Supplementary Information

Design and development of fMRI Dicom file automatic collating software

Xinlin Huang [1,#], Xuanze Wang [2,#], Feng Gao[3], Jing Qin[4], Hongyan Sun [3,*]

[1] *Clinical Medicine college, Jiamusi University School, Heilongjiang Jiamusi 154000, China*

[2] *Molecular Electrochemistry for Energy Laboratory, School of Energy Science and Engineering, Vidyasirimedhi Institute of Science and Technology (VISTEC), Rayong, 21210, Thailand.*

[3] *Suzhou Guangji Hospital, Jiangsu Suzhou 215000, China*

[4] *Department of radiology, Shanghai Anting Hospital, Shanghai 20000*

[#] These authors contributed equally

*Address all correspondence to Hongyan Sun: hoya12@126.com

File Processor Guide
Abbreviation: DFP

Contents

1 System overview

1.1 Development background

For a long time, in the field of medical imaging research, to do a scientific research project of big data research and observation, the workload of data collation in the early stage is heavy, cumbersome, and complicated, error-prone, time-consuming, and labor-intensive. In the era of the modern Internet and big data, "Internet + scientific research" puts forward a new scientific research organization form and method for interdisciplinary cooperation. Mistakes will lead to deviations in the final research results and can also greatly improve work efficiency.

1.2 Development framework

MATLAB is a collection of numerous computational algorithms. It has more than 600 mathematical operation functions used in projects, which can conveniently realize various calculation functions required by users. The algorithms used in the function are the latest research results in scientific research and engineering calculations and have undergone various optimizations and fault-tolerant processing. MATLAB includes many toolboxes. Each toolbox is a collection of functions dedicated to a specific problem. From this perspective, a MATLAB toolbox is a library of functions dedicated to a problem. These function libraries are carefully designed by professionals, and their performance and quality are guaranteed.

At present, most of the image processing and analysis software for medical imaging is developed based on MATLAB, and organizing image data files is also for better processing and analysis of image data. Based on the above situation, write codes based on MATLAB, package It is a software that can organize Dicom files. It has a simple operation interface and automatic batch sorting, which can effectively reduce the preparation time of scientific research work, as well as the operation errors caused by manual work, and improve the efficiency of scientific research.

2 About DFP

The full name of the software is Dicom File Processor, or DFP for short, which is software for automatically batch sorting Dicom file formats.

DFP can realize the batch extraction and renaming of files or folders, extract the specified functional image and structural image from the folder of the data set, store them in the corresponding folders, and rename them according to the serial number, the same case data The new folder names of the Bold image and the T1 image are the same, realizing the one-to-one correspondence between the Bold image and the T1 image; it is also possible to realize the batch extraction of a specified file or folder.

The DFP software can classify the data downloaded from the magnetic resonance machine without sequence classification (such as the data downloaded from the CD), and sort them according to the scanning sequence conditions, which is also a batch sorting scan sequence.

3 Computer system requirements

3.1 Hardware

RAM at least 8GB

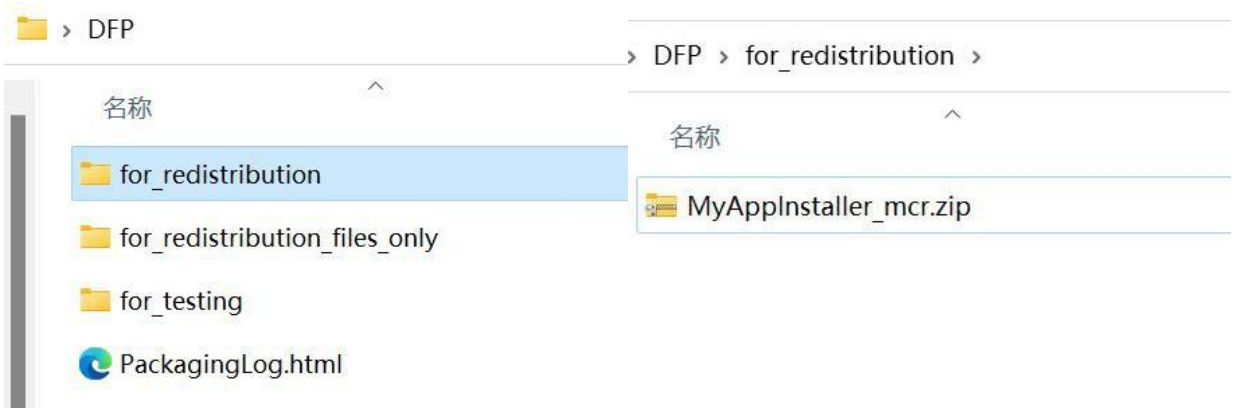
Free hard disk space at least 4GB

3.2 Operating system

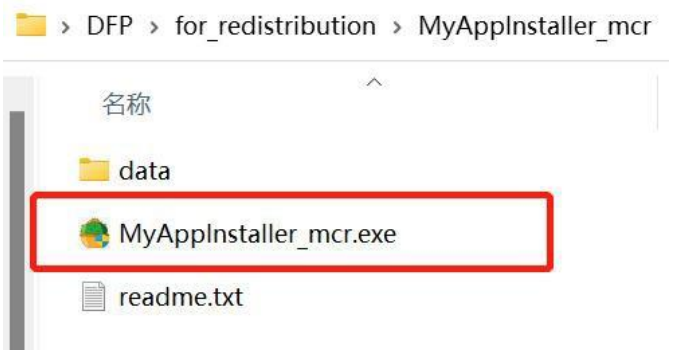
Microsoft Windows 10, 64 bit version and above

4 Software installation steps

① Open the DFP installation package, select the first folder "for_redistribution", and unzip the folder "MyAppInstaller_mcr.zip".

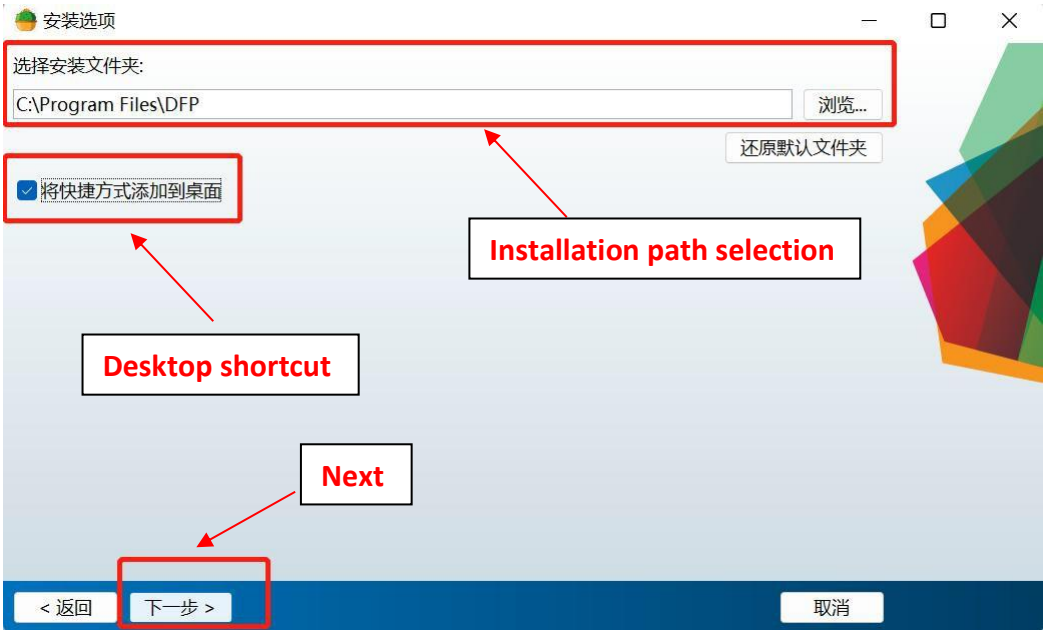


② Enter the decompressed folder "MyAppInstaller_mcr", double-click "MyAppInstaller_mcr.exe" to install the software operating environment, the installation starts slowly, wait patiently after double-clicking.



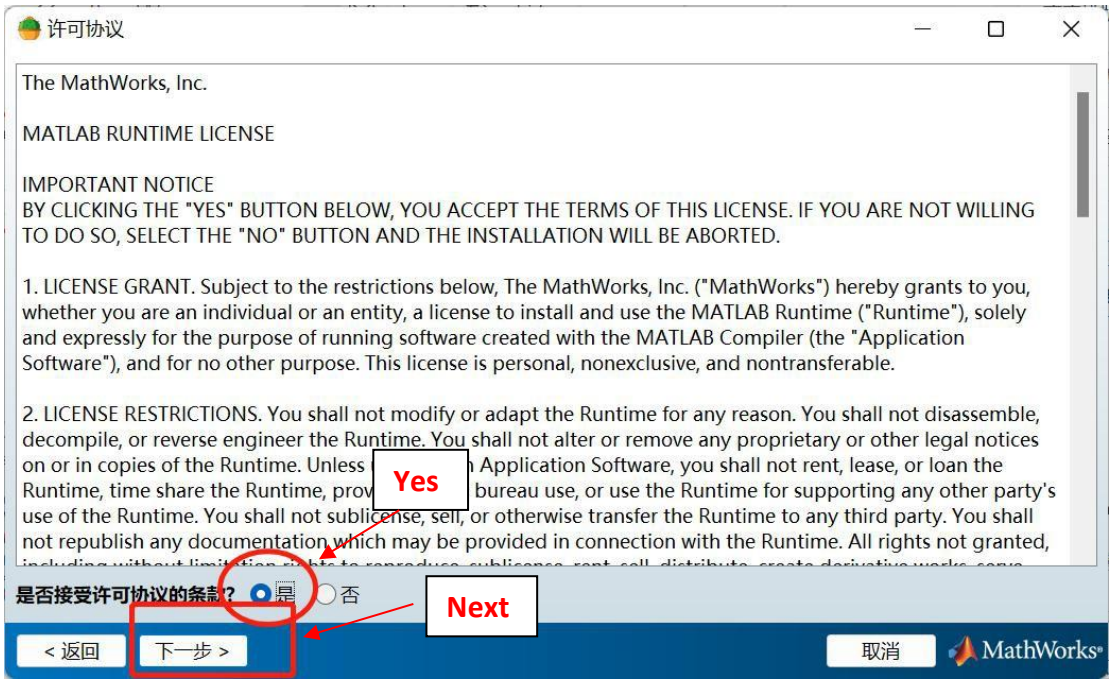
③ After double-clicking, the following interface appears, click "Next".

