Napari-Assistant Dask Plugin User Guide

Version 0.0.1

Artem Tomilo

Nafisa Anjum

Himanshu Manoj Kaloni

September 24, 2023

Contents

1 Introduction

In the dynamic realm of scientific image analysis, the demand for tools that embody efficiency, scalability, and accessibility has reached a paramount status. Within this context emerges the Napari-assistant dask plugin[1], representing an integration of notable significance. It empowers researchers and scientists with the capability to harness the computational prowess of Dask while operating within Napari—an open-source, versatile platform designed for the interactive exploration and analysis of biological images. The amalgamation of these two resources confers a multitude of advantages, not only streamlining the analytical workflow but also facilitating the handling of extensive and intricate data sets with remarkable ease. Dask, a flexible and high-performance parallel computing library, is renowned for its ability to seamlessly scale computation from a single machine to a cluster of servers. By marrying Napari, a cutting-edge tool for interactive exploration and visualization of multidimensional data, with the parallel computing capabilities of Dask, the Napari Assistant Dask Plugin opens up new horizons in the field of image analysis. This dynamic collaboration not only enhances the user experience but also significantly expands the scope and efficiency of image analysis workflows.
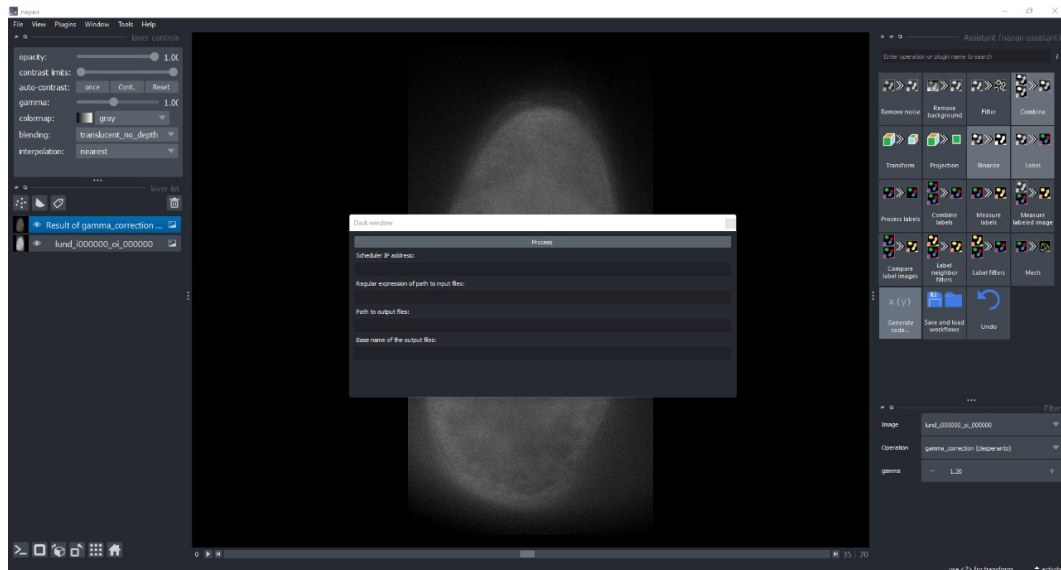

Figure 1: Napari-assistant dask plugin overview

---

2   Step-by-step instructions on how to use this plugin

It is recommended that this plugin be installed in a clean virtual environment using an environment manager like conda or venv with the file environment.yml. Here are the steps:

1. Download the file environment.yml from the repository, and save it to the same directory where you want to install the napari-assistant-dask plugin.

