Spatial-Live: A **li**ghtweight and **ve**rsatile tool for single cell spatial-omics data visualization

Zhenqing Ye[1,2*], Zhao Lai[1,3], Siyuan Zheng[1,2], Yidong Chen[1,2*]

1. Greehey Children's Cancer Research Institute, University of Texas Health San Antonio, San Antonio, Texas, 78229, USA

2. Department of Population Health Sciences, University of Texas Health San Antonio, San Antonio, Texas, 78229, USA

3. Department of Molecular Medicine, University of Texas Health San Antonio, San Antonio, Texas, 78229, USA

* To whom correspondence should be addressed.
Contact: yez@uthscsa.edu or cheny8@uthscsa.edu

Abstract: Single cell spatial-omics data visualization plays a pivotal role in unraveling the intricate spatial organization and heterogeneity of cellular systems. Various software tools and packages have been developed to effectively visualize and interpret single-cell spatial-omics data. However, challenges still exist in areas such as user-friendly accessibility, multiple modal data integration, and responsively interactive features. In this study, we introduce Spatial-Live, a **li**ghtweight and **ve**rsatile viewer tool specifically designed for flexible single-cell spatial-omics data visualization. Spatial-Live seamlessly integrates and stacks multiple layers of data into a unified three-dimensional (3D) environment, providing natural compatibility for visualizing multi-type spatial data in a single interface. By leveraging the GPU rendering capability, Spatial-Live excels at efficiently processing large datasets while offering interactive, responsive features and a wide range of visualization effects achieved through the stacking of multiple layers in one 3D space.

Introduction:

The field of single-cell spatial genomics research has experienced remarkable advances in recent years, driven by innovative platforms, diverse tools, and various techniques [1-5]. Among them, data visualization is vital for understanding complex spatial organization and cellular heterogeneity within tissue patterns. Several software tools and frameworks have emerged to effectively visualize and interpret single-cell spatial-omics data [6-8]. For instance, TissUUmaps1 [9] and the recently updated TissUUmaps3 [10] address challenges associated with analyzing large-scale spatial-omics data. This tool emphasizes interactive visualization capabilities, enabling intuitive exploration and

interaction with data, while also providing features for quality assessment. Another notable framework is Vitessce [11], which offers a comprehensive platform for visualizing and exploring diverse types of spatially resolved single-cell data. With its user-friendly interface and a wide range of visualization techniques, Vitessce empowers researchers to unravel complex cellular dynamics.

While significant progress has been made in the development of data visualization toolkits, persistent challenges remain in areas such as user accessibility and the integration of multiple modalities of data [12, 13]. A particularly noteworthy challenge is the increasing emergence of multiple-modal single-cell spatial data, which demands the integration of diverse data types within a unified framework. For instance, the inclusion of pathological imaging data, immunofluorescent staining data, expression profiles, chromatin data, and more requires a flexible visual platform capable of accommodating these diverse datasets simultaneously. Although tools like TissUUmaps [9, 10] and Viteessce [11] have made strides in addressing some of these challenges, they primarily rely on 2D layout visualization, in which each layer overlays on another layer in an orthographic way in 2D space.

While the orthographic way of visualizing data in a 2D space has proven to be valuable in many spatial genomics applications, it comes with a few inherent limitations and drawbacks for 2D layer overlay. For example, in an orthographic view, the lack of perspective and depth perception can make it challenging to accurately interpret the relative positions and distances between features in different layers. Furthermore, when multiple layers are overlaid in a 2D orthographic view, it can be difficult to visualize and distinguish overlapping features. This can lead to potential ambiguity in identifying and analyzing specific regions of interest, particularly when there are complex spatial arrangements or dense overlapping structures. Additionally, in an orthographic 2D view, the ability to dynamically explore and manipulate the data may be limited compared to a 3D environment. The absence of rotation, zooming, and other interactive features can limit researchers' ability to thoroughly analyze the spatial genomics data from different angles and perspectives.