④ Select the installation folder, this is the installation path of your DFP software, just select the installation path according to your own situation, check the shortcut, or not, and then click "Next”.

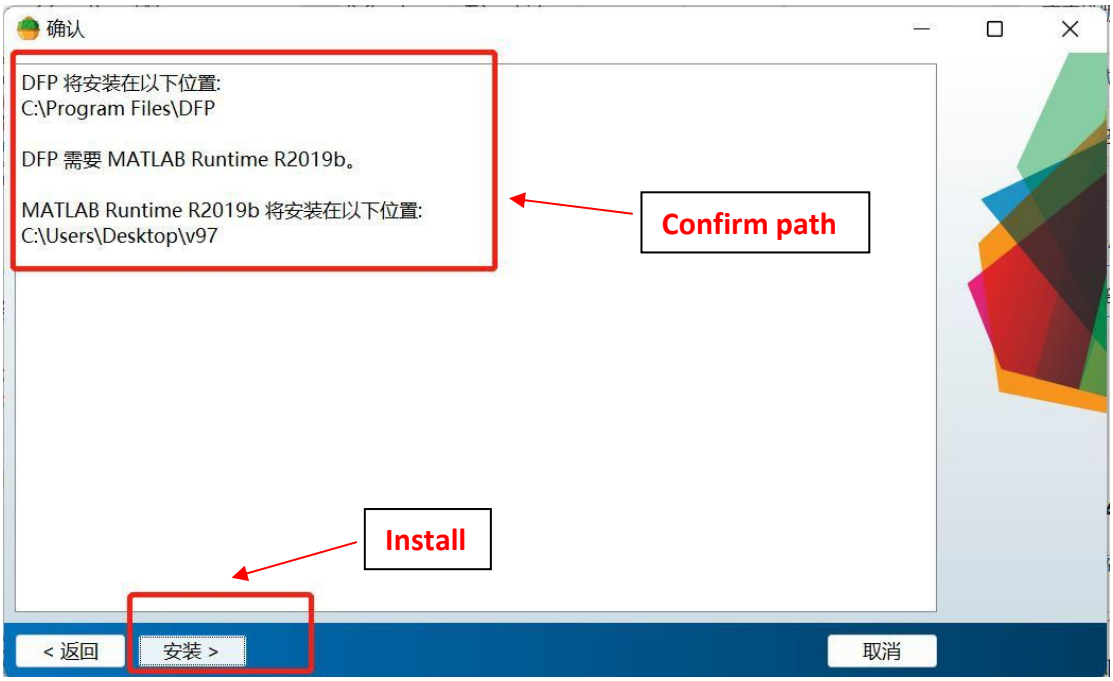


⑤DFP needs MATLAB Runtime as the running environment, select the MATLAB Runtime path and installation location, and then click "Next”.

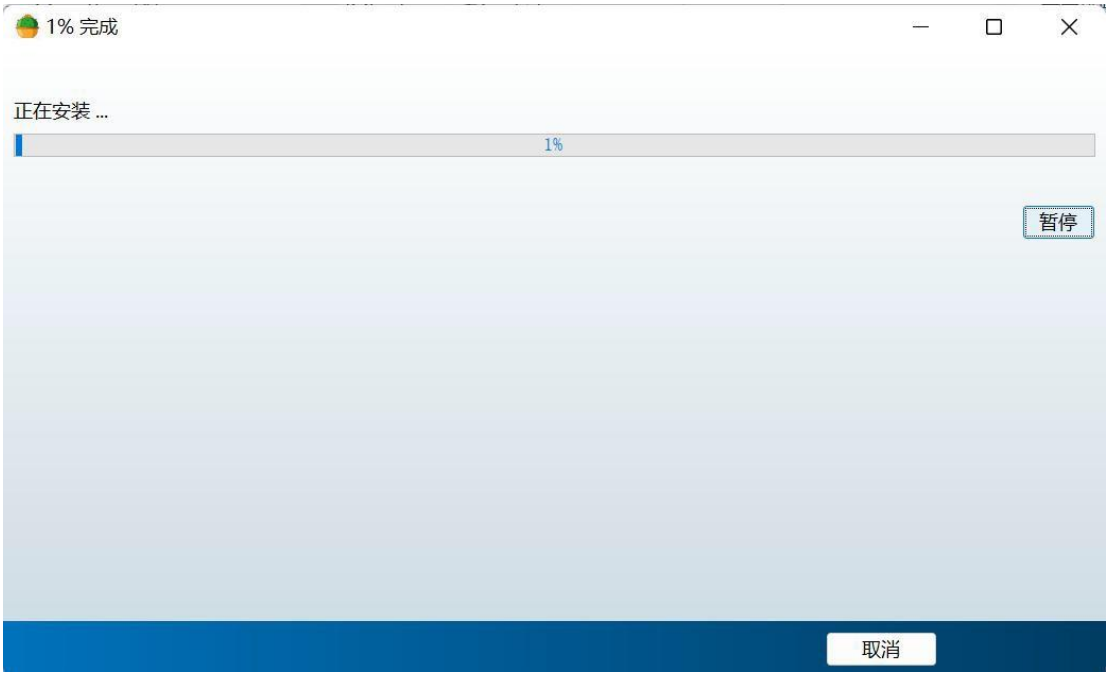
⑥Accept the license agreement, select "Yes", and click "Next".



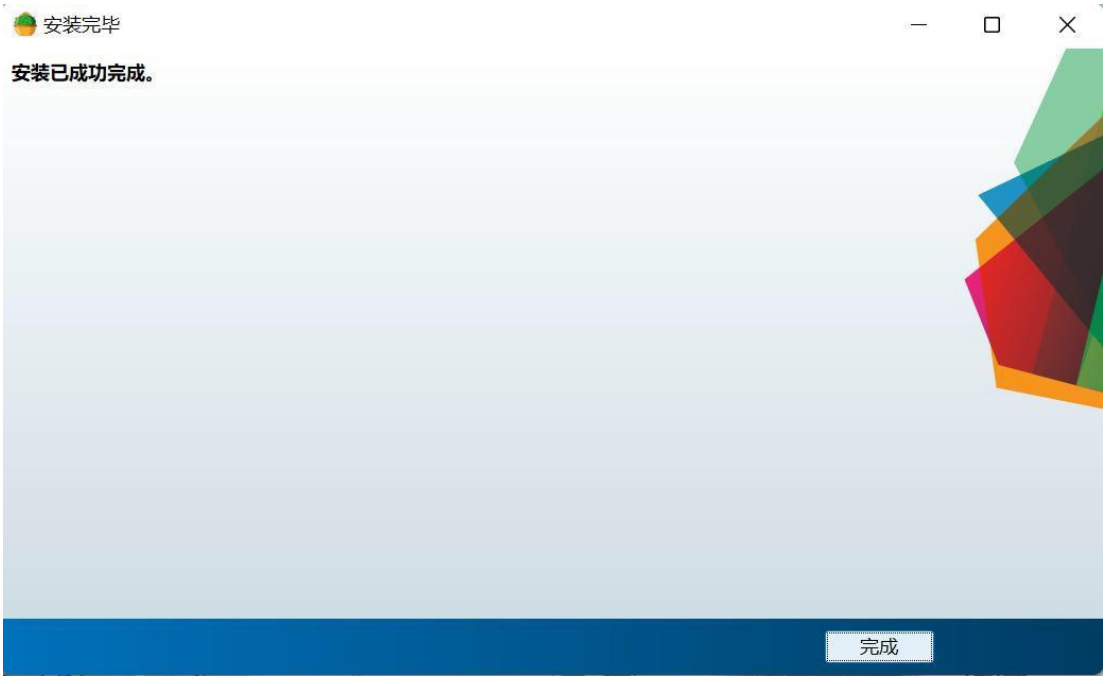
⑦Reconfirm the installation path of the DFP software and the operating environment, which can be the same path or different paths. If there is any problem, can choose "Return" to modify it. After confirming that the path is correct, click "Install".



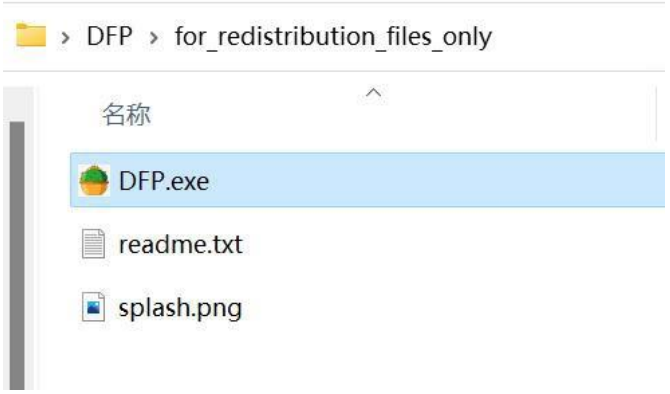
⑧ During the installation, wait patiently, the installation time may be long.



