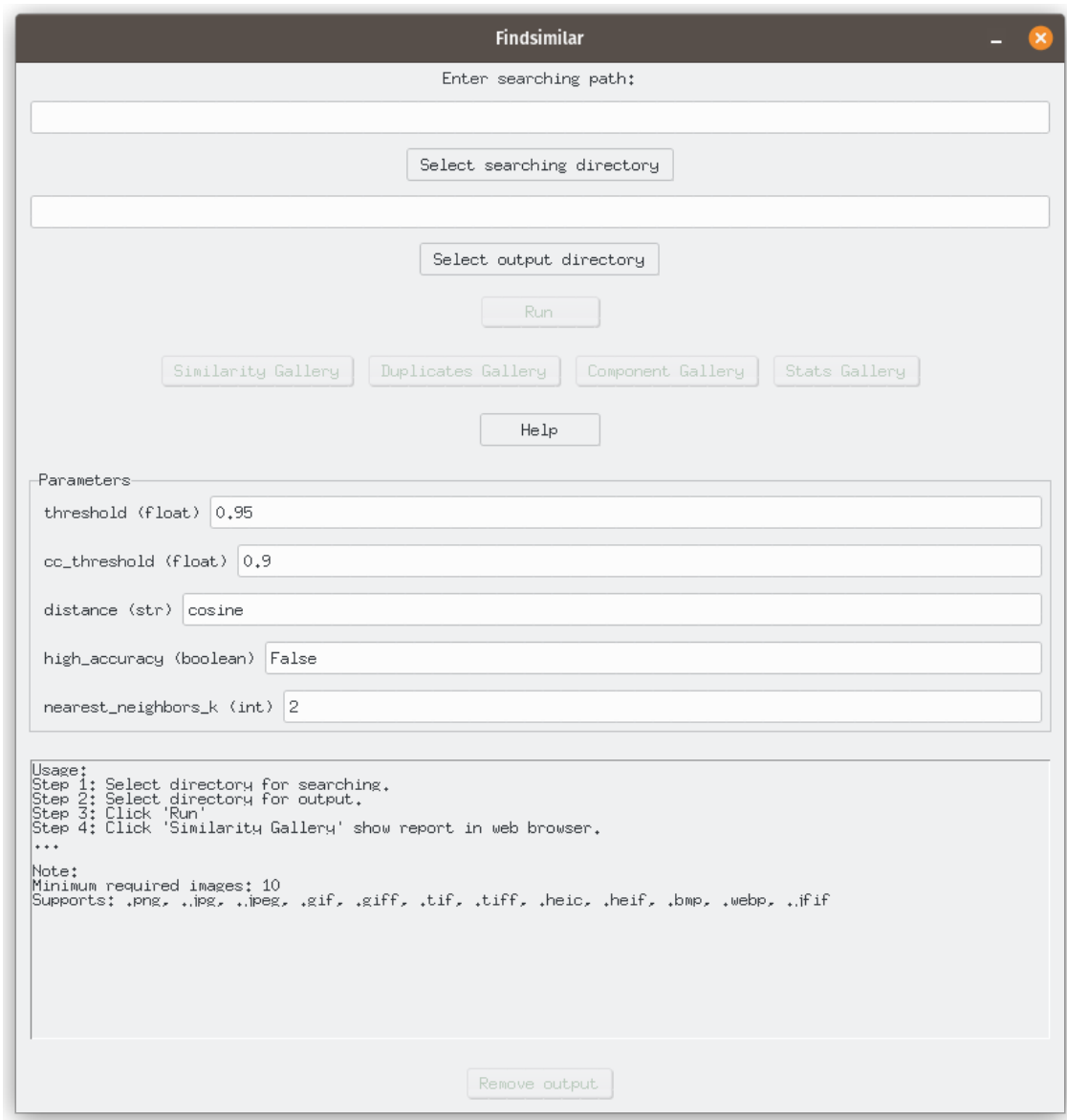


Figure 1: Panel

2 安装

2.1 Python 3.9.0

请下载并安装 Python 3.9.0:

<https://www.python.org/ftp/python/3.9.0/python-3.9.0-amd64.exe>

2.2 findsimilar

2.2.1 (Option 1) 运行 bat 文件安装 (未测试)

请确认已安装完成 Python, 然后打开文件夹



‘All files’ 数据已全部提供。

(对应文件为 ./use_for_install/)

注：文件夹./use_for_install/共包含 3 个文件。

1. findsimilar
2. get_shortcut.bat
3. install.bat



1. 双击 install.bat, 等待安装完成。
2. 双击 get_shortcut.bat, 这会在桌面生成 findsimilar.bat。
3. 完成。

2.2.2 (Option 2) 通过命令安装

按 Win + R, 输入 cmd, 确认打开 cmd 界面, 输入以下安装。

cmd input

```
# 请确认 cmd 已经切换到 findsimilar 安装包的路径
pip install findsimilar -i https://pypi.tuna.tsinghua.edu.cn/simple
```

cmd input

```
echo $(which findsimilar) > "%USERPROFILE%\Desktop\findsimilar.bat"
```

3 使用示例

可以使用如下文件夹作为测试。



‘Test files’ 数据已全部提供。

(对应文件为 `./test/`)

注：文件夹`./test/`共包含 10 个文件。

1. Control-1.tif
2. Control.tif
3. DAPIPKH67.tif
4. Figure 4 revise.tif
5. Figure 5 revise.tif
6. ...



新建一个空的文件夹作为输出目录。



Figure 2 (下方图) 为图 unnamed chunk 11 概览。

(对应文件为 `Figure+Table/unnamed-chunk-11.png`)