On the other hand, nowadays there are many different technical platforms for spatial genomics study, offering diverse technologies and approaches [14, 15], including Vizgen's MERFISH, 10x Genomics' Visium and Xenium, and Nanostring's CosMx. The dynamic and ever-evolving nature of spatial-omics research calls for tools that are agile, adaptable, and capable of accommodating diverse datasets from various platforms. While comprehensive tools like TissUUmaps [9, 10] and Vitesscee [11] offer an extensive range of features, they often come with increased complexity, resource requirements, and steep learning curve. In contrast, a simple and adaptable tool offers better flexibility,

enabling researchers to quickly adapt to evolving research needs and efficiently visualize and analyze spatial genomics data.

In this study, we introduce Spatial-Live, a lightweight and versatile viewer specifically designed for the flexible visualization of single-cell spatial-omics data. This tool is finely tuned to focus on the essential data variables necessary for your analysis only. Spatial-Live can seamlessly integrate and stack multiple layers of data into a unified 3D environment, overcoming those drawbacks presented in 2D orthographic view mode, and providing natural compatibility for visualizing diverse spatial data types in a single interface. By leveraging the GPU rendering capability, Spatial-Live efficiently processes large datasets while offering interactive, responsive features and a wide range of visualization effects achieved through the stacking of multiple layers in one 3D space.

Results:

The conceptual framework and schema of Spatial-Live

Numerous platforms aid in exploring cellular spatial organization, but also pose challenges in terms of data integration for achieving unified visualization and exploration. Spatial-Live addresses these challenges by being platform-agnostic and prioritizing a data-type-oriented approach. Regardless of the diverse origins of spatial data sources, from the Spatial-Live perspective, datasets from various platforms can be uniformly preprocessed and consolidated into three input files, namely the image PNG file, the data CSV file, and the polygon-shape JSON file. As shown in Fig.1, the image file (in PNG format) is mandatory and plays a pivotal role in establishing the pixel coordinate space for all other data plotting. While the JSON file is an optional component, it can prove invaluable in specific scenarios, allowing for annotation of regions of interest (ROIs) on the image. These JSON files can represent various features [16], such as cancer cell-enriched regions or abnormal tissue zones, as collections of geometric polygons rendered as a GeoJson Layer.

The Comma-Separated Value (CSV) file assumes a central role in Spatial-Live, serving as the primary means for structuring and accommodating variables. It facilitates the organization of data from various platforms after preprocessing, segregating them into distinct variable types: categorical, numerical, and gene variables. These variable types are represented by different column headers in the CSV file, as illustrated in Fig.1, such as ‘char:[variable]’, ‘num:[variable]’, and ‘gene:[variable]’. Keep its lightweight nature in mind, Spatial-Live prefer to load only the essential data variables necessary for your analysis, rather than the entire spectrum of the dataset. In addition to these, three fundamental columns are essential: ‘id:spot', 'pos:pixel_x', and 'pos:pixel_y'. These columns specify the unique spot IDs and pixel positions for each spot (or cell), cooperating with the underlying image to establish the pixel coordinate base. By

analyzing these columns, Spatial-Live's rendering engine efficiently generates corresponding visual layers for each variable type, leveraging a GPU-powered backend. In Fig.1, categorical variables appear as ScatterLayers, numerical variables as ColumnLayers, and gene variables as HeatmapLayers. Furthermore, the Spatial-Live engine offers an array of interactive controls, including 2D orthographic or 3D orbit view modes, pan/zoom functionality, and color platters. More details can be found in "Materials and Methods" and the online documentation.

Implementation and graphical user interface:

To extend and overcome the limitations of the 2D orthographic view as mentioned previously, we integrated a 3D orbit perspective view mode into Spatial-Live. To implement this feature, as well as the structural layout as shown in Fig.1, we utilized the deck.gl library [17], a powerful WebGL2 (GPU baked) data visualization infrastructure, which can offer many visualization layers that can be customized and adapted, allowing us to create visually appealing and interactive data visualizations. By leveraging the capabilities of deck.gl, along with other modern JavaScript frameworks like Vue3 and React, we forged the Spatial-Live Rendering Engine, which offers users an immersive 3D viewing experience, elevating the exploration and analysis of spatial-omics data, through an appealing multiple-layer stacking approach. By adjusting the z-height level of each layer, Spatial-Live allows for proper positioning and view angles, ensuring the effective overlay of multiple layers in 3D space.