⑨ The installation is complete.



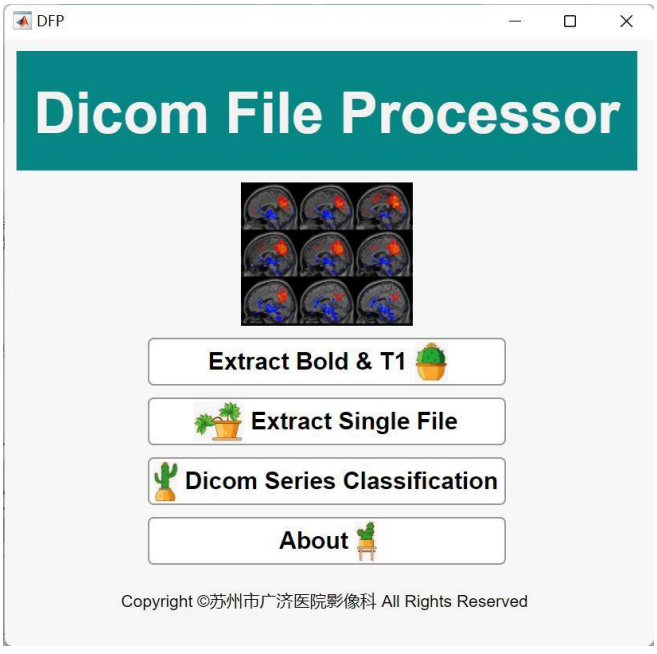
⑩ Double-click the software icon to open the software. If the shortcut is not checked in step ④ to add to the desktop, can select "DFP.exe" in the "for_redistribution_files_only" folder, right-click to select the shortcut to send to the desktop, or directly copy the Drag it to the desktop, double-click to open the software, and the software installation is complete.



5 DFP interface operation

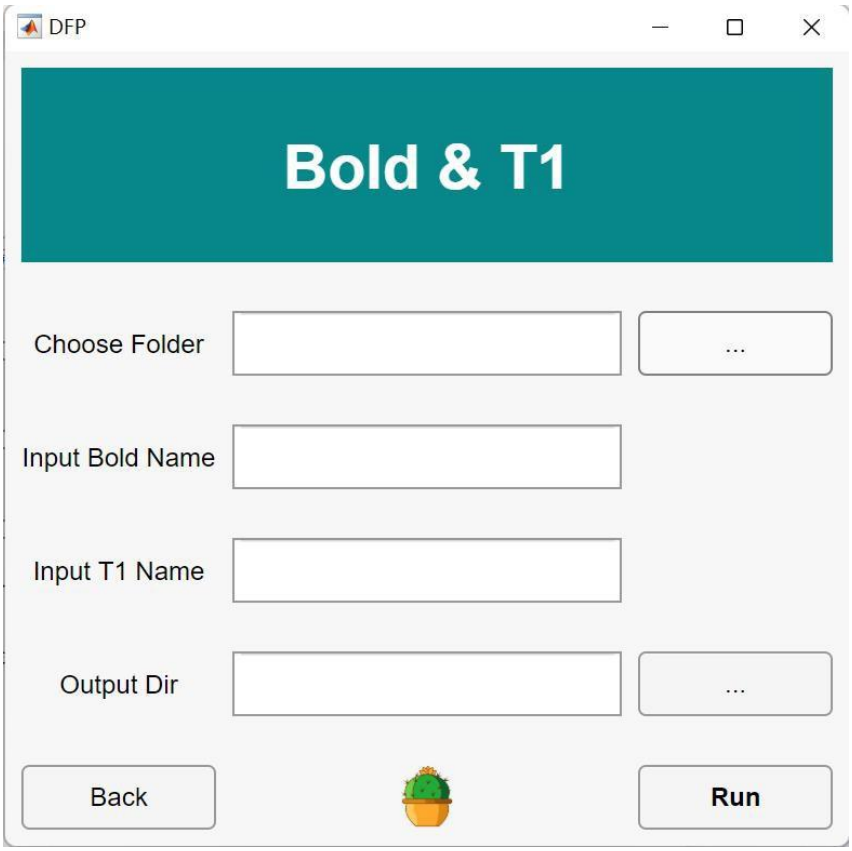
5.1 Main interface

The main interface of the software is divided into four sections, the section for extracting Bold and T1 files at the same time: Extract Bold & T1; the section for extracting a single specified file: Extract Single File; the section for sorting and archiving Dicom images: Dicom Series Classification; and the section about the software : About;

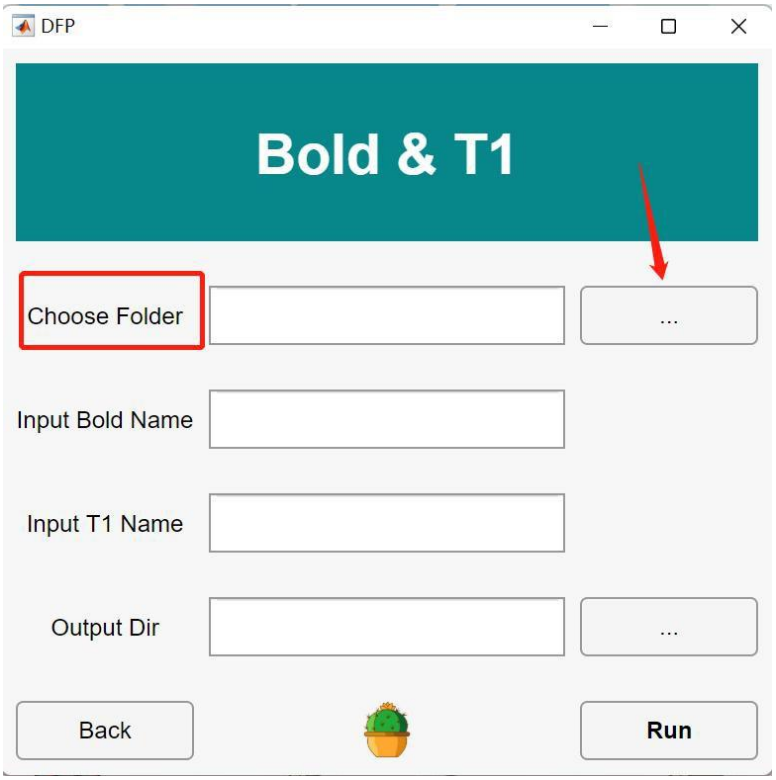


5.2 Extract Bold & T1 interface

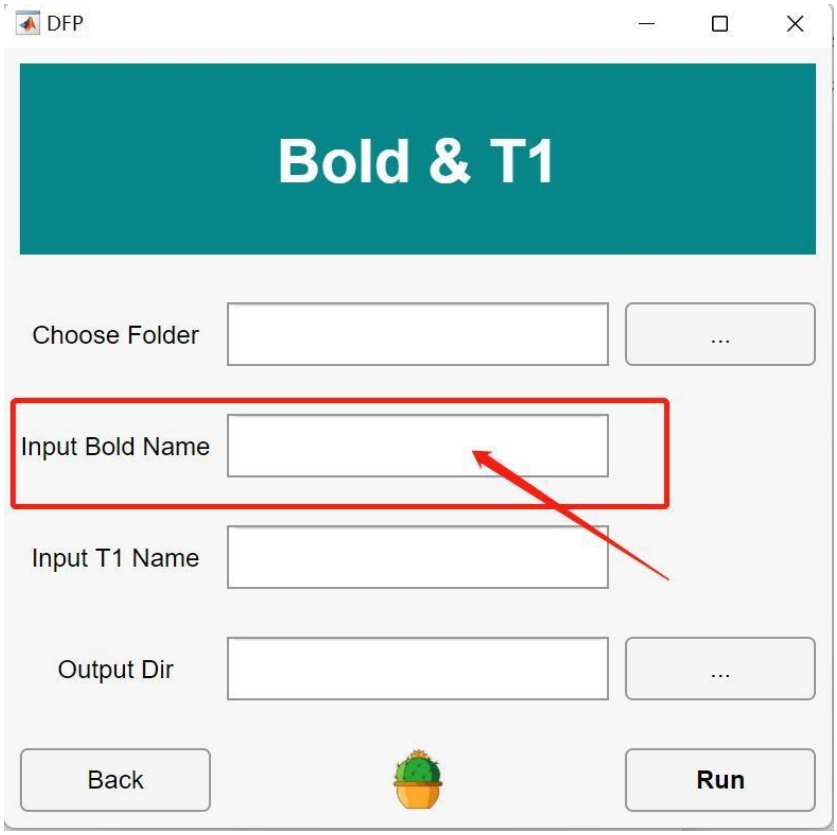
Click "Extract Bold & T1" to enter the "Bold & T1" operation interface.