2. Create the environment using the following command:
   `conda create -y -n myenv -c conda-forge python=3.9`

3. Activate the environment by using the following command
   `conda activate myenv`

4. Install napari by using the following command `pip install devbio-napari`

5. Install `napari-assistant` plugin using the following command:
   `pip install napari-parallel`

6. Start the user interface with the following command:
   `napari`

7. Start Napari Assistant as a user interface from the menu Tools > Utilities > Assistant (na)
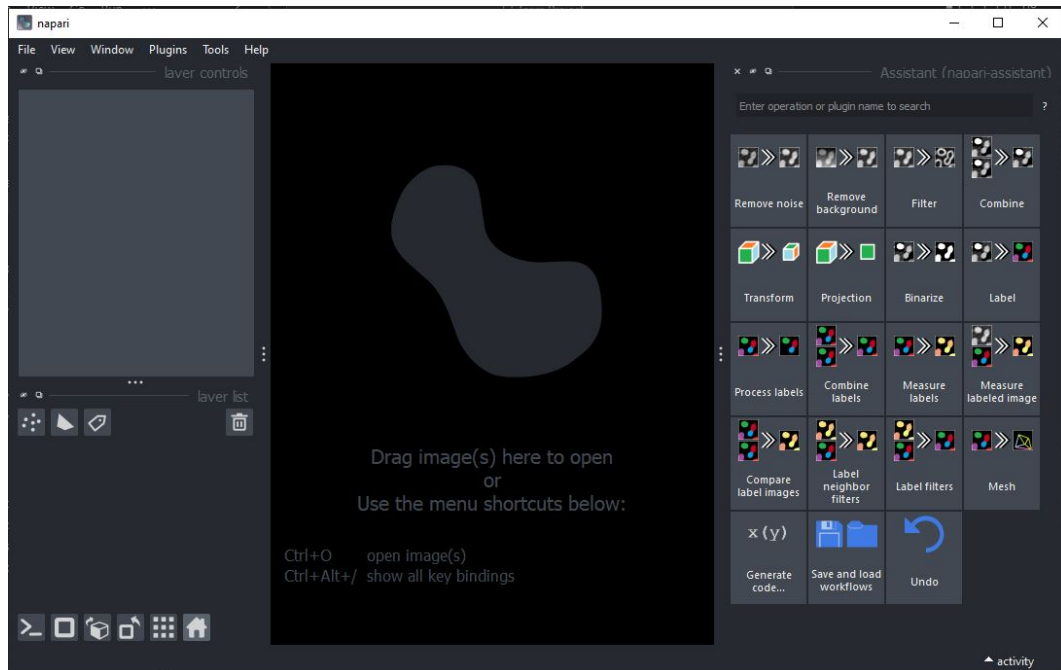
Figure 2: Napari-assistant overview

8. Perform your desired image processing workflow by loading one image from the file.

2.1 Dask Cluster setup

1. Install dask as `python -m pip install "dask[complete]"` (to install everything.)

2. Run the library scheduler using the command: `dask-scheduler`

3. Using IP-address from `dask-scheduler` command output (see Fig. 3), start the `dask-worker` using the command (see Fig. 4): `dask-worker <ip-address>`

2.2 Network File System (NFS)

This project assumes that the dataset would be stored in a central Network File System (NFS) which can be mounted on a Windows device. Generally

Figure 3: Scheduler

Figure 4: Worker

this process would be already done by the IT department. However if that is not the case, please follow the steps to create your own shared directory.

2.2.1 Getting NFS ready

We first have to turn on the NFS services in the Windows machine. This will let us create a sharable directory over the network which can be mounted onto different computers.

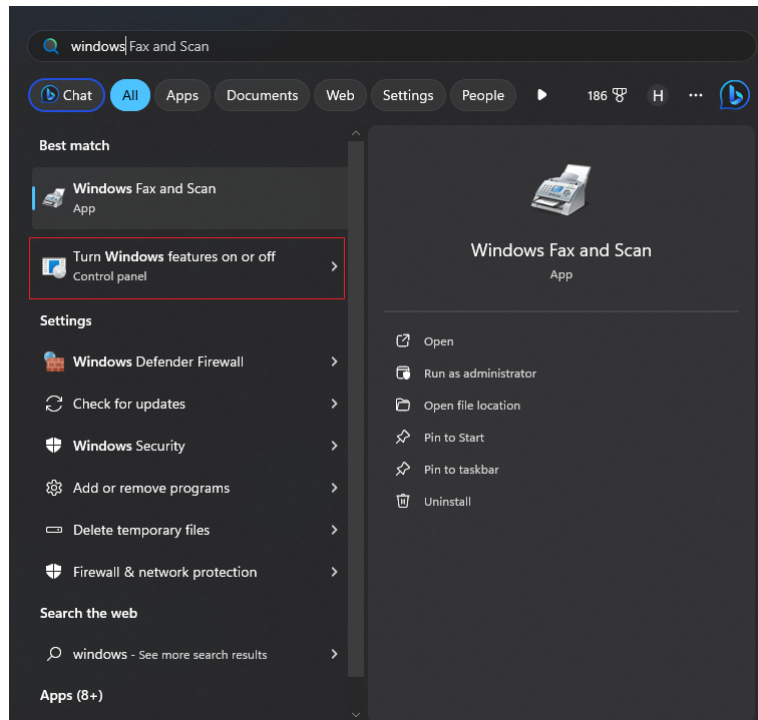1. Search for Turn Windows features on or off from the start menu.