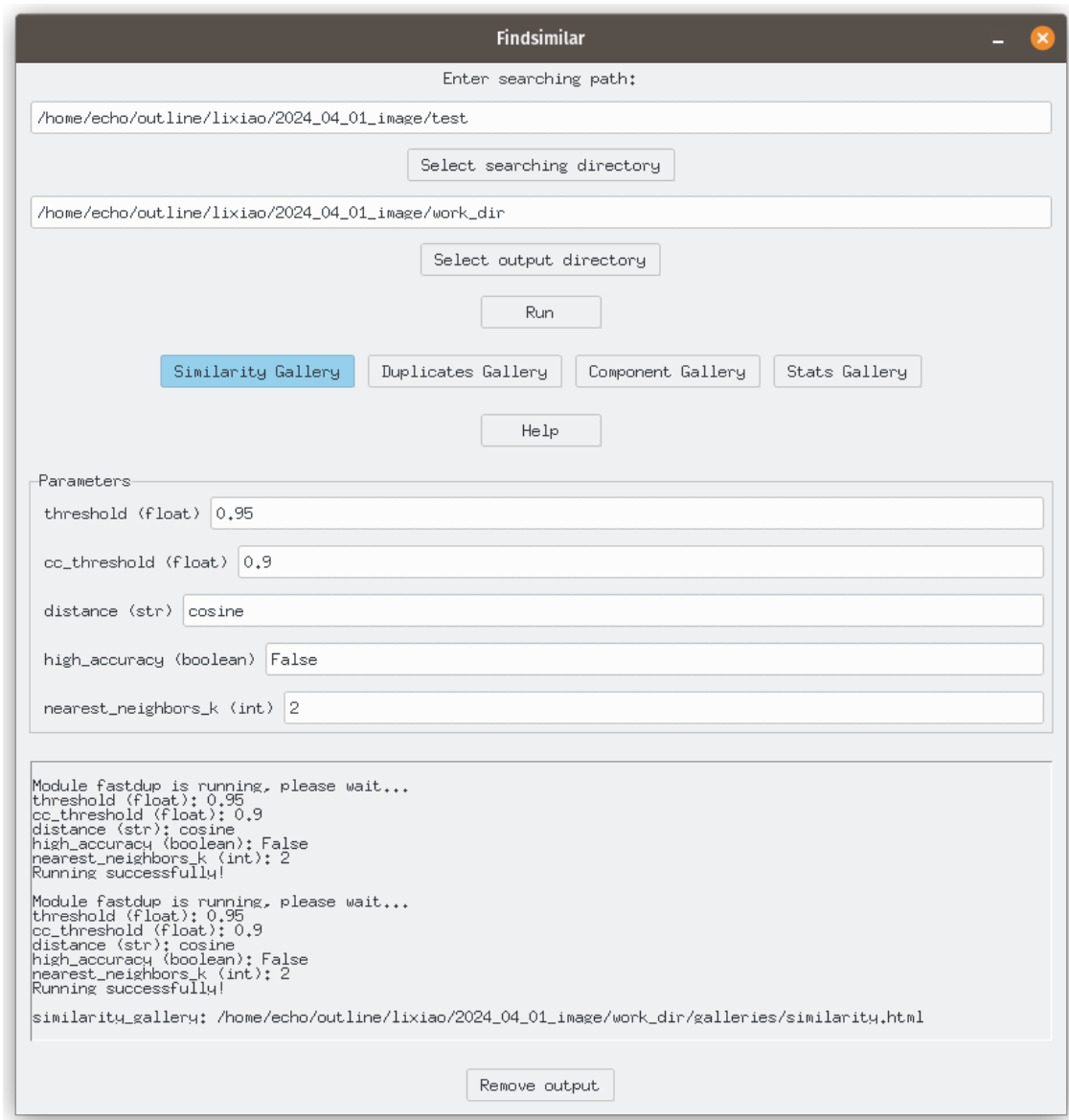


Figure 2: Unnamed chunk 11



如上运行完成后，点击“Similarity Gallery”，得到 HTML 报告。



Figure 3 (下方图) 为图 unnamed chunk 12 概览。

(对应文件为 Figure+Table/unnamed-chunk-12.png)

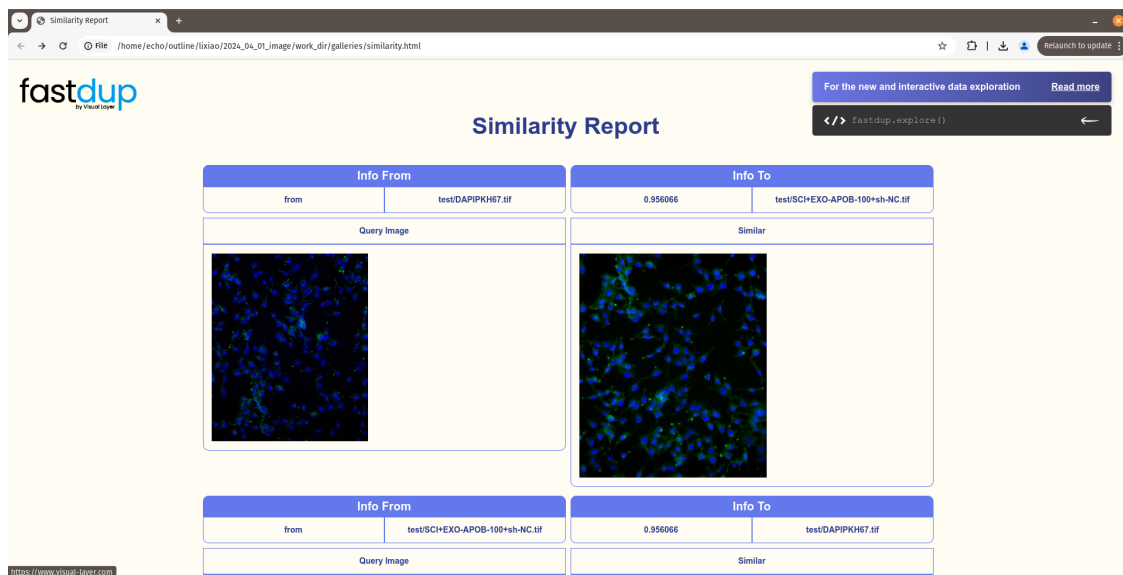


Figure 3: Unnamed chunk 12

