Automated data-set generation pipeline for 3D Neuron visualization and cloud processing in Kbrain-map DB station portal

Nam Uk Kim[1,2], Byeong Soo Kang[3], and Sung-Jin Jeong[1,2,*]

Neural Circuit Research Group, Korea Brain Research Institute, Daegu, 41068, Korea,
[2]*Molecular Aging & Development Laboratory, Korea Brain Research Institute, Daegu, 41068, Korea,*
[3]*R&D Center, SYSOFT, Daegu, 42988, Korea*
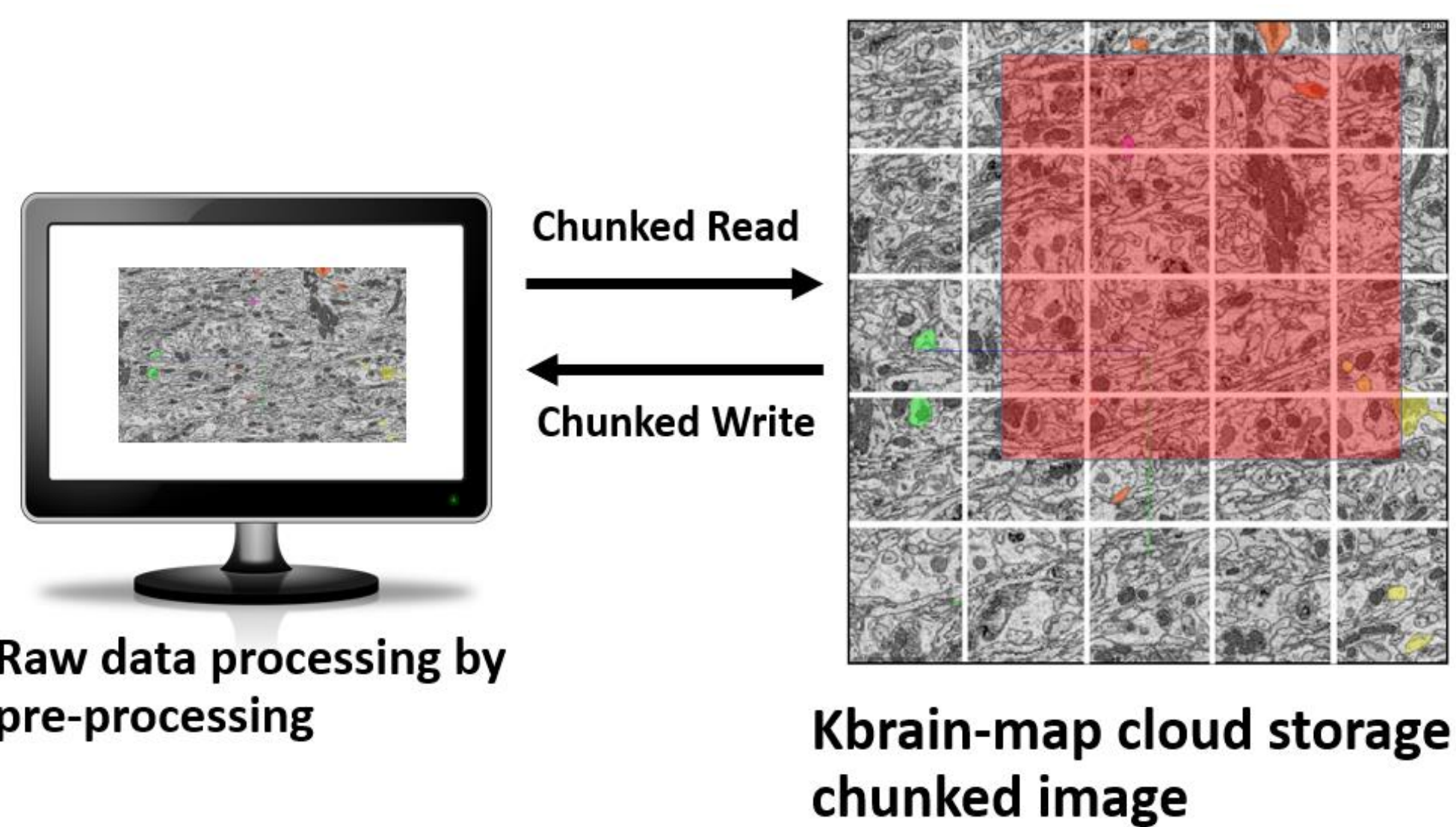**Corresponding author: sjjeong@kbri.re.kr*