Figure 5: Turn Windows features on or off
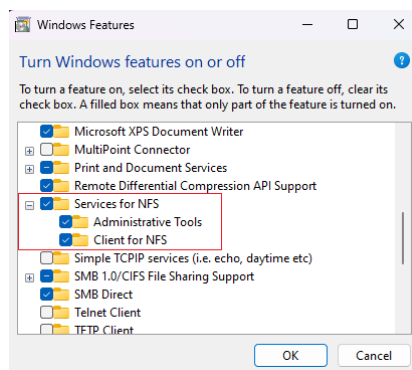
2. Find and select NFS services from the list.

Figure 6: Turn Windows features on or off

2.2.2 Sharing directory as NFS

After enabling the Services for NFS the next step is to actually create a shared directory. You can do it with the following steps.

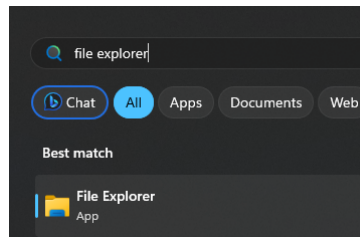1. Open your start menu and search for File Explorer.

Figure 7: File Explorer

2. Navigate to the location where you want to create the shared directory. (In the example we will create in the Desktop. The name of the shared directory is files.)
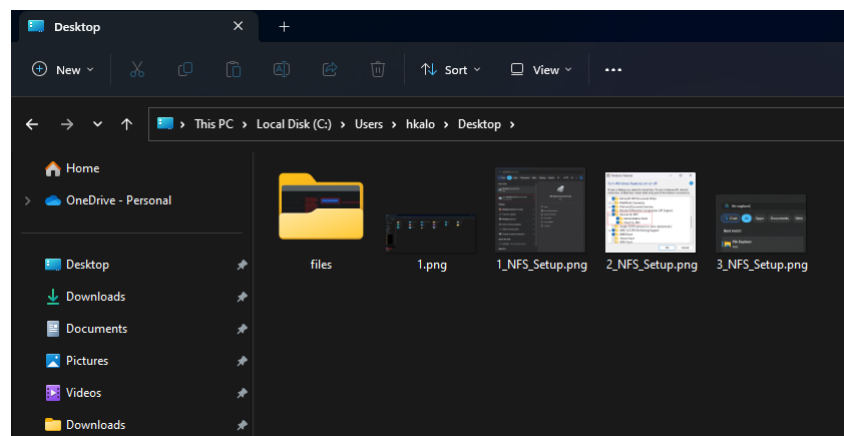

Figure 8: Create Directory for Sharing

3. Right click on the directory and select properties. (Alternatively you can select the directory and press Alt+Enter key.)
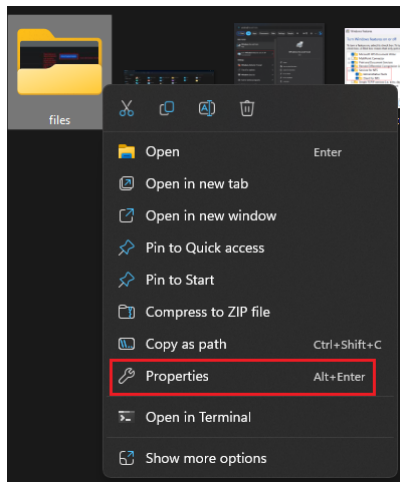
Figure 9: Right Click Folder Menu