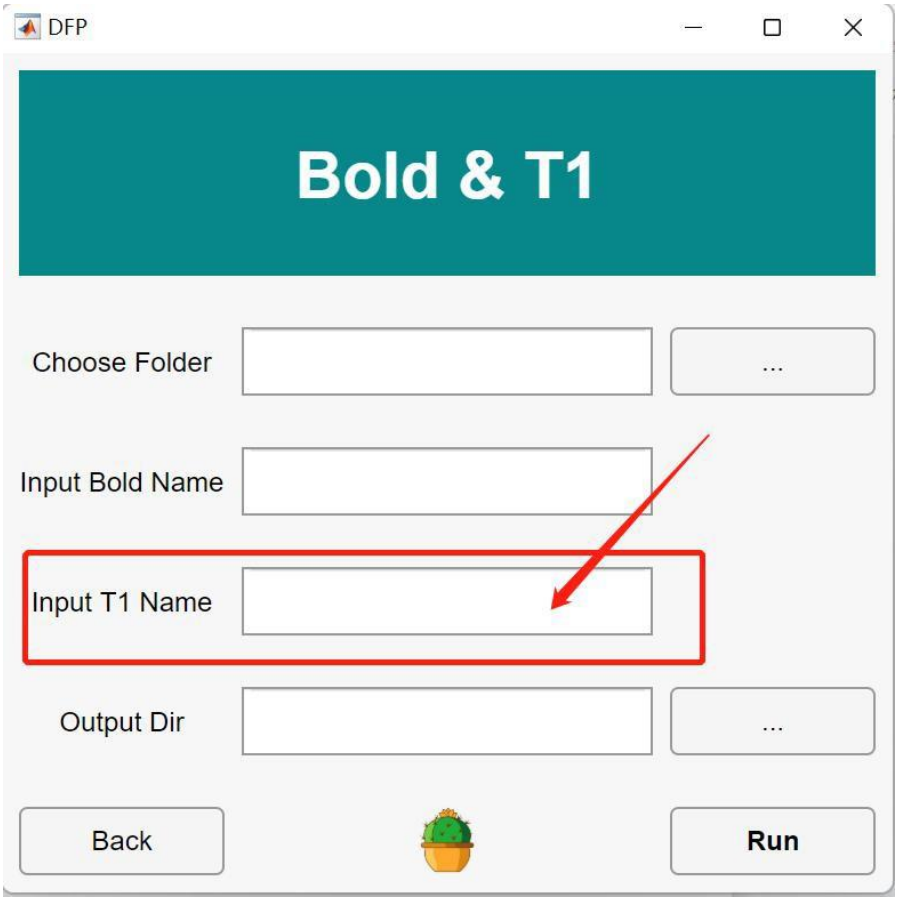
①Click this button to select the total data folder you want to organize (note: if there is only one case of data to be extracted, you also need to put it into a folder, select this folder instead of directly selecting that case of data).



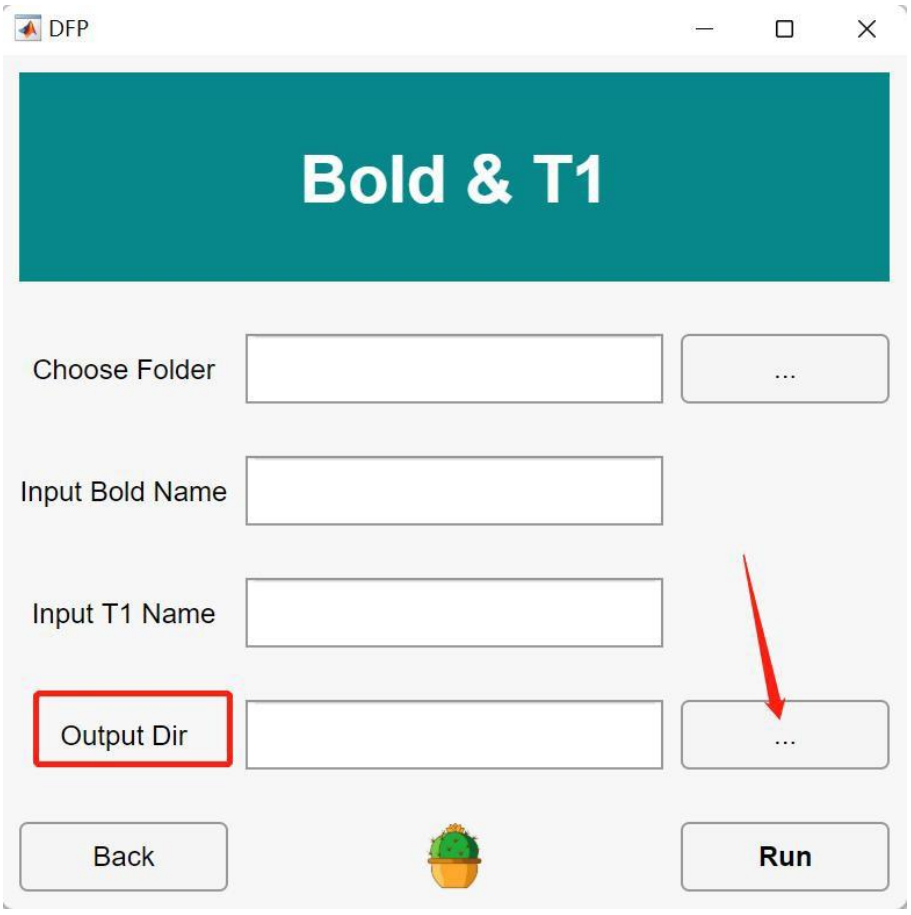
②Enter the folder name of the Bold sequence to be extracted here, case-sensitive, and enter the full name of the folder (Note: If you do not enter the full name of the folder to be extracted, it will be filtered according to the entered information, if there are more folders that meet the conditions, and there are numbers in the folder name, the number with the largest number will be selected for extraction according to the number in the folder name; if there is no number and you cannot select a single one, it will prompt "failure").



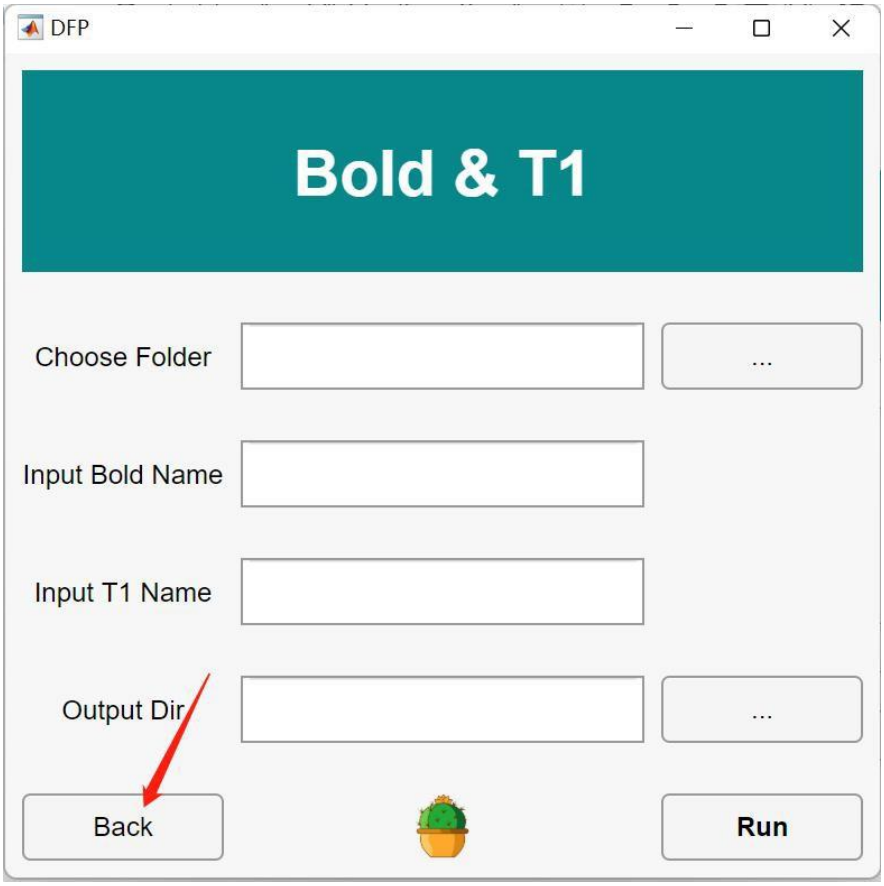
③Enter the folder name of the T1 sequence to be extracted here, case-sensitive, and enter the full name of the folder (Note: If you do not enter the full name of the folder to be extracted, it will be filtered according to the entered information, if there are more folders that meet the conditions, and there are numbers in the folder name, the number with the largest number will be selected for extraction according to the number in the folder name; if there is no number and you cannot select a single one, it will prompt "failure").



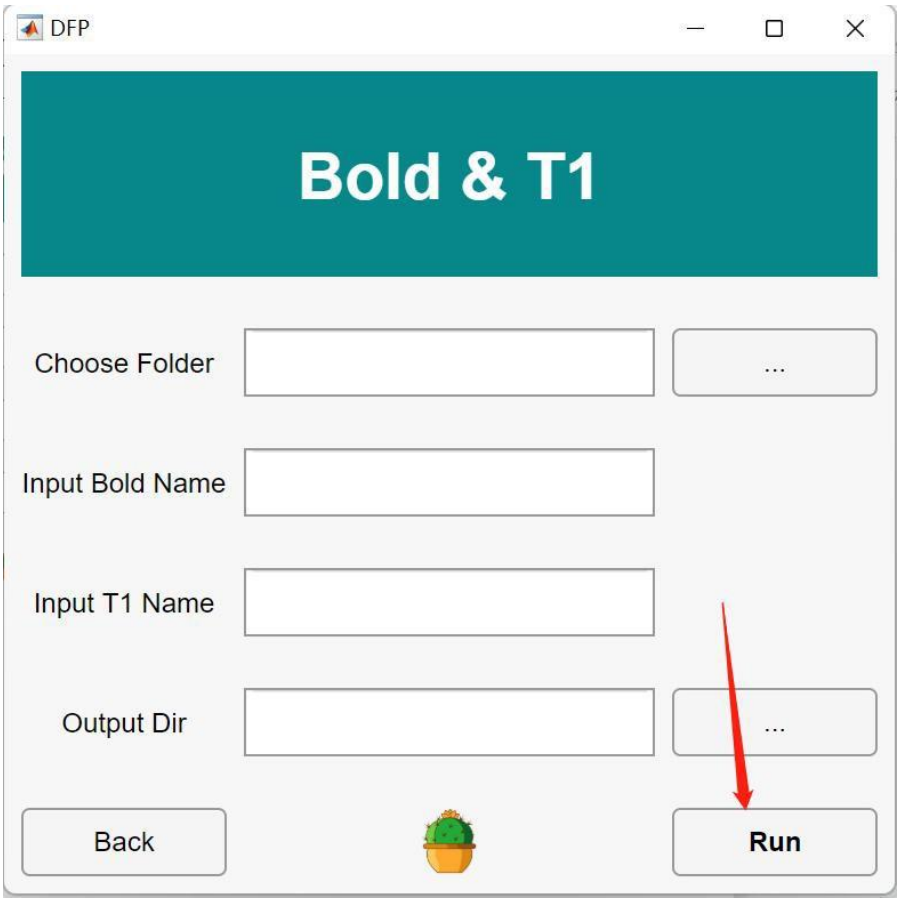
④Click this button to select the output path of the extracted folder.



