

NeuroMorpho Access Tool: A Simple GUI for Accessing NeuroMorpho.Org Database

Ivan Zaletel (✉ ivan.zaletel@med.bg.ac.rs)

University of Belgrade

Marko Vidaković

Richard S. Nowakowski

Florida State University

Research Article

Keywords: Python, API, GUI, morphometry, NeuroMorpho.Org

Posted Date: April 7th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1515346/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

The use of open-access data has become increasingly prevalent in modern neuroscience research. These data are commonly used as a supplement to practical experimental research, but there is also an increasing body of literature which relies exclusively on the use of such publicly available data. Although freely available, these databases usually require an adequate level of training in bioinformatics and/or programming for a researcher to access them, which limits their usability. One of the oldest, most utilized and maintained databases of neuron and glia cell morphology is the NeuroMorpho.Org database, which can be accessed through the use of application programming interface (API).

In the present paper we introduce the NeuroMorpho Access Tool (NMAT), a simple and user-friendly graphical user interface (GUI) program, which allows easy and fast access to the NeuroMorpho.Org database to any neuroscientist, regardless of their previous bioinformatics knowledge. The core of NMAT was built in Python, utilizes the NeuroMorpho.Org API in order to access the database, and has a GUI built in Qt environment. The program is lightweight, consists of only one executable file and does not require any extensive accompanying instructional or technical documentation. The current version of the program enables the users to obtain cell reconstruction data together with morphometry data and also acquire images of cell reconstructions. The main advantage of NMAT is its multithreading capability, which allows the users to gain access to a large amount of data in a very short time period. Thus, NMAT may prove as a useful tool which would complement the highly utilized NeuroMorpho.Org database.

1 Introduction

Open-access databases can be used for various forms of neuroscientific research, such as morphometric analysis of nervous system cells (Ascoli et al., 2007), neuroimaging analysis of MRI datasets (Poldrack & Gorgolewski, 2014), genomic, transcriptomic and proteomic expression analysis (Keil et al., 2018; Pereira et al., 2014), etc. However, to adequately utilize the majority of open-access databases, a researcher needs to be proficient in handling a variety of software. The access to these databases is usually through the use of an Application Programming Interface (API), usually accessed with Python or R (Freeman, 2015). Python usage has seen a dramatic increase in the previous decade, as the scientific community has avidly accepted this programming language due to its high applicability, ease of use and modularity (Muller et al., 2015; Harris et al., 2020). Thus, training in bioinformatics and programming has now become a necessity for every life scientist to allow researchers to handle and analyze their data efficiently (Via et al., 2013; Moore et al., 2021). However, in low-income countries and countries with poor science funding it is very difficult to incorporate that kind of training at different stages of scientist research career (Moore et al., 2021). This is just one of multiple reasons why many open-access databases were and are still underutilized and reserved only for laboratories that have people with adequate programming expertise (Akil et al., 2011).

One of the oldest, most utilized and maintained databases of neuron and glia cell morphology is the NeuroMorpho.Org database (<http://neuromorpho.org/>), that currently contains cell reconstruction and

morphometric data for more than 170 000 nervous system cells from many species and collected by hundreds of labs (Ascoli, 2006; Ascoli et al., 2007). As NeuroMorpho.Org requires API access to obtain the cell morphometric data, its usefulness is limited to researchers with preexisting bioinformatics programming skills. Thus, we sought out to create a simple and user-friendly graphical user interface (GUI) program, which will allow easy and fast access to the NeuroMorpho.Org database to any neuroscientist, regardless of their bioinformatics knowledge. In the present paper we introduce NeuroMorpho Access Tool (NMAT), a GUI program. In its current form this program allows users to quickly download cell reconstruction data and morphometry data and acquire image files of reconstructed cells. The present paper describes the background of NMAT, its multithread architecture structure, shows options relevant to the functionality of the program and provides future perspectives on its development.