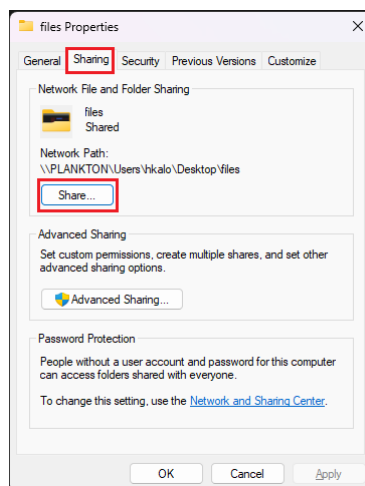4. Go to Sharing tab and then click Share.

Figure 10: Sharing Tab

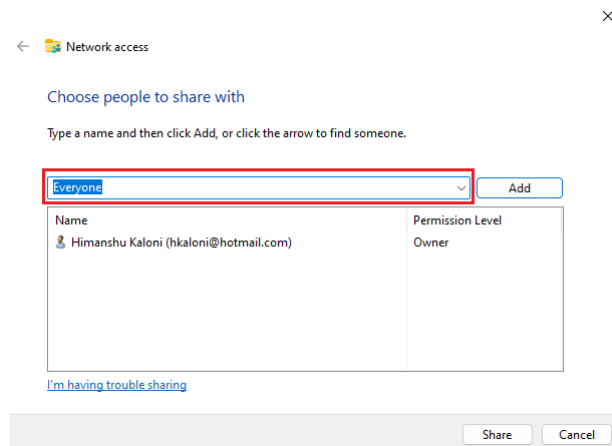5. From the drop-down menu select everyone and then click on Add button.

Figure 11: Network Access

6. Change the permission level from Read to Read/Write. We have to update it so that the dask worker is able to write the processed images back into the directory.
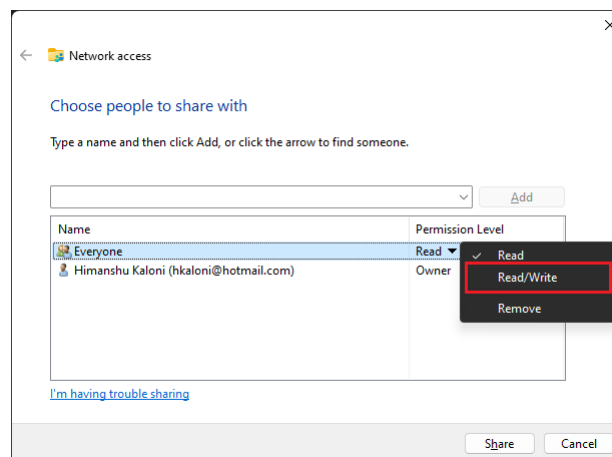

Figure 12: Network Access

7. Finally you can see this window which means the sharing is successful.
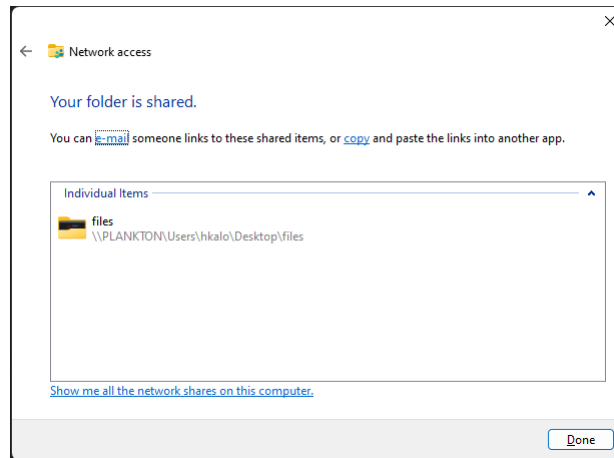
Figure 13: Network Access Success Window