⑤Click the "Back" button to return to the main interface;

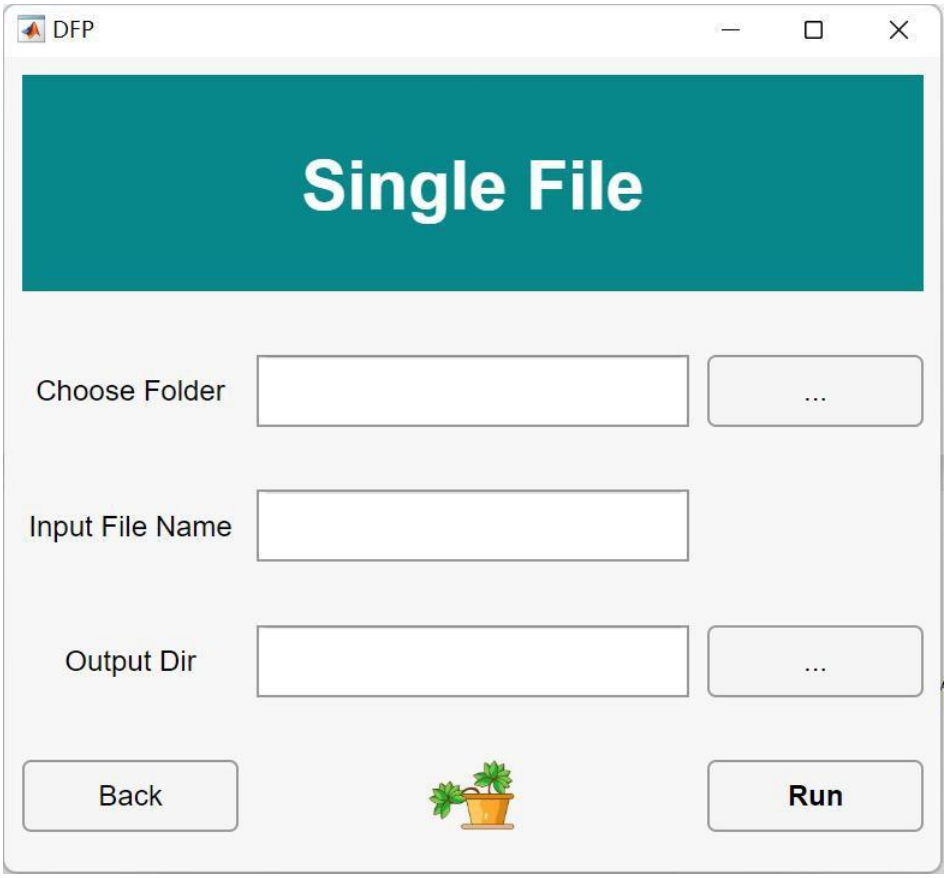


⑥After the information input is complete, click "Run" to start data processing.

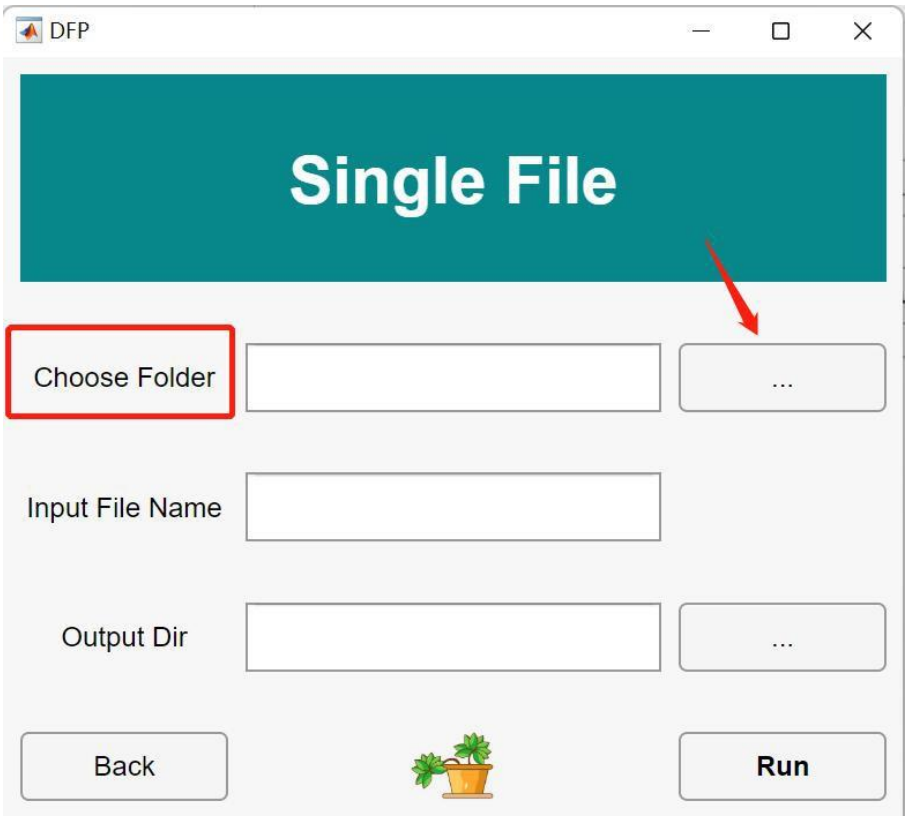


5.3 Extract Single File interface

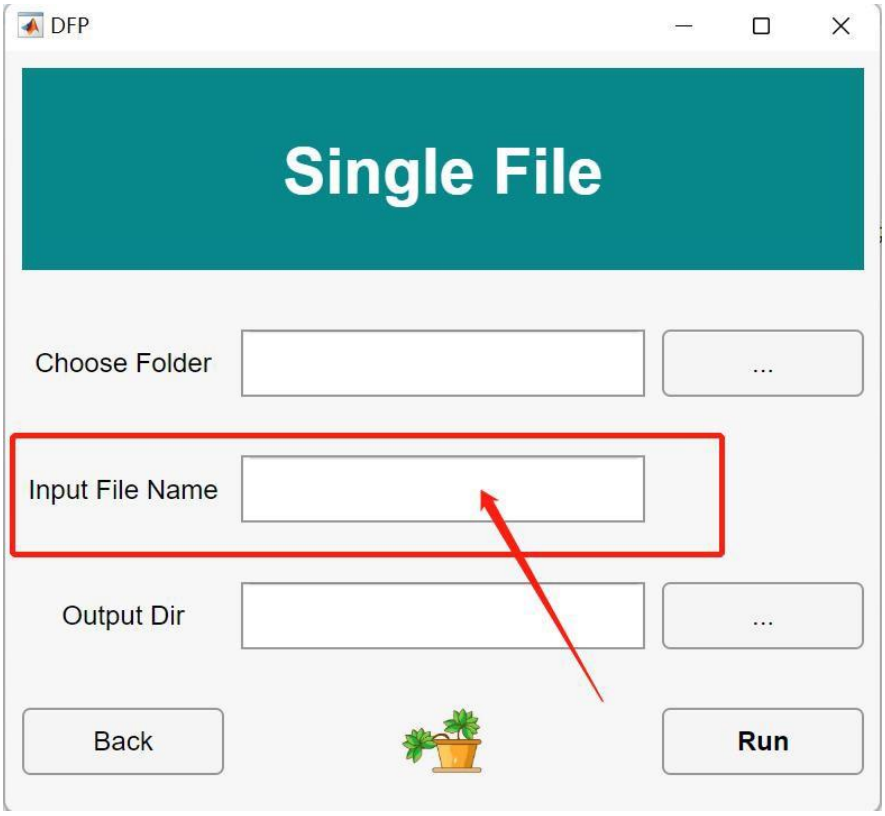
Click "Extract Single File" to enter the "Single File" operation interface.



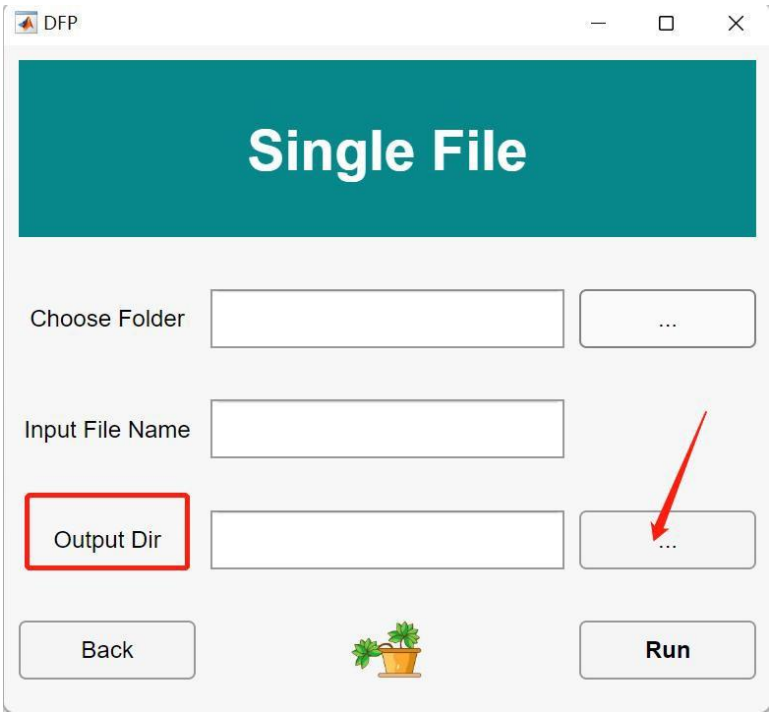
①Click this button to select the total data folder you want to organize (note: if there is only one case of data to be extracted, you also need to put it into a folder, select this folder instead of directly selecting that case of data).