Abstract

Recently, high-resolution image and video data are being increased exponentially in the field of life sciences by the advanced imaging technologies. Among them, connectomics images of brain tissues converted nano and micrometer high-resolution into 3D image produce big data characteristics. In order to visualize such connectomics image effectively, there are resource constraints such as memory and disk. In addition, as Web services grow in scale, sudden increase in traffic causes server bottlenecks as well as degradation of server performance due to server overloading problems. It is essential to establish a Web-based platform which allows the connectomics images to be visualized and analyzed in 3D at anytime and anywhere without restriction on the spatial environment. Also, in order to improve the processing time and speed of the entire image, it is necessary to divide into the image pieces and distribute them to the system. In this study, we propose pre-computed pipeline and methodology generating an automated data set and providing an advantage of block storage in a cloud environment, which are eventually utilized for visualization and analysis of connectomics images in three dimensions through a web browser. The 3D image visualization utilizes WebGL-based 3D open source. The 3D image visualization system was customized to effectively visualize the data generated by the pipeline. The pipeline sets the bounding box space of x, y, z axis and divides into 3D chunk units by slicing work for each area. The divided 3D chunk dataset and information files are kept in the block storage of the cloud and the dataset is converted to the KNeuroViz format for efficient I/O operations. This system is aimed to build an interactive database for brain connectome convergence research based on a user interface that can be integrated with various analysis modules.

Method : Automated dataset generation pipeline

### Cut-outed cubic Image and KNeuroViz 3D viewer flow chart
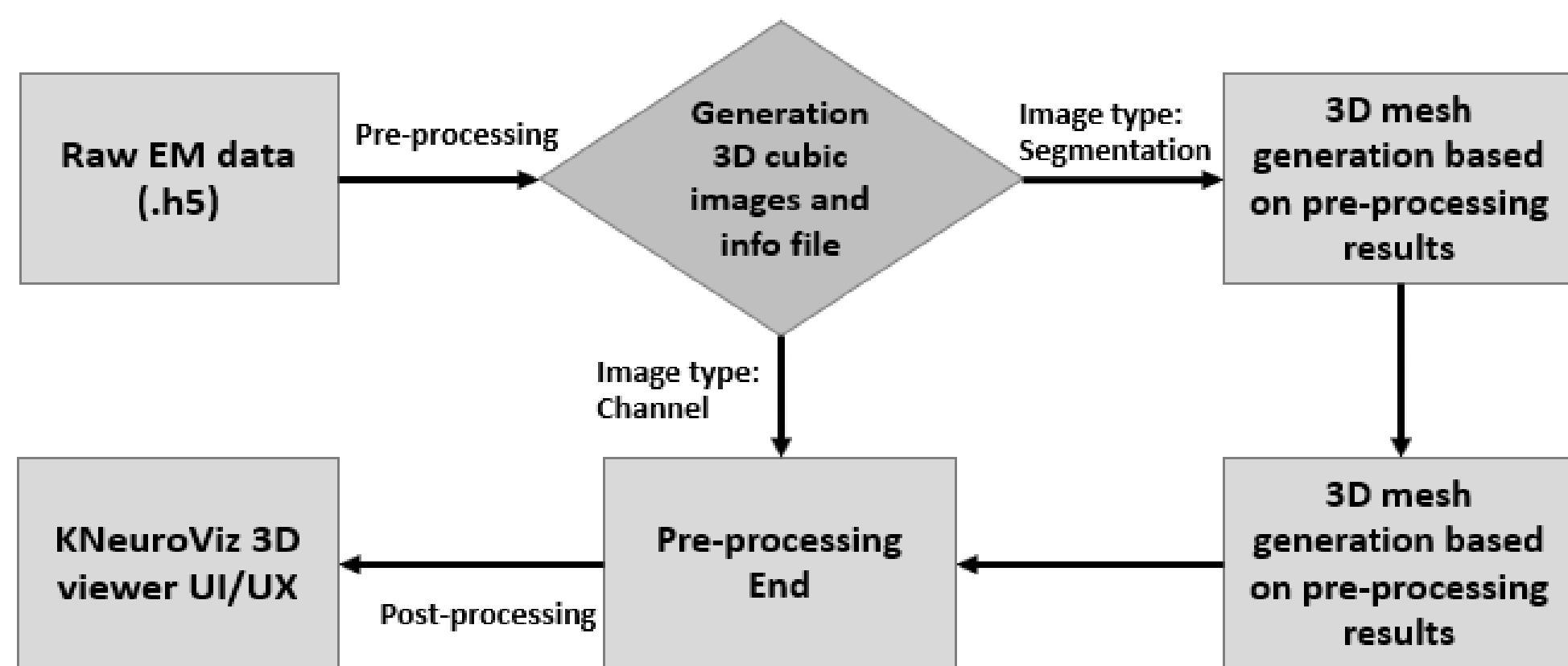
The pipeline splits the raw data into chunks and processes them into a data format that can be read by KNeuroViz Web.