2.2.3 Connecting to NFS

After creation of the NFS, it is necessary to connect it to the client system. Following are the steps to map the network drive:

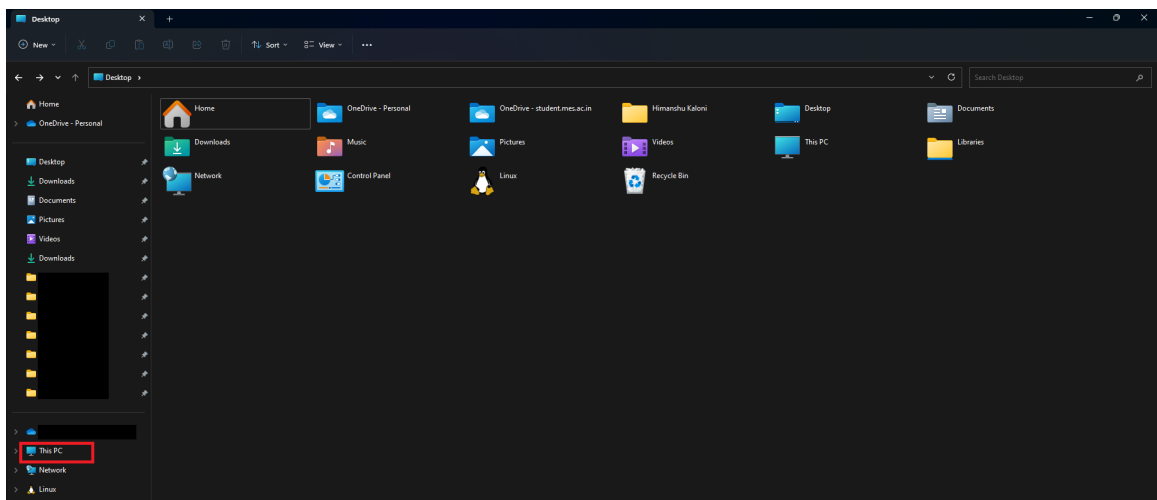1. Open the file explorer in your system. Find This PC on the left hand side of the file Explorer.

Figure 14: Open File Explorer and find This PC

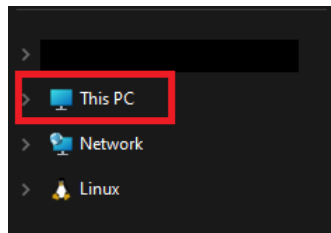2. Right Click on This PC.

Figure 15: Right click on This PC

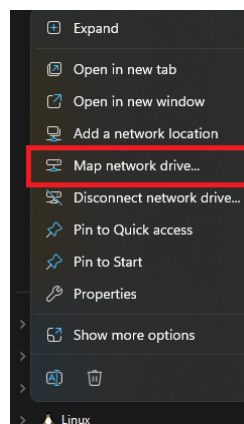3. Click on Map network drive to connect your system to the NFS location.

Figure 16: Select Map Network Drive from the menu

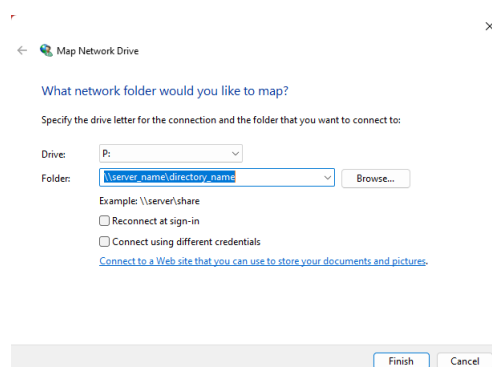4. Fill in the details and click Finish.

Figure 17: Enter the details of the network drive and the drive letter to mount on.