2 Methods And Results

21. NMAT background

NMAT (version 1.0) is a simple GUI interface primarily used to easily and rapidly access and obtain cell data stored in the NeuroMorpho.Org database. The NMAT program, together with its source code is freely available for download at GitHub (<https://github.com/izaletel/NeuroMorpho-Access-Tools>). Users can choose to download Windows or Linux compatible versions of the program. The source code is also available for advanced users who wish to modify the functionality of the program according to their own needs. The core of NMAT was built in Python and utilizes the NeuroMorpho.Org API to access the database. NMAT has a GUI built in Qt environment (PySide6, version 6.2.3, <https://pypi.org/project/PySide6/>). The entire NMAT program is lightweight, consists of one executable file and has no dependencies. It possesses a simple, friendly and intuitive interface and thus does not require any extensive accompanying instructional or technical documentation. Additionally, NMAT does not need to be installed on the computer and can very easily be transferred from one device to another, which allows easy access to the database from any computer.

2.2 Overview of NMAT interface

2.2.1 Querying and acquiring cell data

The current version of NMAT consists of a single 3-tab window, which allow users to obtain cell data and acquire cell image reconstructions (Figure 1). Each tab consists of different combo boxes and a text box, which gives user feedback information about the tasks that the user has given the program through its combo box.

The first tab called “Generate CSV” is the main component of the NMAT program and is used to generate a .csv file of cell reconstruction data and morphometry based on the user’s query input. Note that .csv

format was chosen because it can be easily imported into a spreadsheet, Python or R program for subsequent analysis. The NeuroMorpho.Org database allows the users to explore the database using multiple search criteria. Here we have decided to simplify the query process and thus have opted for three main criteria that the user can query: 1) Brain Region, 2) Species, and 3) Cell Type. These three combo boxes contain the identical values/options found in the NeuroMorpho.Org database. After selecting the appropriate values and pressing the “Get CSV” button, the program will make an API connection to the NeuroMorpho.Org server, find and then download the cells based on the search criteria. The program will first show the available number of cells and begin with a 10 second countdown during which the user can decide to continue with the generation of the .csv file or cancel the current query and acquisition. After the 10 second countdown the program will automatically continue with the query after which the NMAT will create a joined Pandas Data Frame with both cell reconstruction data and cell morphometry data. The user can follow the cells that are being queried in the text box and get feedback about the remaining time need to generate the .csv file by looking at the progress bar at the top of the window. At the end of process, the text box will show the structure of the newly created Data Frame (No. rows x No. columns) together with the time that was needed to generate the .csv file. By using the buttons “Open CSV file” or “Open file location” the user can open the .csv file or access its folder file location. If no cells are available for the given combination the program will return code 404 with the statement “Resource not found or does not exist.”, as it is the default code given by NeuroMorpho.Org server.

The “Generate CSV” tab has an underlying Python script that consists of two main parts: finding the cell reconstruction data and ID numbers of queried cells, and then fetching the morphometric data of queried cells using their ID number. This results in the creation of 3 .pkl files (“pickled” data files) which are needed to store the information about cell reconstruction and morphometric data and that will appear in the /output folder of NMAT after the first search. These are generated/overwritten for each new search that the user requests from the program. The last part of the script is used to save the generated Pandas Data Frame in the form of a .csv file. By default, each generated .csv file will be saved in the following format “NM_BrainRegion_Species_CellType”, which will correspond to the values queried by the user in the beginning of the search selection. The user also has an option to manually choose and enter the name of the .csv file that is going to be generated.

2.2.2 Getting images of downloaded cells

The NMAT also provides the ability to download an image (.png) file of each reconstructed cell. This is done in the second tab named “Get Images”. After opening this tab, the user will be shown all of the .csv files that were generated earlier and just by simply selecting the appropriate .csv file and pressing the button “Get Images”, the program will begin downloading the .png reconstructions. The files will be saved in the /output/images folder, where a new folder will be created with a name identical to the .csv file used to download image reconstructions. It is also possible to select multiple .csv files at once and download images files from more than one.csv file at the same time. This will lead to the creation of multiple new /output/images folders, with the same names as the selected .csv files.

The “Get Images” option can also work if the user modifies the previously downloaded .csv files. If the user does not need all of the image reconstructions contained in the .csv file, the user can simply delete the rows which contain those cells, save the .csv file and then select it from the “Get Images” tab of the NMAT program. The only requirement for this option to work is to not modify the “Png URL” column in the downloaded .csv file.

2.3 Multithreading architecture and additional notes

Through the use of the NMAT program it is possible to simultaneously download cell reconstruction and morphometric data and download image reconstructions based on previously generated .csv files. Each function uses multiple threads. Thus, all of the tabs work independently and allow users to perform multiple actions at the same time.