Raw data processing by pre-processing

Chunked Read

Chunked Write

Kbrain-map cloud storage chunked image

The image sliced by the bounding box is stored in the cloud storage as a chunk image. (KNeuroViz pre-computed format). Read in chunks and crop to ROI. The Writing object divides the image into chunks. Also, in technical side, Writing requires aligning chunks to avoid race conditions.

Pre-processing dataset generation pipeline for 3D volume visualization (Front-end)
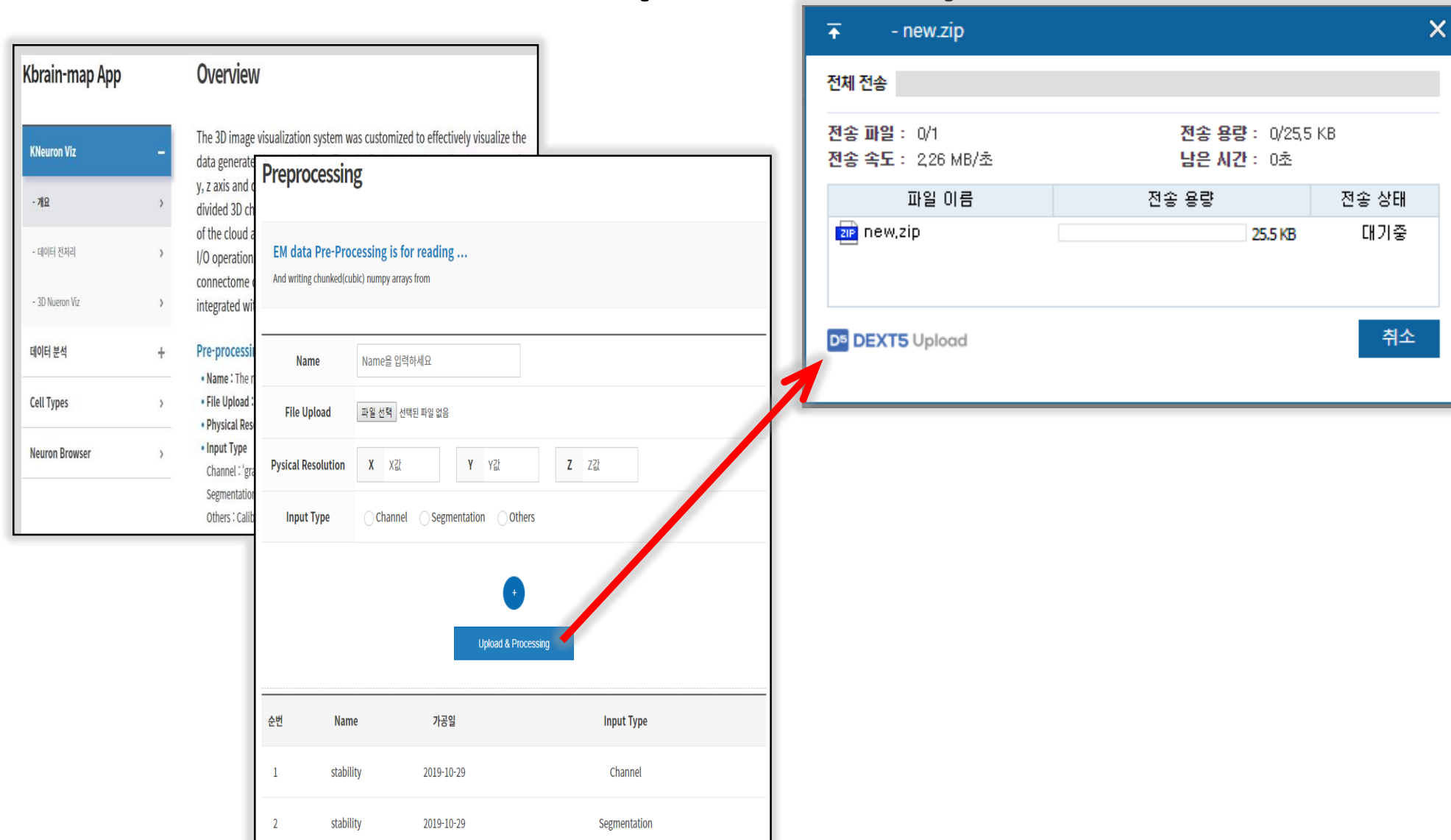
The pre-processing pipeline is divided into chunk data by bounding boxes for each region of the x, y, and z axes. By creating a divided data set for each x, y, z axis area and related info file, It performs pre-processing to efficiently load and visualize in the web's 3D visualization platform.

