Eurographics Workshop on Visual Computing for Biology and Medicine (2022)
T. Schultz and H.-Y. Wu (Editors)

Supplementary Information: A Stratification Matrix Viewer for Analysis of Neural Network Data

Philipp Harth[1], Sumit Vohra[1], Daniel Udvary[2], Marcel Oberlaender[2], Hans-Christian Hege[1] and Daniel Baum[1]

[1]Zuse Institute Berlin (ZIB), Department of Visual and Data-Centric Computing, Germany
[2]Max Planck Institute for Neurobiology of Behavior – caesar, Bonn, Germany

Abstract

This document provides supplementary information to the paper titled "A Stratification Matrix Viewer for Analysis of Neural Network Data". In particular, it describes at a more technical level how cells in the stratification matrix encode data (Sect. S1). In addition, it provides further information on the neural network datasets used in the case studies of the main paper (Sect. S2). Finally, it shows how the tool can be applied to a dataset from a different application domain, namely epidemiology (Sect. S3).

S1. Encoding connectional and non-connectional information

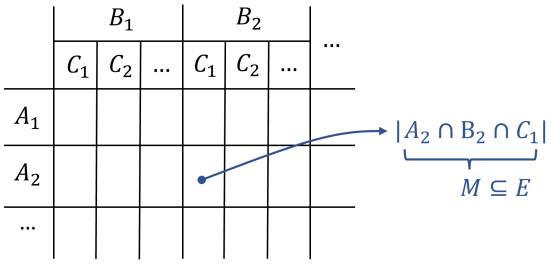
Cells in the stratification matrix show the number of entities (e.g., neurons, synapses) that are filtered by the row and column annotations of the matrix, which correspond to categorical properties of the entities (e.g., cell type, soma location). Let $E = \{e_1, e_2, \ldots, e_n\}$ be the set of entities and X a categorical property that takes n_X discrete values and that is defined for every entity e through map σ_X,

$$\sigma_X : E \to \{1, 2, \ldots, n_X\}.$$

We define X_k to be the subset of entities that have property value k,

$$X_k = \{e_i \in E \mid \sigma_X(e_i) = k\}.$$

Assume that the entities of the dataset have three properties A, B, C and that the user has assigned property A to the rows of the matrix and properties B and C to the nested columns. Then matrix cells show the number of entities $|M \subseteq E|$ that satisfy all of the respective property values in rows and columns.



The stratification matrix displays connectional or non-connectional information depending on the nature of the entities E and the semantics of their properties $A, B, \ldots$ that are assigned to rows and columns of the matrix by the user. If the entities represent connectional information, such as synaptic contacts, the stratification matrix can be interpreted as a connectivity matrix, otherwise as a generic tabular layout. We illustrate this with two examples.

S1.1. Example for connectional information

- Entities E: synapses in human temporal cortex dataset (synapses are the connection points between neurons and signal flow is directed; therefore every synapse has a presynaptic and a postsynaptic neuron)
- Property $A = \{\text{L2, L3, L4, L5}\}$ layer in which soma of presynaptic neuron is located (red, assigned to rows)
- Property $B = \{\text{INTER, PYR}\}$ cell type (inter- or pyramidal neuron) of presynaptic neuron (blue, assigned to rows)
- Property $C = \{\text{L2, L3, L4, L5}\}$ layer in which soma of postsynaptic neuron is located (cyan, assigned to columns)
- Property $D = \{\text{INTER, PYR}\}$ cell type (inter- or pyramidal neuron) of postsynaptic neuron (orange, assigned to columns)

		L2		L3		L4		L5	
		PYR	INTER	PYR	INTER	PYR	INTER	PYR	INTER
L2	PYR	3384	6022	2149	2151	187	54	86	11
	INTER	1342	586	499	280	12	2	7	0
L3	PYR	560	298	2397	1999	557	366	103	33
	INTER	357	111	1672	397	269	44	24	0
L4	PYR	32	5	640	240	1534	2002	344	439
	INTER	17	2	256	50	808	93	53	7
L5	PYR	25	4	50	11	113	91	414	578
	INTER	4	1	1	1	116	21	360	55

The highlighted matrix cell shows the number of synapses that connect PYR neurons from L3 with INTER neurons from L4. Here, the stratification matrix represents a hierarchical connectivity matrix.

S1.2. Example for non-connectional information

	L1	L2	L3	L4	L5	L6	WM
neuron	555	4573	2639	3388	2268	1045	612
glial	2670	1826	3384	3043	4056	2651	11660

- Entities E: cells in human temporal cortex dataset
- $A = \{\text{neuron, glial}\}$ cell type (orange, assigned to rows)
- $B = \{\text{L1, L2, L3, L4, L5, L6, WM}\}$ cortical layer in which cell body is located (blue, assigned to columns)

The highlighted matrix cell shows the number of glial cells that reside in L5. Here, the stratification matrix represents a generic table.