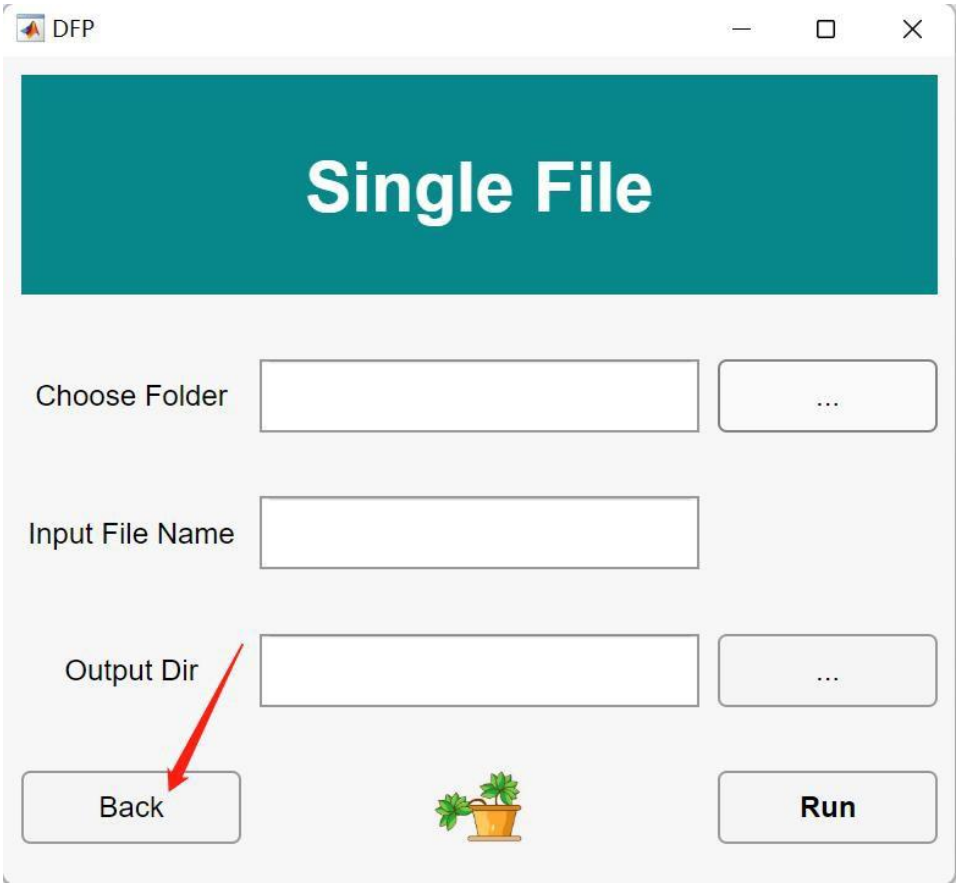
②Enter the name of the folder to be extracted here, case-sensitive, and enter the full name of the folder (Note: If you do not enter the full name of the folder to be extracted, it will be filtered according to the entered information, if more than one meets the conditions If there is a number in the folder name, the folder with the largest number will be selected for extraction according to the number in the folder name; if there is no number and cannot be selected individually, it will prompt "failure").



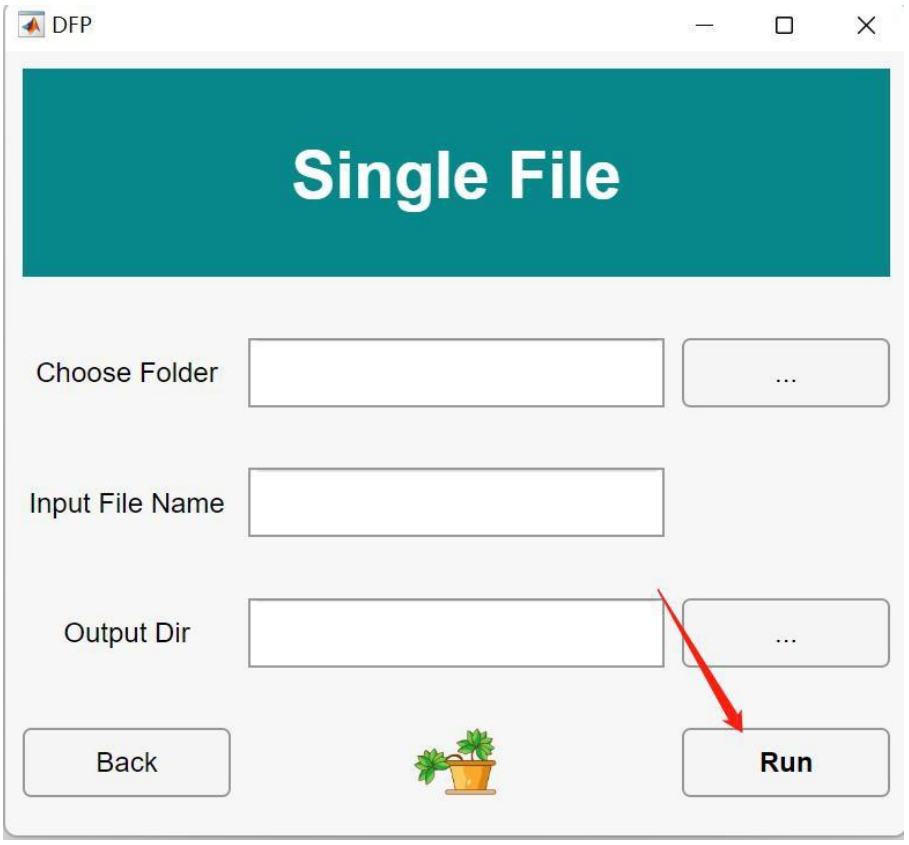
③ Click this button to select the output path of the extracted folder.



④Click the "Back" button to return to the main interface.

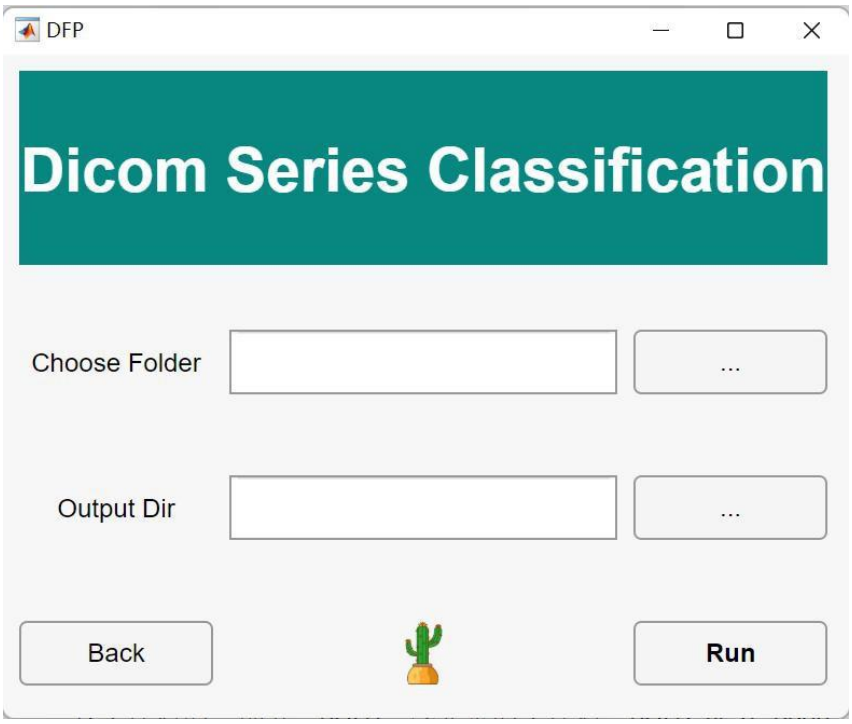


⑤After the information input is complete, click "Run" to start data processing.