Structure : Web based DB station technology

Architecture : Interactive integrated DB station Kbrain-map and cloud system

Integrated DB station infrastructure has the advantage of providing optimized IaaS by creating virtual machine for cluster-based hybrid big data analysis environment and big data processing. Also, **the integrated DB station system** provides a platform to share data for integration that can collaborate on data processing between domestic and foreign researchers.

KNeuroViz : Structure and principle
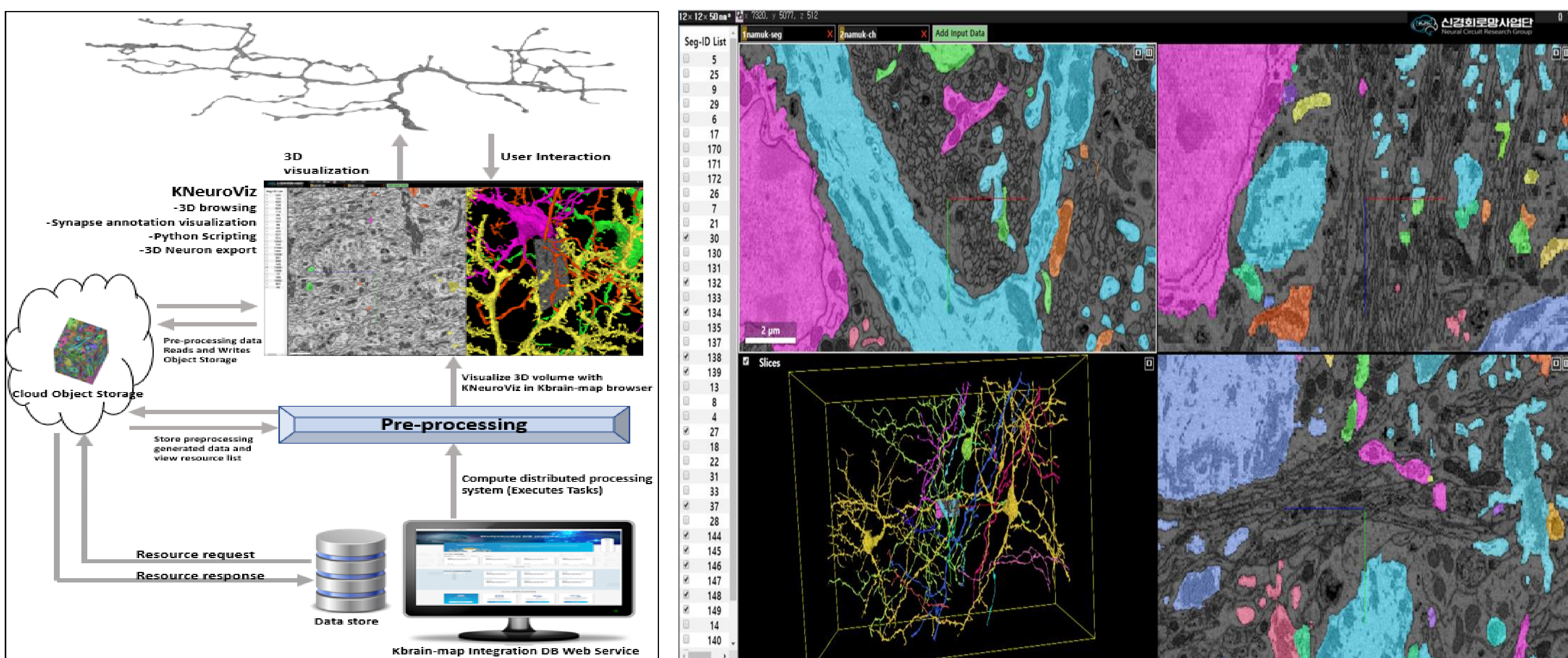
Visualization of the data stored in cloud using KNeuroViz, one the S/W in Kbrain-map web server. KNeuroViz providers a benefits a real-time streaming of the chunked image data and their visualization in 3D. 3D rendering speed is improved due to the proposed technology (where available).