S2. Neural network datasets

S2.1. Rat barrel cortex

We created a detailed anatomical model of the rat barrel cortex (Fig. S1) that is based on experimental measurements of barrel cortex geometry, cellular composition, and neuronal morphology [UHM*22]. The barrel cortex covers a volume of approximately $6.5\,mm^3$, contains about 550,000 neurons, and approximately 5 billion synapses. The biological function of the barrel cortex is to process sensory information that is received from the whiskers on the snout of the rat. Relevant properties by which neurons in this dataset can be grouped and filtered are:

- **cell type**: L2PY, L3PY, L4PY, L4sp, L5IT, L5PT, L6CC, L6INV, L6CT, VPM, and INH,
- **cortical column**: A1-E4, α-δ (24 values in total),

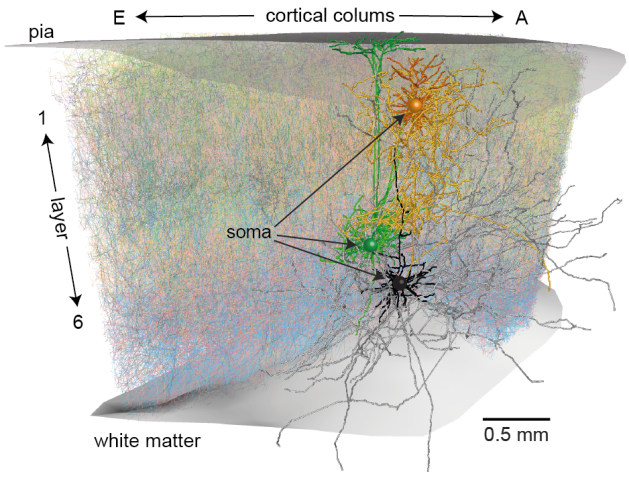


Figure S1: *Dense model of the rat barrel cortex [UHM*22].*

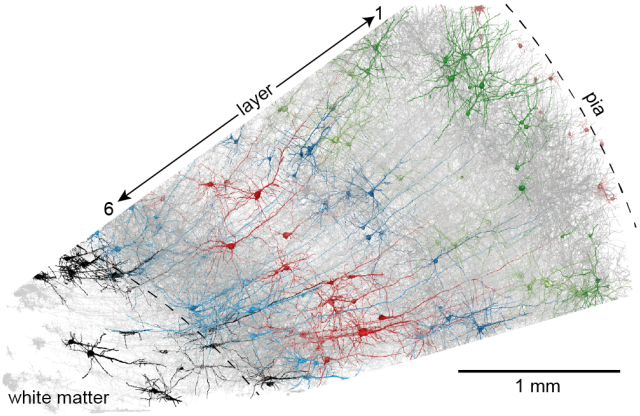


Figure S2: *Dense reconstruction of a slice of the human temporal cortex [SCJB*21].*

- **subregion**: inside, septum, other (indicates whether a neuron resides inside the column, between columns, or in thalamic barreloids),
- **cortical depth**: 100, 200, ..., 2200, other (depth of the soma in relation to the pia, binned at intervals of 100 microns).

S2.2. Human temporal cortex

The human cortex dataset is an electron microscopy-based dense reconstruction of $1\,mm^3$ of the human temporal cortex (Fig. S2). The dataset contains 57,000 cells, 16,000 fully annotated neurons and more than 133 million synapses [SCJB*21]. Relevant properties by which neurons in this dataset can be grouped and filtered are:

- **cell type**: neuron, glial, blood vessels, n/a,
- **neuron annotation**: pyramidal, interneuron, ...,
- **layer**: L1, L2, L3, L4, L5, L6, WM (white matter),
- **axon nodes**: number of axon skeleton nodes—can be seen as a proxy for axon length,
- **dendrite nodes**: number of dendrite skeleton nodes—can be seen as a proxy for dendrite length,
- **synapses outgoing**: number of outgoing synapses from this neuron to other neurons,
- **synapses incoming**: number of incoming synapses from other neurons to this neuron.