Due to its multithread-based architecture, NMAT is capable of downloading a large amount of morphometric and reconstruction data in a short time period, which represent the main advantage of the program. The multithreading capabilities were implemented through the use of QThreadPool threads for GUI threading, and concurrent.futures.ThreadPoolExecutor threads for the worker threading. GUI threading was implemented in order to allow the uninterrupted use of the program when an operation is running. Current operations are the acquisition of neuron morphometry (Acquisition) and the download of neuron images (Image). Upon starting an operation (clicking a button on the related tab) a QThread is created. The subsequent QThread activates worker threads to complete its function. Worker threading is implemented in order to speed up operations. Current maximum for threading is set to 4 threads, with the possibility of changing this number in further versions. Worker threads split up the acquiring of morphometry, and the download of images in order to speed up the requests for data.

The entire code for accessing the NeuroMorpho.Org API server and the NMAT program code are available at the NMAT project GitHub page and can be freely downloaded and modified by other users to fit their specific needs.

3 Discussion And Perspectives

The use of open-access data has become increasingly prevalent in the previous decade, with many diverse neuroscience repositories and databases being continuously created and maintained, all with the aim of enabling researchers worldwide to easily access and use different types of data. The advantages of open data in diverse biomedical areas are numerous, starting with its increased transparency, accessibility, sharing ability and possibility to combine these data with other unpublished data in order to generate new results (Huston et al., 2019). NeuroMorpho.Org, one of the largest public repositories of digital reconstructions of glial and neuronal morphology, enables researcher to gain access to tens of thousands different reconstruction and morphometric data and to freely use them for their experimental work. Hundreds of peer-reviewed articles have been published from this database and presumably this trend will continue as the database is constantly updated with new reconstructions (Ascoli, 2015).

The initial idea behind the NMAT program was not to provide novel computational techniques in neuroscience research, but to simplify and speed up the access and acquisition of cell data from the NeuroMorpho.Org database. Researchers with no or modest knowledge of API access can now easily gain access to the cells deposited in the NeuroMorpho.Org database, while the neuroscientist with advanced knowledge of Python programming language, can also find NMAT very useful as it can facilitate their current research and shorten the time need for different morphometric data acquisition. Thus, the existence of such a program is beneficial to neuroscientist with different levels of bioinformatics knowledge. Its ease of use means that NMAT may be useful even for non-scientists who, for example, simply need an illustration of a dendritic tree or axonal arbor from a typical neuron. Additionally, one of the main benefits of using the NMAT program is its multithreading architecture which allows any user to obtain a large amount of data in a short period of time. This can greatly shorten the time needed to acquire large datasets consisting of several thousand cells. As an example, a 5-year-old everyday laptop computer, running Windows version 10, with an i3 core is able to generate a .csv file of more than 1000 cells in approximately one minute. Thus, even researchers working on computers with older configuration can have quick and easy access to the NeuroMorpho.Org database.

Although the development of different Python packages and tools for neuroscientific research is now a common and widespread process, the development of easy-to-use GUI based program is becoming increasingly rare. GUI programs such as L-Measure (Scorcioni et al., 2008) and Neuromantic (Myatt et al., 2012), which are used for digital reconstruction, tracing and quantification of neuronal morphology are good examples of how different programs can be integrated with the NeuroMorpho.Org database. The current version of NMAT is primarily intended to easily access and obtain data from the NeuroMorpho.Org website. Future upgrades are planned which would further expand the usability and scope of this program. This is primarily related to the implementation of various biostatistical methods in the NMAT, which would allow the users to analyze the obtain data by means of simple and complex statistical techniques and at the same time graphically plot results. Furthermore, current additional work is underway towards increasing the usability of the program, in terms of importing of new query options, Data Frame merging, scheduled download timer, as well as a macOS version of the program. Thus, NMAT may prove as a usefully tool which would complement the NeuroMorpho.Org database and hopefully even more increase its usability.

Declarations

Information Sharing Statement

The NeuroMorpho Access Tool program, together with its source code is freely available for download at <https://github.com/izaletel/NeuroMorpho-Access-Tool>. The NMAT utilizes the NeuroMorpho.Org API to access its publicly available database (<http://neuromorpho.org/>). Python 3 (<https://www.python.org/download/releases/3.0/>) and PySide (<https://pypi.org/project/PySide6/>) were used to construct the code and GUI.