2.3 Process with Dask setup

1. After image processing with one image click on the button "Save and load workflow" and click "Process with Dask library".
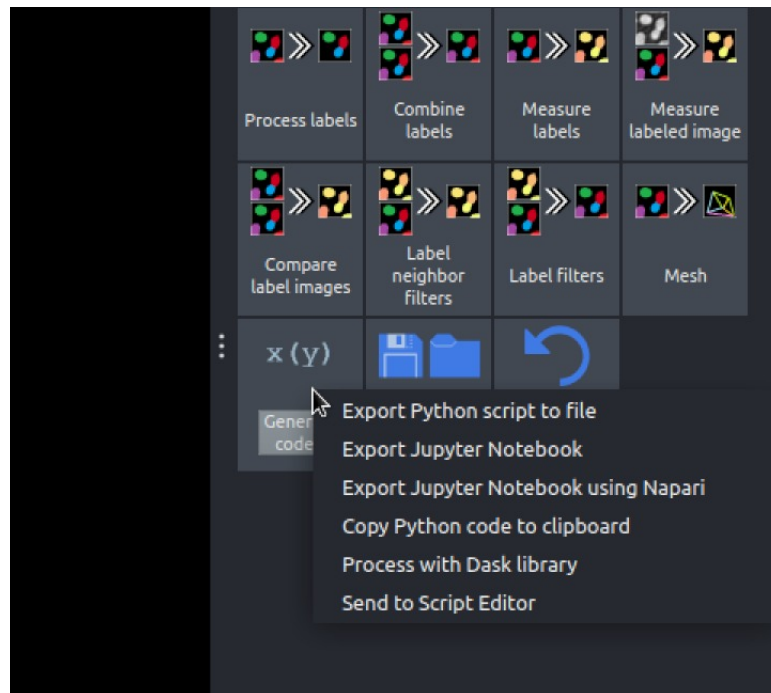

Figure 18: Process with dask

2. After clicking the following window will open: Where you have to fill in the following field:

   (a) Scheduler IP address: tcp://XXX.XXX.X.XXX:YYYYY

   (b) Regular expression of path to input files: After setting up the network file sharing system you will have the path to the folder.

   (c) Path to output files: Enter the path to the folder where you want to save the results.

   (d) Base name of the output files: Write the base name of the output files.

   (e) Extension for output files: Write one of the following extensions (".jpg", ".jpeg", ".png", ".gif", ".tif")

3. Click on the button "Process" to process the all images in the file with the dask cluster.
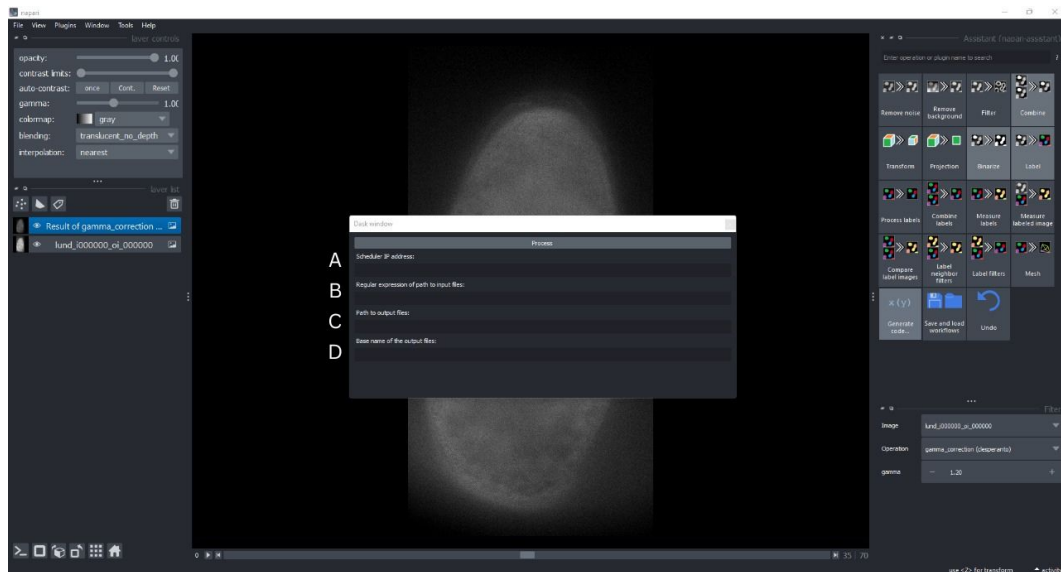
Figure 19: Process with Dask