S3. Application to epidemiological dataset

We also applied our tool to a dataset from a different domain (epidemiology) that accounts for all reported COVID-19 infections and deaths in Germany from the first calendar week of 2020 to the 25th calendar week of 2022 [RKI22]. Matrix cells encode the number of reported infections (channel 1) and COVID-19-related deaths (channel 2). The goal of the user is to visualize the temporal evolution of the pandemic and to understand its impact on different subpopulations. In Fig. S3A, the matrix cells encode the number of reported cases and deaths for the respective subpopulation at

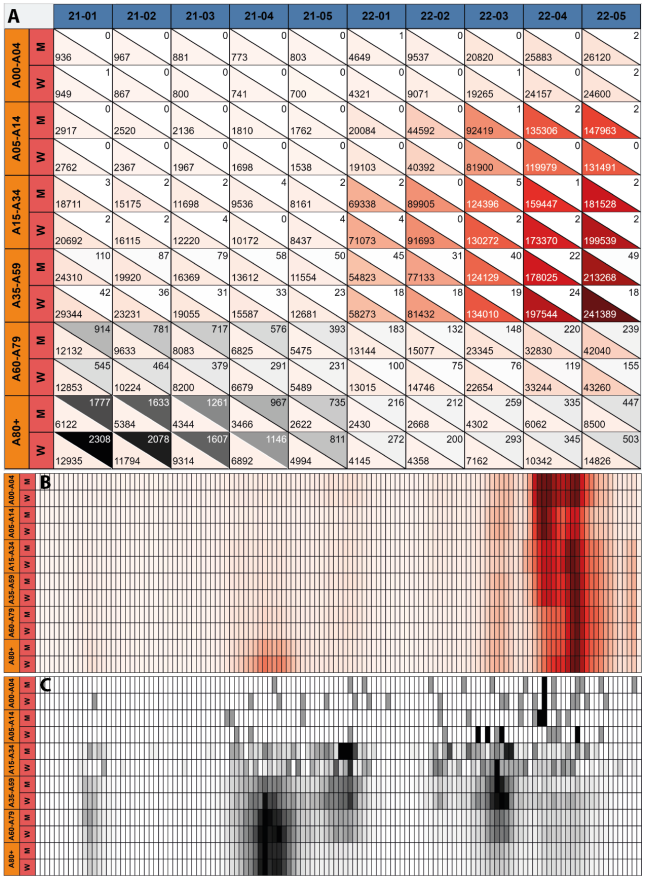


Figure S3: ***A****: COVID-19 cases in Germany [RKI22] stratified by age group, gender (rows) and week of year (columns), covering the calendar weeks 2021-w01 to 2021-w05 and 2022-w01 to 2022-w05. Red-colored triangles show the number of reported cases, gray colored triangles the number of COVID-19-related deaths (red and gray color scales were normalized by respective max values in matrix).* ***B,C****: Reported cases/deaths for all weeks from 2020-w01 to 2022-w25, with color scales normalized separately for each row by the maximum value in the row.*

the calendar weeks from 2021-w01 to 2021-w05 and from 2022-w01 to 2022-w05. In terms of reported cases, the wave in the first weeks of 2022 exceeds all previous waves. However, most of the casualties in higher age groups occurred in the second wave in the last weeks of 2020 and the first weeks of 2021. This also becomes apparent when displaying cases and deaths separately for the entire course of the pandemic and normalizing the color scale separately for each row. The number of reported cases was substantially higher in the 2021-2022 wave (Fig. S3B). Presumably due to vaccination and different variants of the virus, casualty rates for the older population were lower in this wave than in the preceding second wave of 2020-2021. In relative terms, deaths among the young and middle-aged in the 2021-2022 wave were more frequent than in preceding waves (Fig. S3C). In absolute terms, however, deaths were much lower in the young and middle-aged groups than in the elderly population, as is easily checked by inspecting the matrix cells in (Fig. S3A), where absolute numeric values are displayed in addition to the normalized color scale. [video 9:10]

References

[RKI22] ROBERT-KOCH-INSTITUT: SARS-CoV-2 Infektionen in Deutschland, June 2022. doi:10.5281/zenodo.6672879. S2, S3

[SCJB*21] SHAPSON-COE A., JANUSZEWSKI M., BERGER D. R., POPE A., WU Y., BLAKELY T., SCHALEK R. L., LI P. H., WANG S., MAITIN-SHEPARD J., KARLUPIA N., DORKENWALD S., SJOSTEDT E., LEAVITT L., LEE D., BAILEY L., FITZMAURICE A., KAR R., FIELD B., WU H., WAGNER-CARENA J., ALEY D., LAU J., LIN Z., WEI D., PFISTER H., PELEG A., JAIN V., LICHTMAN J. W.: A connectomic study of a petascale fragment of human cerebral cortex. *bioRxiv* (2021). doi:10.1101/2021.05.29.446289. S2

[UHM*22] UDVARY D., HARTH P., MACKE J. H., HEGE H.-C., DE KOCK C. P. J., SAKMANN B., OBERLAENDER M.: The impact of neuron morphology on cortical network architecture. *Cell Reports 39*, 2 (2022). doi:10.1016/j.celrep.2022.110677. S2