The Spatial-Live user interface consists of three main components: the left menu pane, the middle main visual pane, and the right control panel, as shown in Fig.2. The left panel controls visibility of different variable layers. When toggled to the "on" state (checkbox), the corresponding layer is added to the visual output, with the active layer highlighted in yellow. The middle pane dynamically renders the visual output in response to parameter changes and enables interactive manipulation through mouse actions, including zooming, rotating, and dragging. This functionality provides users with the ability to explore and visualize data from various angles and perspectives. The top-right panel provides a convenient interface for adjusting parameters to control the appearance of the active layer and optimize the layout. Notably, users can switch between 2D orthographic and 3D orbiting perspective modes to suit their viewing preferences.

Furthermore, Spatial-Live has the option to enable tooltips for most layers, except for the Gene Heatmap layer. The tooltips can provide useful hints and additional information in certain cases. To enable tooltips, simply toggle the "Tooltip" button located in the right control panel. Additionally, once all the layers are finalized and properly stacked, users can export the visualization as an external image by clicking the

"Export Image" button. For a more comprehensive guideline of using Spatial-Live, we have provided a quick demo and instructional video. Please refer to the online documentation for detailed instructions and further information.

Practical biological applications:

To maintain platform-agnosticism, we have deliberately decoupled data processing from the Spatial-Live tool, recognizing that it may be tied to specific single-cell spatial platforms. This separation allows us to concentrate exclusively on visualization. Nonetheless, for optimal utilization of Spatial-Live, it is essential to supply properly formatted input files. To facilitate this process, comprehensive tutorials have been curated to exemplify the preparation of these files. These resources can be readily accessed via our online documentation.

In this context, an illustrative example has already been presented in Fig.2, based on a study in which the authors induced acute kidney injury (AKI) in mice using the CLP model (cecal ligation puncture) [18]. The spatial transcriptomic data were generated from the 10x Visium platform. As evident from the central portion of Fig.2, the careful arrangement of multiple layers offers a lucid depiction of distinct kidney regions, including the cortex and inner medulla. These regions are discerned using annotated polygon shapes displayed via the Geo-JSON layer, as well as the “leiden” clusters highlighted through the scatter layer. Moreover, the Column bar plot of a specific gene (e.g. Sprr1a) vividly demonstrates gene expression fluctuations within these areas. With the capabilities of Spatial-Live, we can delve into individual cell spots beneath the column. As we descend along the column, we can seamlessly zoom in on the adjacent spots, allowing for a comprehensive examination of histopathological details within the tissue image.

To further demonstrate the versatility of Spatial-Live in handling data from various platforms, we processed and visualized mouse liver MERFISH data from the Vizgen platform. Specifically, cell filtering was based on marker genes persistently expressed in hepatocyte peri-central and peri-portal zones [19]. In addition, we prepared the required input PNG image and CSV data files for Spatial-Live, aptly integrating two numerical variables and two gene variables. As demonstrated in Fig.3, a screenshot captured from Spatial-Live effectively highlights the tool’s capacity to convey abundant information via a singular plot. The light-blue and orange heatmaps represent the gene variables ‘gene:Aldh3a2’ and ‘gene:Hsd17b6’, capturing hepatocyte peri-portal and peri-central spatial distributions while unveiling a unique, mutually exclusive jagged pattern. The two numerical variables, ‘num:Vwf’ and ‘num:Axin2’, plotted as purple and green column bars, respectively, presenting discernible spatial trends within hepatocyte zones; notably,

enriched Vwf expression borders blood vessels. Through strategic layer stacking, an engaging visual depiction of processed mouse liver MERFISH data can be achieved.

Conclusion:

Recent advances in single-cell spatial genomics research have been greatly accelerated by various innovative tools and platforms [15, 20]. Such advancements are crucial for unraveling the complex spatial organization and heterogeneity within tissue cellular patterns. Effective data visualization plays a pivotal role in data exploration and interpretation, serving as a key component in gaining insights from complex datasets. Spatial-Live represents a valuable tool for achieving these objectives, especially for single-cell spatial-omics data visualization.