4. After the process is done the following message will be there as shown in Figure 20.
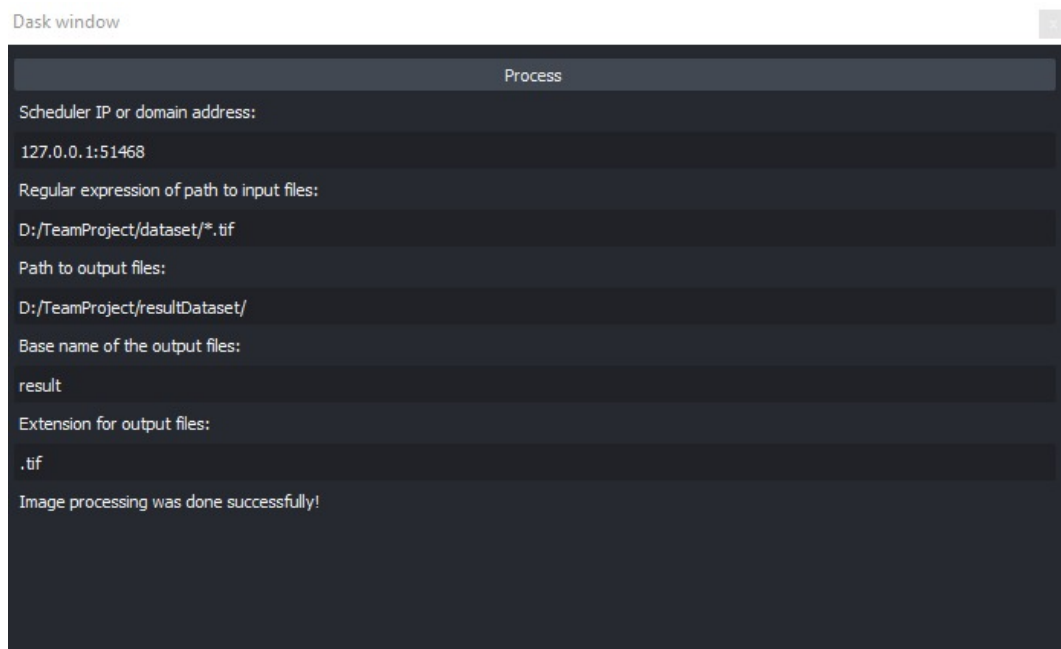

Figure 20: Output

2.4 Common errors while sharing directory or dask scheduler-worker connections

Sometimes due to network setting the shared directory may not be visible or dask scheduler is unable to connect with the workers and client machine. It mainly happens due to identification of the network as Private or Public. Private networks are the easiest to work with since the Network Discovery is enabled by default. If there are any issues then update the Network Setting by following the below steps.

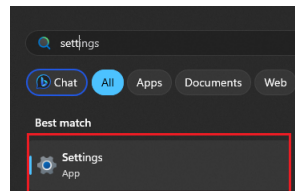1. Search Setting App from the Start Menu.

Figure 21: Setting App

2. Select Network & Internet from the menu on the left.
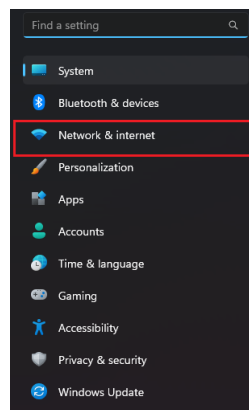

Figure 22: Setting App Menu

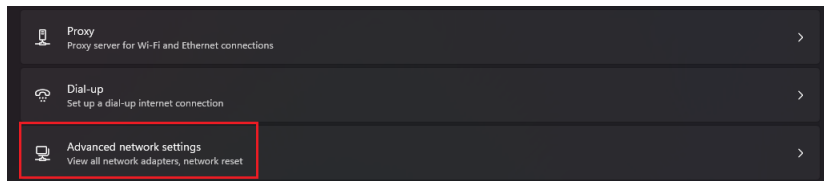3. Select Advanced Network Settings.