Acknowledgment

Ivan Zaletel is a Fulbright Visiting Scholar at the College of Medicine, Florida State University, Tallahassee, Florida, USA. His fellowship is supported by the Bureau of Education and Cultural Affairs of the U.S. Department of State through the Fulbright Program.

Author contributions

IZ and RSN contributed to the conception and the design of the project. MV and IZ developed the software. IZ wrote the first draft of the manuscript. RSN revised various version of the manuscript. All of the authors read, and approved the final submitted version.

References

1. Akil, H., Martone, M. E., & Van Essen, D. C. (2011). Challenges and Opportunities in Mining Neuroscience Data. *Science*, 331(6018), 708–712. <https://doi.org/10.1126/science.1199305>
2. Ascoli, G. A. (2006). Mobilizing the base of neuroscience data: the case of neuronal morphologies. *Nature Reviews. Neuroscience*, 7(4), 318–324. <https://doi.org/10.1038/nrn1885>
3. Ascoli, G. A., Donohue, D. E., & Halavi, M. (2007). NeuroMorpho.Org: A Central Resource for Neuronal Morphologies. *Journal of Neuroscience*, 27(35), 9247–9251. <https://doi.org/10.1523/jneurosci.2055-07.2007>
4. Ascoli, G. A. (2015). Sharing Neuron Data: Carrots, Sticks, and Digital Records. *PLoS Biology*, 13(10), e1002275. <http://doi.org/10.1371/journal.pbio.1002275>
5. Freeman, J. (2015). Open source tools for large-scale neuroscience. *Current Opinion in Neurobiology*, 32, 156–163. <https://doi.org/10.1016/j.conb.2015.04.002>
6. Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., et al. (2020). Array programming with NumPy. *Nature*, 585(7825), 357–362. <https://doi.org/10.1038/s41586-020-2649-2>
7. Huston, P., Edge, V., & Bernier, E. (2019). Reaping the benefits of Open Data in public health. *Canadian Communicable Disease Report*, 45(11), 252–256. <https://doi.org/10.14745/ccdr.v45i10a01>
8. Keil, J. M., Qalieh, A., & Kwan, K. Y. (2018). Brain Transcriptome Databases: A User's Guide. *Journal of Neuroscience*, 38(10), 2399–2412. <https://doi.org/10.1523/jneurosci.1930-17.2018>
9. Pereira, S., Gibbs, R. A., & McGuire, A. L. (2014). Open Access Data Sharing in Genomic Research. *Genes (Basel)*, 5(3), 739–747. <https://doi.org/10.3390/genes5030739>
10. Poldrack, R. A., & Gorgolewski, K. J. (2014). Making big data open: data sharing in neuroimaging. *Nature. Neuroscience*, 17(11), 1510–1517. <https://doi.org/10.1038/nn.3818>
11. Moore, B., Carvajal-López, P., Chauke, P. A., Cristancho, M., Angel, V. D. D., Fernandez-Valverde, S. L., et al. (2021). Ten simple rules for organizing a bioinformatics training course in low- and middle-

- income countries. *PLoS Computational Biology*, 17(8), e1009218.
<http://doi.org/10.1371/journal.pcbi.1009218>
12. Myatt, D. R., Hadlington, T., Ascoli, G. A., & Nasuto, S. J. (2012). Neuromantic – from Semi-Manual to Semi-Automatic Reconstruction of Neuron Morphology. *Frontiers in Neuroinformatics*, 6, 4.
<http://doi.org/10.3389/fninf.2012.00004>
 13. Muller, E., Bednar, J. A., Diesmann, M., Gewaltig, M. O., Hines, M., & Davison, A. P. (2015). Python in neuroscience. *Frontiers in Neuroinformatics*, 9, 11. <http://doi.org/10.3389/fninf.2015.00011>
 14. Scorcioni, R., Polavaram, S., & Ascoli, G. A. (2008). L-Measure: a Web-accessible tool for the analysis, comparison, and search of digital reconstructions of neuronal morphologies. *Nature Protocols*, 3(5), 866–876. <https://doi.org/10.1038/nprot.2008.51>
 15. Via, A., Blicher, T., Bongcam-Rudloff, E., Brazas, M. D., Brooksbank, C., Budd, A., et al. (2013). Best practices in bioinformatics training for life scientists. *Briefings in Bioinformatics*, 14(5), 528–537. <https://doi.org/10.1093/bib/bbt043>