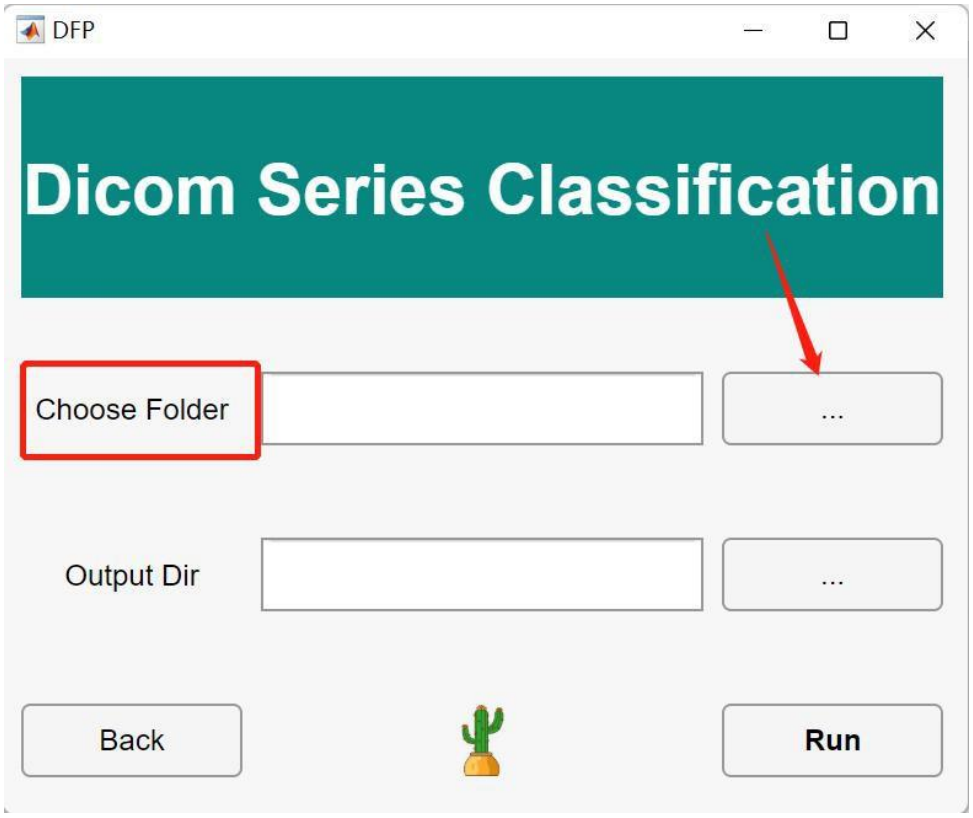


5.4 Dicom Series Classification interface

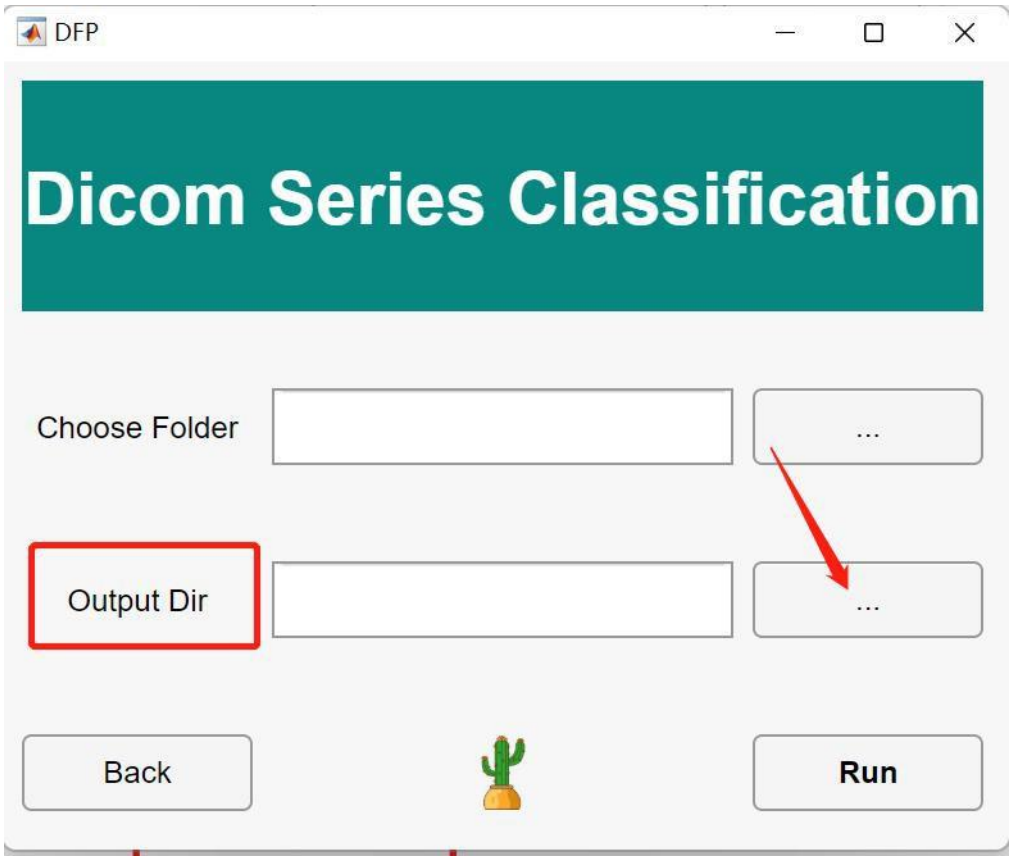
Click "Dicom Series Classification" to enter the "Dicom Series Classification" operation interface.



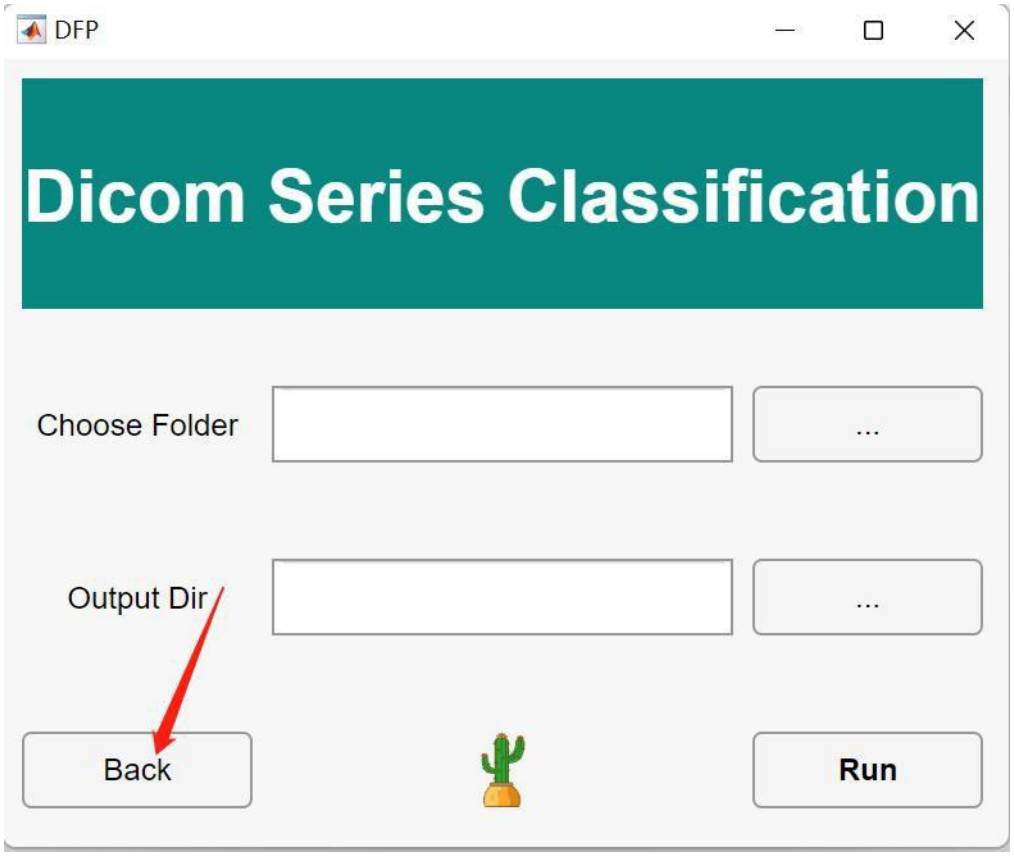
① Click this button to select the total folder of data that needs to be sorted (note: if only one case of data needs to be extracted, it also needs to be put into a folder, select this folder instead of directly selecting that case of data)



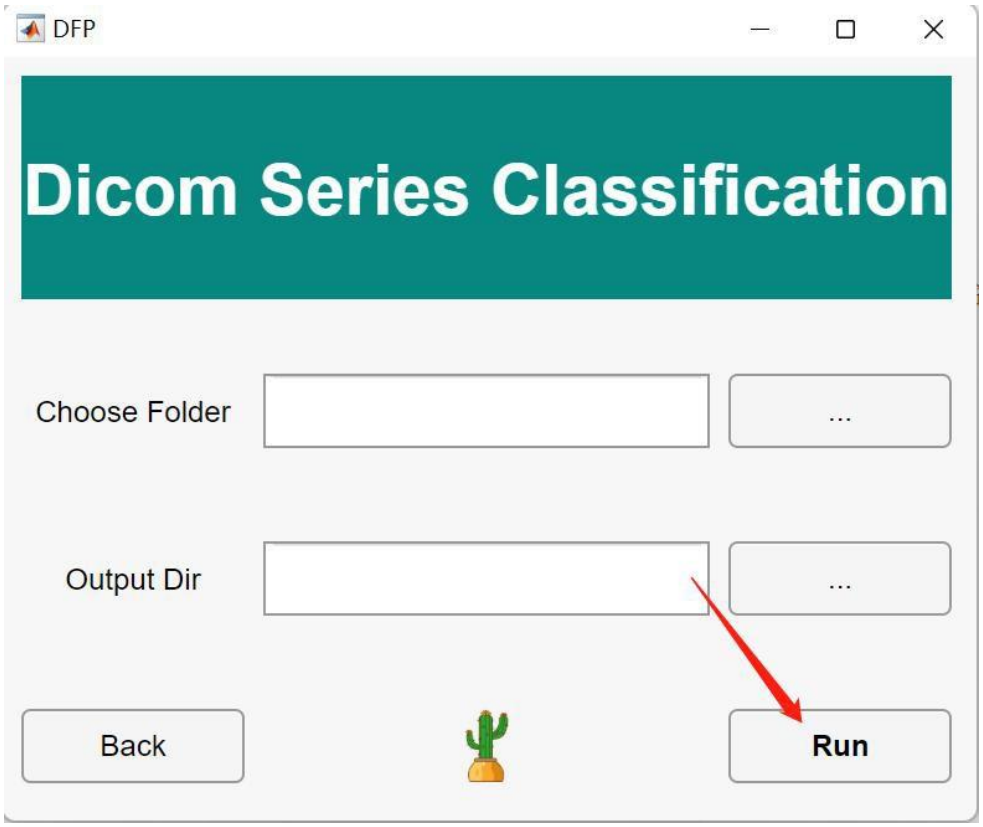
②Click this button to select the output path of the sequence folder.