Figure 23: Advanced Network Setting
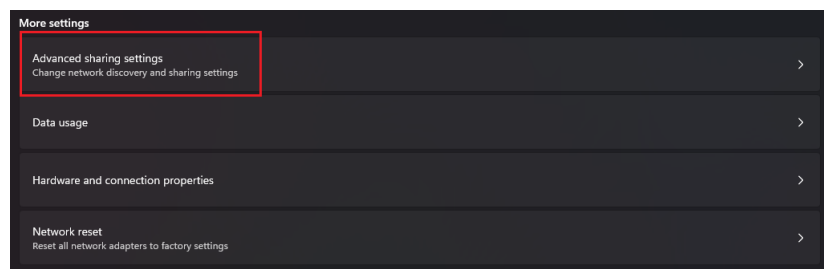
4. Go to Advanced Sharing Settings.

Figure 24: Advanced Network Setting menu

5. Make sure that the setting looks as in the figure below. Please note to update the setting for Private Network as shown in the figure.
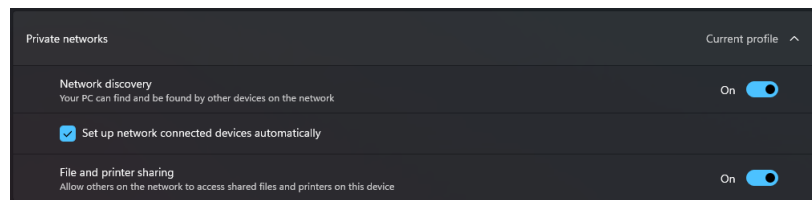

Figure 25: Advanced Sharing Setting

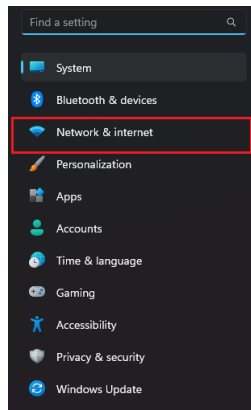6. Navigate back to Network & Internet from the menu on the left.

Figure 26: Setting App Menu
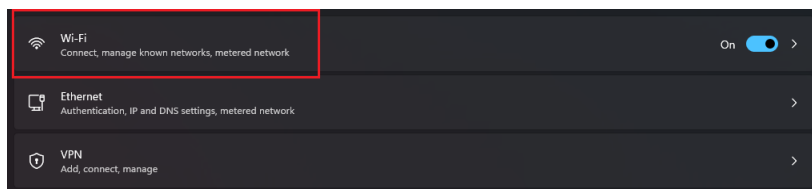
7. Select the Wifi setting.

Figure 27: Setting App Menu
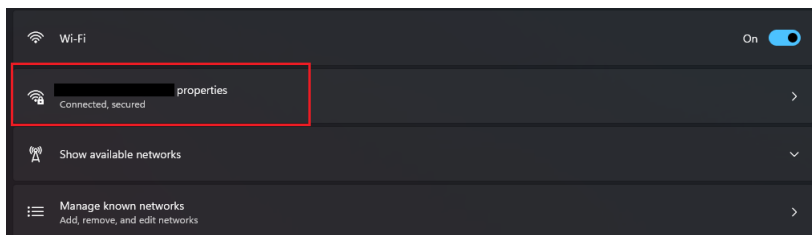
8. Go to the connected network.

Figure 28: Connected Network
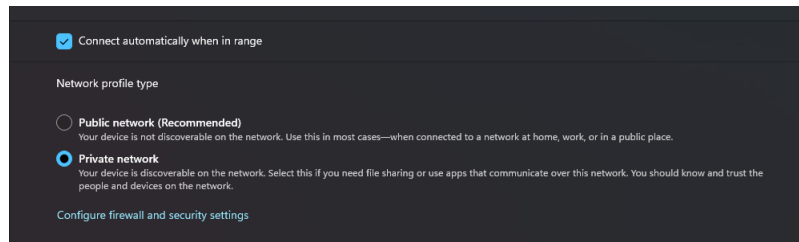
9. Make sure that the connected network profile is a Private Network.

Figure 29: Network Profile Type