In summary, traditional 2D orthographic visualization, while valuable in spatial genomics, has limitations compared to the immersive 3D orbit view. Orthographic projection's fixed perspective and limited depth perception may not fully capture complex spatial relationships in tissue context. Conversely, the 3D orbit view in Spatial-Live allows for a more intuitive exploration of spatial genomics data, offering dynamic rotation and viewpoint manipulation. This immersive perspective provides a rich understanding of the intricate spatial patterns and cell-tissue interactions. Demonstrated as a proof of concept, Spatial-Live extends the conventional 2D orthographic view to a 3D orbit perspective mode with interactive multiple-layer stacking. This tool represents a significant advancement in spatial genomics data exploration, providing greater flexibility and ease of use for researchers.

In addition, Spatial-Live was designed with a platform-agnostic approach, prioritizing a data-type oriented strategy. This design philosophy ensures easy extensibility, allowing for seamless integration of diverse data types and enhancing the flexibility of the tool. Furthermore, although Spatial-Live is specifically developed for spatial-omics data visualization, its principles and capabilities can be applied to other fields that involve data visualization and plotting within image-pixel coordinate space. If the proper input files are prepared following the guidelines, Spatial-Live can be utilized in areas such as pathological imaging and various other domains for lightweight tasks, extending its potential utility. Despite all these functionalities we implemented, Spatial-Live has ample room for improvement and can be further expanded with additional functional modules to accommodate new data types, including network-type data and more, in the future.

Materials and Methods:

Rendering variables into distinct visual layers

Irrespective of their diverse origins, datasets sourced from different spatial omics platforms are consistently preprocessed and organized into three key input files: a PNG image file, a CSV data file, and an optional JSON geometry-shape file. These files form the foundation for rendering into image-based pixel coordinates, subsequentially accommodating the four distinct data types: categorical variables, numerical variables, gene variables, and geometric shape variables (see Fig.1). The data preparation and visualization rendering in Spatial-Live mainly revolve around these four data types, ensuring that the tool is tailored to accommodate and effectively handle each of them. In line with this objective, we have established certain constraints on their formats. And by leveraging the capabilities of deck.gl [17], Spatial-Live can seamlessly generate the appropriate visual layers for diverse variables, as illustrated below:

- Categorial variable -> Scatter Layer

  Each categorical variable needs to start with the prefix of "char:", and will be rendered as a ScatterLayer in deck.gl, where each data element is represented by a small circle dot (spatially resolved spot) at given coordinates with a filled color assigned based on its categorical value. The radius of these circle points can be adjusted to match the spatial resolution of the spots.

- Numerical variable -> Column Layer

  For numerical variables, the requirement is to prefix them with "num:". During visualization, each variable can be translated into a ColumnLayer from the rendering engine, wherein each data element is portrayed as an extruded cylinder column. The height of each column is proportional to the range of numerical values. Just like in the case of scatter plotting layer, the radius of the cylinder can be adjusted to align with the spatial resolution of the spots.

- Gene variable -> Heatmap Layer

  Gene variables, bearing the obligatory prefix of "gene:", while essentially numerical in nature, are distinctively handled. This differentiation arises from the spatial resolution of spots on the image, particularly the inter-spot gap distance. In the case of gene variables, Spatial-Live generates a continuous heatmap layer through Gaussian estimation to fill those spatial gaps based on the gene expression values as weights. The process is streamlined using the fast Gaussian kernel density estimation (fast-kde [21]), ensuring efficient and accurate gene heatmap rendering.

- Geometric shape variable -> GeoJson Layer

Each geometric shape variable is associated with a JSON file that encapsulates a diverse collections of geometric shapes adhering to the GeoJSON specification [16]. These shapes are then visualized using a GeoJson Layer in rendering engine. While this layer remains optional within Spatial-Live, it can prove beneficial for incorporating custom annotations to designated regions of interest (ROI) within the image. These annotations could encompass regions enriched with cancer cells or cellular segmentation boundaries, enhancing the tool's versatility.

In addition to its exceptional rendering capabilities across multiple layers via deck.gl as mentioned above, Spatial-Live harnesses the power of several outstanding third-party JavaScript libraries. Notably, vue3 and the react framework are employed to establish the inherent logical flow and control panel, as well as a range of functionalities including orbit view mode, pan/zoom features, and color platters, enhancing the overall user experience. Comprehensive implementation details can be found from the online source codes and documentations.