③Click the "Back" button to return to the main interface.



④After selecting the folder to be sorted and the output path, click "Run" to start data processing.



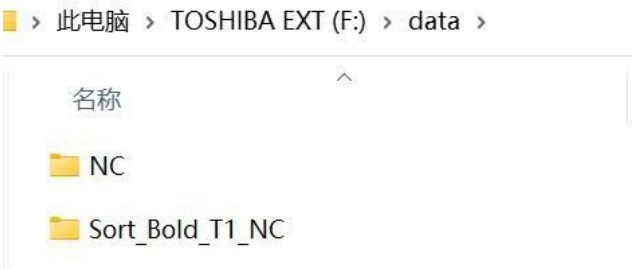
5.5 About interface

This section introduces the software version and the contact information when there is a problem with the software.

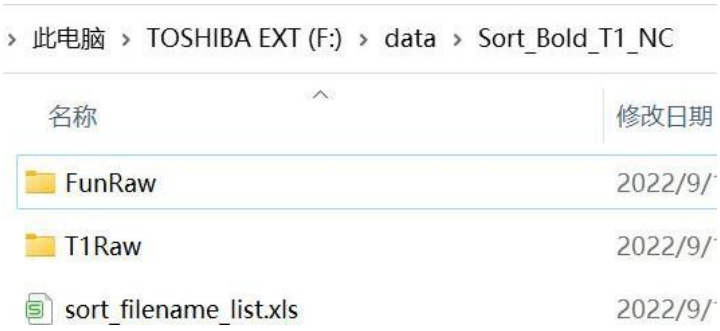
6 Running results

6.1 Extract Bold & T1 Running results

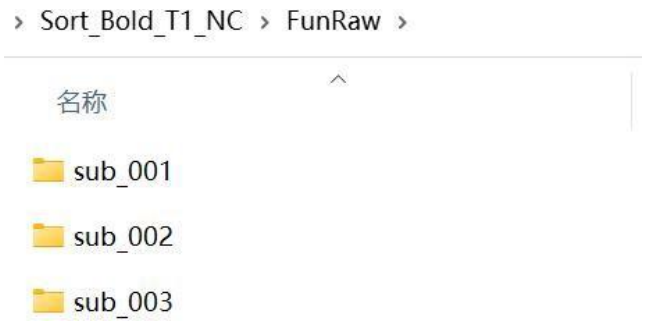
①Under the specified output path, a new folder is generated, and the new folder name is composed of the selected total folder name plus the prefix "sort_Bold_T1_".



②Then create two new folders under the general folder, which are FunRaw and T1Raw folders respectively. The FunRaw folder stores the Bold function image, and the T1Raw folder stores the T1 structure image, and generates an excel sheet of the file sorting results, named for "sort_filename_list".



③Extract the Bold and T1 data of each case of data in a set of data, and rename them in sequence according to the order of "sub_001, sub-002, sub_003", the new folder names of the same case of data Bold and T1 are the same, and then save them in the FunRaw and T1Raw folders respectively.



④During the data processing, if a case of data extraction fails and its Bold or T1 data cannot be extracted, an error pop-up window will appear.

⑤After all the data processing in the folder is completed, the following pop-up window will appear as a reminder.

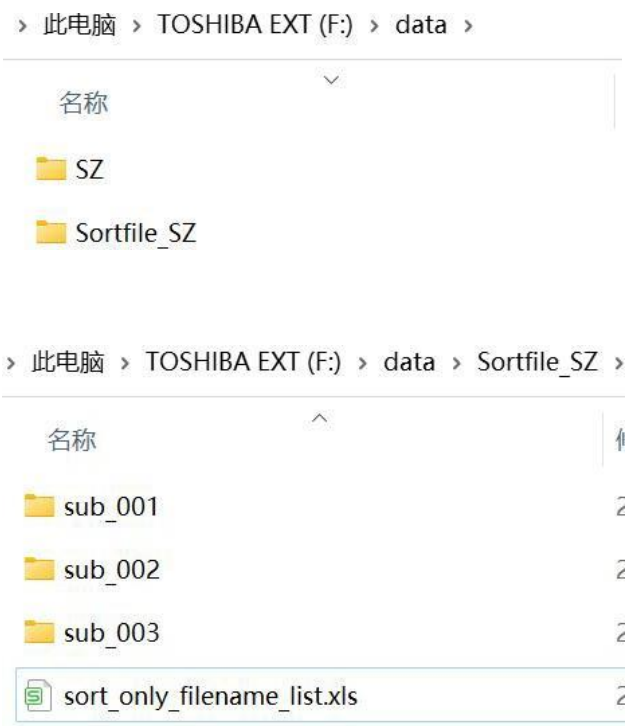
⑥The generated excel table contains the original and new names of the data, as well as the specific processing conditions of Bold and T1 data, to check the specific information of the extracted data. If the sorting status appears "failure", you can check the data content according to the original name of the data. The specific case of Bold or T1.

A	B	C	D
新的序号名	已整理好的数据原名称	Bold数据整理状态	T1数据整理状态
sub_001	zhangsan	成功	成功
sub_002	lisi	成功	成功
sub_003	wang	成功	成功

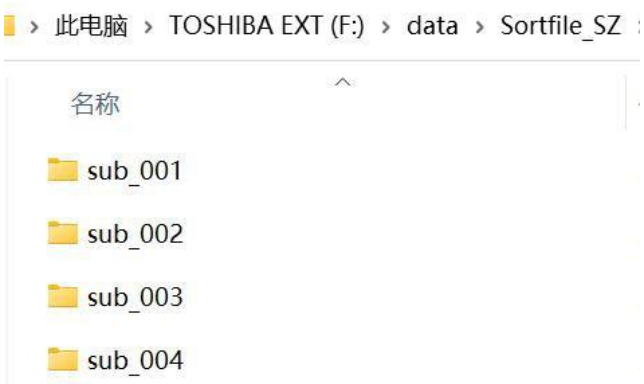
A	B	C	D
新的序号名	已整理好的数据原名称	Bold数据整理状态	T1数据整理状态
sub_001	zhangsan	成功	成功
sub_002	lisi	失败	成功
sub_003	lisan	成功	失败

6.2 Extract Single File Running results

①Under the specified output path, a new folder is generated. The new folder name is composed of the selected total folder name plus the prefix "Sortfile_", and an excel table is generated in the folder, named "sort_only_filename_list".



②According to actual needs, the name of the file or folder to be extracted is input in the interface, extracted in batches from a set of data, and renamed in order according to the sorting methods such as "sub_001, sub_-002, sub_003...".



③During the data processing, if the extraction of the specified folder of a certain case of data fails, an error prompt pop-up window will appear, which will be reflected in the generated excel form:

④After all the data processing in the folder is completed, the following pop-up window will appear as a reminder.

⑤The generated excel table contains the original and new names of the data files, as well as the extraction status of the specified folder, to check the specific information of the extracted files. If the sorting status appears "failed (**失败**)", you can check the details of the files according to the original name of the data Condition.

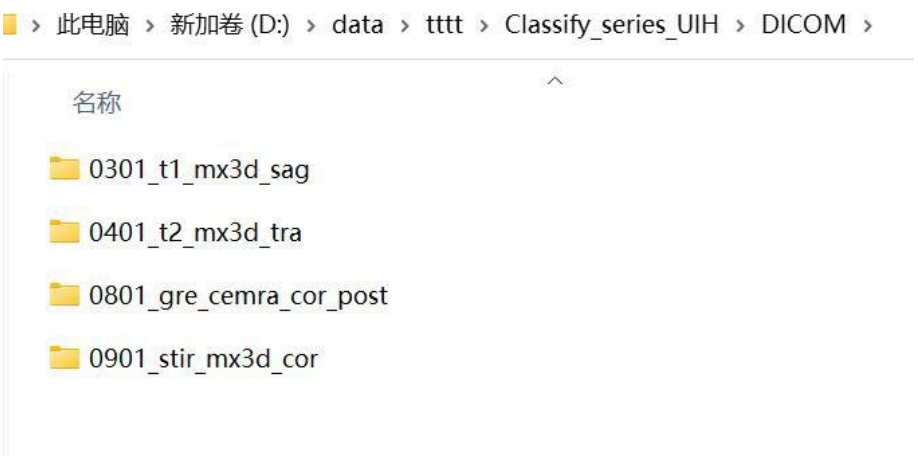
新的序号名	已整理好的数据原名称	数据整理状态
sub_001	zhangsan	成功
sub_002	lisi	成功
sub_003	xiaoming	成功

新的序号名	已整理好的数据原名称	数据整理状态
sub_001	zhangsan	成功
sub_002	lisi	失败
sub_003	xiaoming	成功

6.3 Dicom Series Classification Running Result

①Under the specified output path, a new folder is generated. The new folder name is based on the selected total folder name plus the prefix "Classify_series_". All sorted sequence folders will be stored in this newly created total file folder.

②Classify the Dicom images without distinguishing sequence according to their sequence information, and store them in the corresponding sequence folder;



③After all the Dicom data are sorted according to the scanning sequence, the following pop-up window will appear as a reminder.