Kbrain-map App: Neuron browser for visualization and analysis of 3D neuron meshes
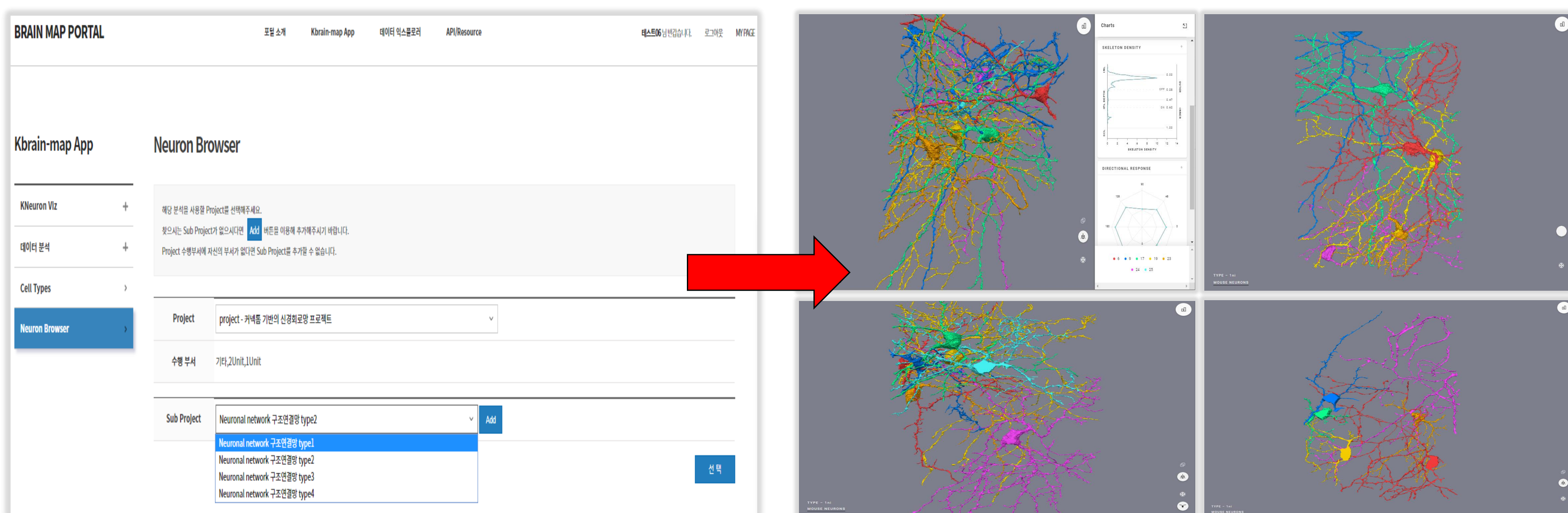
The Neuron-browser connects with KNueronViz and provides the ability to find and search cell classifications to determine the structure of connections between neurons.

Features of KNeuroViz

• Development of function to set MIP level in pre-processing

• System integration for visualization and numerical analysis of specific 3D neurons in Kbrain-map

• Extended to handle various raw data

• KNeuronViz system and GPU memory allocation UI implementation

• Integrated database for 3D neuron information in KNeuroViz

• Support for skeleton 3D reconstruction in pre-processing

Reference

1. Vogelstein, Joshua T., et al. "To the cloud! A grassroots proposal to accelerate brain science discovery." *Neuron* 92.3 (2016)

2. Lichtman, Jeff W., Hanspeter Pfister, and Nir Shavit. "The big data challenges of connectomics." *Nature neuroscience* 17.11 (2014)

3. Seung, H. Sebastian. "Neuroscience: towards functional connectomics." *Nature* 471.7337 (2011)

4. Saalfeld, Stephan, et al. "CATMAID: collaborative annotation toolkit for massive amounts of image data." *Bioinformatics* 25.15 (2009)

5. Haehn, Daniel, et al. "Scalable interactive visualization for connectomics." *Informatics*. Vol. 4. No. 3. Multidisciplinary Digital Publishing Institute (2017)

Acknowledgement