10x Visium mouse kidney data and preprocessing

The spatial transcriptomic dataset for the mouse kidney (CLP model) demonstration was retrieved from GEO (GSM5224981), originating from the 10x Visium platform [18]. We processed the raw sequencing data using the 10x Space Ranger software, followed by standard procedures including quality filtering, highly variable gene identification, UMAP, and clustering assignment. By adhering to the guidelines from Spatial-Live, certain genes were chosen to represent various variables as inputs for Spatial-Live, including two numerical variables (num:Sprr1a and num:Plat) and two gene variables (gene:Cryab and gene:Mrps6), alongside one categorical variable (char:leiden). Additionally, we illustrated the process of creating JSON files, incorporating geometric polygons that adhere to the GeoJSON specification. Comprehensive instructions can be accessed in our online documentation.

Vizgen MERFISH mouse liver data and preprocessing

The spatial profiling for the mouse liver dataset (Liver1Slice1) was sourced from Vizgen's MERFISH Mouse Liver Map available at: https://info.vizgen.com/mouse-liver-access. Within this dataset, MERFISH measurements were conducted on a gene panel comprising 347 genes, spanning a substantial >300,000 liver cells situated within a single tissue slice. We downloaded essential files like cell_by_gene.csv and cell_metadata.csv, in addition to image-related components such as manifest.json and micro_to_mosaic_pixel_transform.csv. Moreover, we acquired DAPI (nuclear staining) and PolyT (RNA labeling) TIF images, which underwent subsequent processing to align with the prerequisites for the Spatial-Live image layer serving as pixel coordinate base.

Recognizing the considerable initial dimensions of the image, we opted to downscale it to a smaller size, ultimately achieving a resolution of 5989×5631. The standard single-cell data processing steps, encompassing quality filtering, clustering, and marker gene identification, were carried out using scanpy [22] and squidpy [23] scripts. Specifically, cell filtering was conducted based on marker genes persistently expressed in hepatocyte peri-central and peri-portal zones. As a result, we obtained the final output files, namely liver_demo.csv and liver_demo.png, which were meticulously tailored to align with the Spatial-Live standards. These files were seamlessly imported into Spatial-Live, facilitating their visualization. For a comprehensive understanding of the step-by-step procedure, refer to the detailed instructions available in the online documentation.

Code availability:

The source code for Spatial-Live is accessible in the GitHub repository, which can be found at https://github.com/yezhenqing/spatial-live. For comprehensive instructions and detailed documentation, please visit https://yezhenqing.github.io/spatial-live/.

Funds and Acknowledgements:

This research and this article's publication costs were supported partially by the National Institutes of Health NCI Cancer Center Shared Resources P30CA54174 to YC & ZL, NCI R50 CA265339 to ZL, and Cancer Prevention and Research Institute of Texas RP220662 to YC. The funding sources has no role in the design of the study; collection, analysis, and interpretation of data; or in writing the manuscript.

References:

1. Satija, R., Farrell, J.A., Gennert, D., Schier, A.F. & Regev, A. Spatial reconstruction of single-cell gene expression data. Nat Biotechnol 33, 495-502 (2015).
2. Conrad, T. & Altmuller, J. Single cell- and spatial 'Omics revolutionize physiology. Acta Physiol (Oxf) 235, e13848 (2022).
3. Vandereyken, K., Sifrim, A., Thienpont, B. & Voet, T. Methods and applications for single-cell and spatial multi-omics. Nat Rev Genet 24, 494-515 (2023).
4. Bressan, D., Battistoni, G. & Hannon, G.J. The dawn of spatial omics. Science 381, eabq4964 (2023).
5. Virshup, I., et al. The scverse project provides a computational ecosystem for single-cell omics data analysis. Nat Biotechnol 41, 604-606 (2023).
6. Dries, R., et al. Giotto: a toolbox for integrative analysis and visualization of spatial expression data. Genome Biol 22, 78 (2021).