Figures

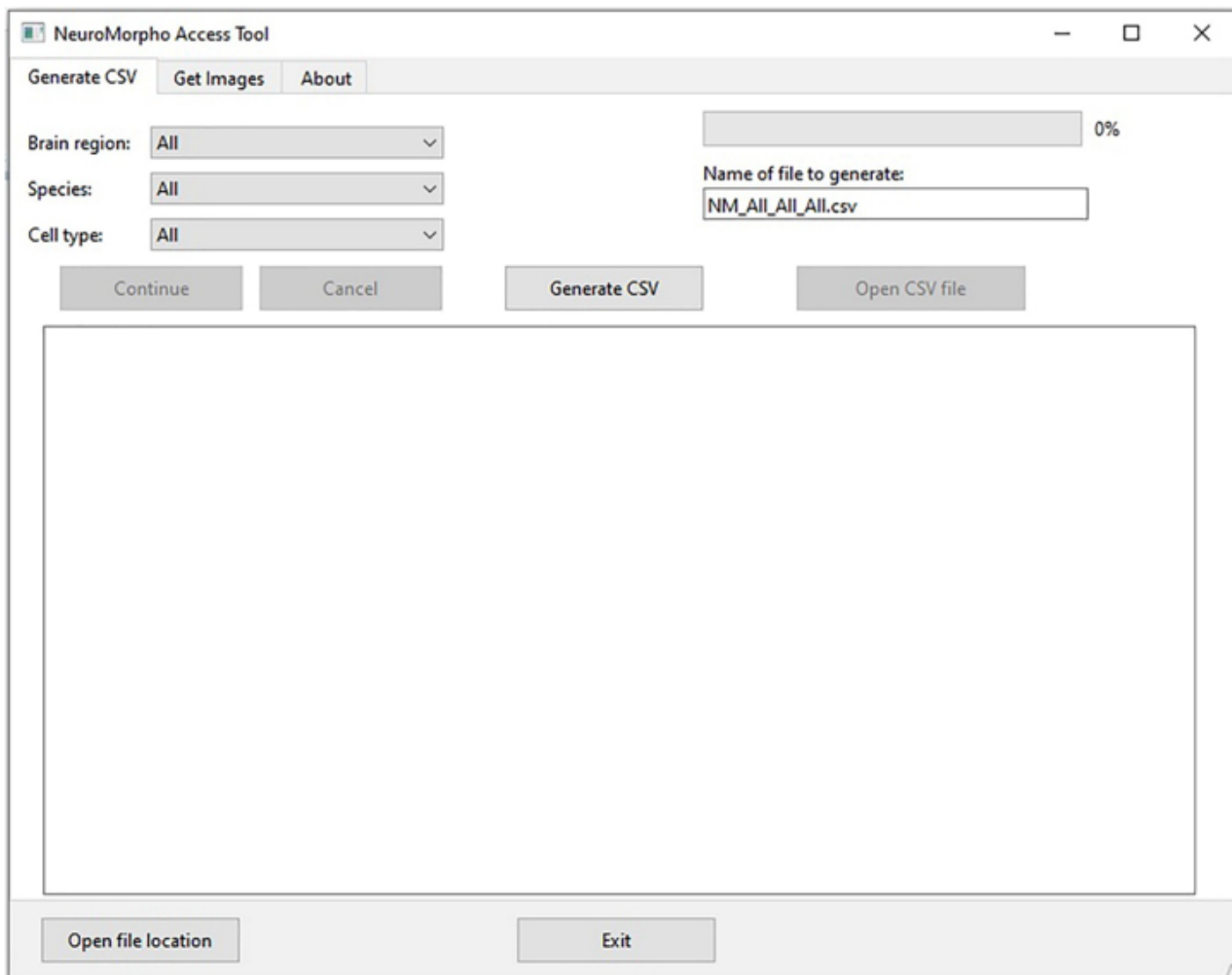


Figure 1

NeuroMorpho Access Tool Program Window. The NMAT possess a simple and friendly user interface which consists of combo boxes and text boxes organized into 2 main working tabs, which are used to obtain cell data in the form of .csv file and acquire cell image reconstructions in the form of .png image files. The main text box shows information about the morphometrical data and images that are being downloaded, which allows the users to follow the progress of the tasks assigned to the NMAT program.