7. Fernandez Navarro, J., Lundeberg, J. & Stahl, P.L. ST viewer: a tool for analysis and visualization of spatial transcriptomics datasets. Bioinformatics 35, 1058-1060 (2019).
8. Manz, T., et al. Viv: multiscale visualization of high-resolution multiplexed bioimaging data on the web. Nat Methods 19, 515-516 (2022).
9. Solorzano, L., Partel, G. & Wahlby, C. TissUUmaps: interactive visualization of large-scale spatial gene expression and tissue morphology data. Bioinformatics 36, 4363-4365 (2020).
10. Pielawski, N., et al. TissUUmaps 3: Improvements in interactive visualization, exploration, and quality assessment of large-scale spatial omics data. Heliyon 9, e15306 (2023).
11. Keller, M.S., et al. Vitessce: a framework for integrative visualization of multi-modal and spatially-resolved single-cell data. *OSF Preprints* doi:10.31219/osf.io/y8thv (2021).
12. Atta, L. & Fan, J. Computational challenges and opportunities in spatially resolved transcriptomic data analysis. Nat Commun 12, 5283 (2021).
13. Choe, K., Pak, U., Pang, Y., Hao, W. & Yang, X. Advances and Challenges in Spatial Transcriptomics for Developmental Biology. Biomolecules 13(2023).
14. Moffitt, J.R., Lundberg, E. & Heyn, H. The emerging landscape of spatial profiling technologies. Nat Rev Genet 23, 741-759 (2022).
15. Williams, C.G., Lee, H.J., Asatsuma, T., Vento-Tormo, R. & Haque, A. An introduction to spatial transcriptomics for biomedical research. Genome Med 14, 68 (2022).
16. GEO-JSON specification: https://geojson.org/
17. Deck.gl library: https://deck.gl/
18. Melo Ferreira, R., et al. Integration of spatial and single-cell transcriptomics localizes epithelial cell-immune cross-talk in kidney injury. JCI Insight 6(2021).
19. Sun, T., et al. AXIN2(+) Pericentral Hepatocytes Have Limited Contributions to Liver Homeostasis and Regeneration. Cell Stem Cell 26, 97-107 e106 (2020).
20. Moses, L. & Pachter, L. Museum of spatial transcriptomics. Nat Methods 19, 534-546 (2022).
21. Heer, J. Fast & Accurate Gaussian Kernel Density Estimation. 2021 IEEE Visualization Conference (VIS), New Orleans, LA, USA, pp. 11-15 (2021).
22. Wolf, F.A., Angerer, P. & Theis, F.J. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol 19, 15 (2018).
23. Palla, G., et al. Squidpy: a scalable framework for spatial omics analysis. Nat Methods 19, 171-178 (2022).

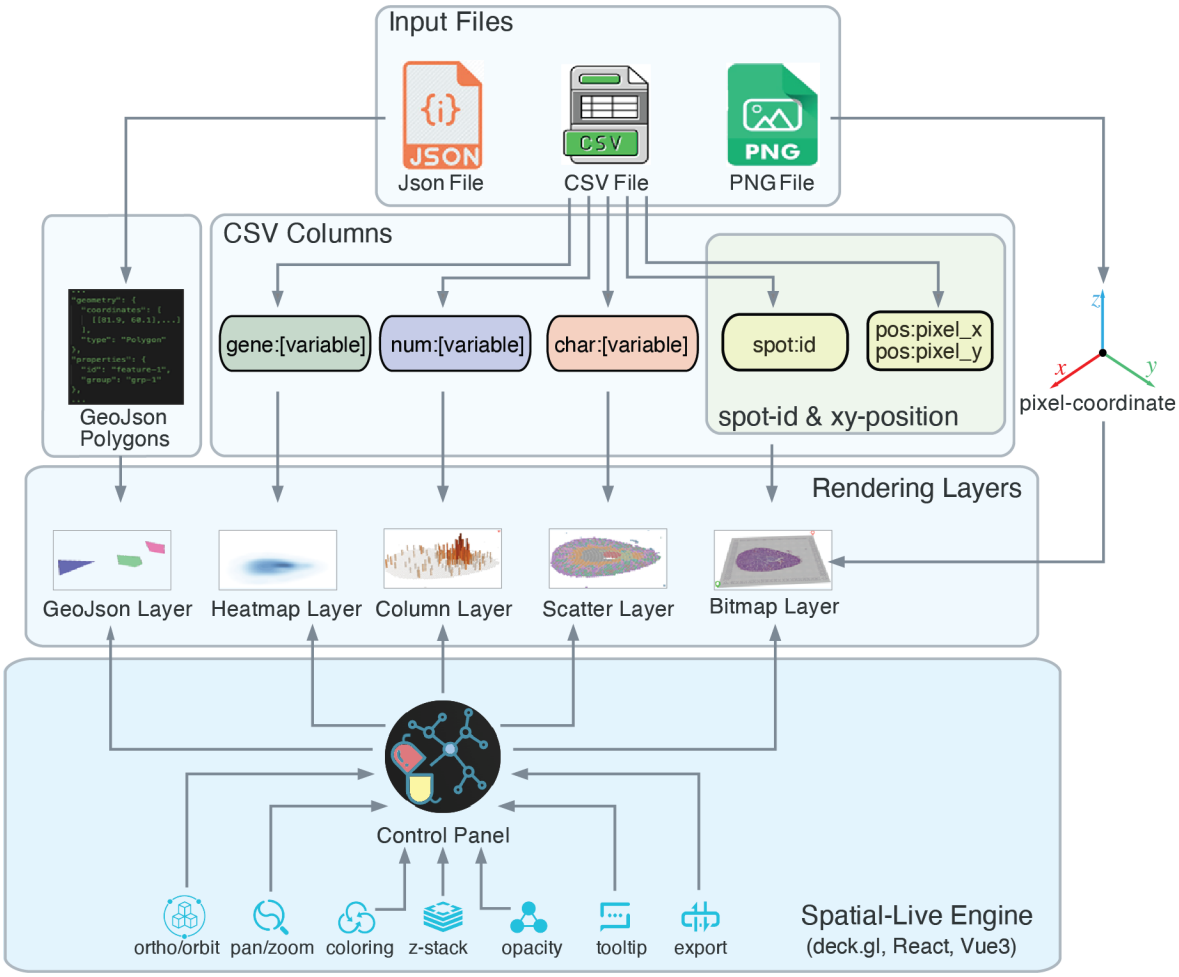

Figure 1: Spatial-Live data preparation and visualization workflow.
Spatial-Live adopts a platform-agnostic approach, transforming diverse spatial omics datasets into three essential input files: an image PNG file, a data CSV file, and an optional geometric shape JSON file. The image file establishes the pixel coordinate space, while the JSON file supports polygon shapes representing geometric variables. The CSV file categorizes data into distinct types: categorical variables, numerical variables, and gene variables. It includes 'id:spot', 'pos:pixel_x', and 'pos:pixel_y' columns, mandatory for establishing the pixel coordinate system together with the image file. Additional columns, signified by 'char:', 'num:', or 'gene:' prefixes, accommodate various data variables. The rendering engine generates corresponding visual layers, such as Scatter, Column, and Heatmap Layers, providing interactive controls like orbit view, pan/zoom, and color settings.

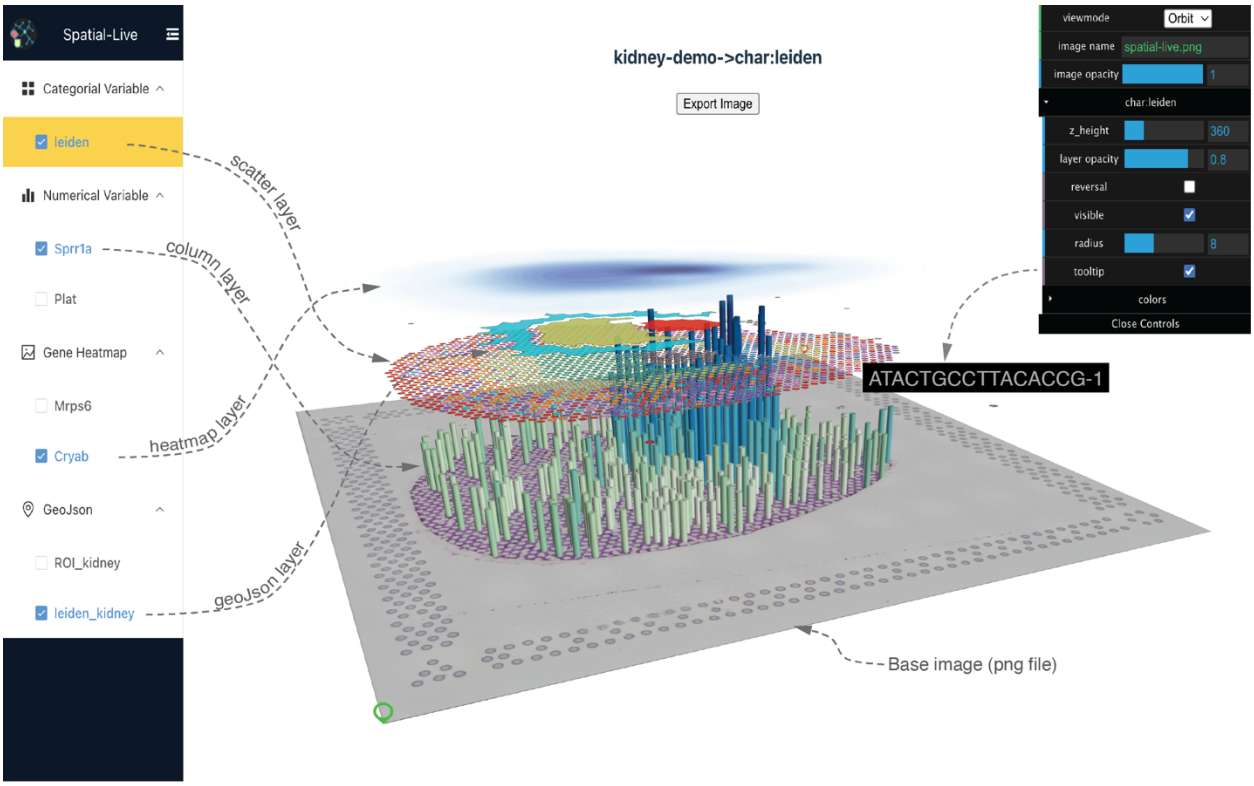

Figure 2: Spatial-Live GUI main layout.
The GUI (Graphical User Interface) of Spatial-Live comprises three integral components: the left menu pane, the central visual pane, and the right control panel. The left pane facilitates the toggling of visibility for various layers. The central pane dynamically generates visual outputs and offers functionalities such as zooming, rotating, and dragging. The top-right panel enables parameter adjustments for the active layer and optimizes the layout. The central visual output depicts a spatial transcriptomic study of mouse acute kidney injury. The thoughtful arrangement of multiple layers provides a clear view of distinct kidney regions and gene expression patterns in their spatial context. Those dotted lines are manually added here to enhance interpretation.

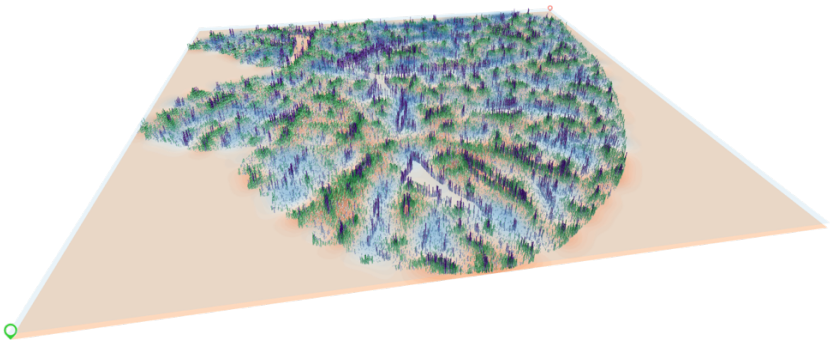

Figure 3: Multi-layer perspective visualization of mouse liver MERFISH data. An illustration showcasing Spatial-Live's capacity to convey abundant information through a single plot, employing multiple layer stacking to create a visually engaging representation of processed mouse liver MERFISH data. The light-blue and orange transparent heatmap layers denote the two gene variables 'gene:Aldh3a2' and 'gene:Hsd17b6', respectively, capturing hepatocyte peri-portal and peri-central spatial distributions while revealing a distinctive, mutually exclusive jagged pattern. The purple and green column bars correspond to the two numerical variables 'num:Vwf' and 'num:Axin2', respectively, presenting discernible spatial trends within hepatocyte zones, notably enriched Vwf expression along blood vessel borders.