Define and visualize pathological architectures of human tissues from spatially resolved transcriptomics using deep learning

Yuzhou Chang[1,*], Fei He[2,*], Juexin Wang[3,*], Shuo Chen[4], Jingyi Li[2], Jixin Liu[5], Yang Yu[2], Li Su[3], Anjun Ma[1], Carter Allen[1], Yu Lin[6], Shaoli Sun[7], Bingqiang Liu[5], Jose Otero[8], Dongjun Chung[1,9], Hongjun Fu[4], Zihai Li[9,$], Dong Xu[3,$], Qin Ma[1,9,$]

[1] Department of Biomedical Informatics, The Ohio State University, Columbus, OH 43210, USA. [2] School of Information Science and Technology, Northeast Normal University, Changchun, Jilin 130117, China. [3] Department of Electrical Engineering and Computer Science, and Christopher S. Bond Life Sciences Center, University of Missouri, Columbia, MO 65211, USA. [4] Department of Neuroscience, The Ohio State University, Columbus, OH 43210, USA. [5] School of Mathematics, Shandong University, Jinan 250100, China. [6] School of Artificial Intelligence, Jilin University, Changchun 130012, China. [7] Department of Pathology, The Ohio State University, Columbus, OH 43210, USA. [8] Departments of Neuroscience, Pathology, Neuropathology, The Ohio State University, Columbus, OH 43210, USA. [9] The Pelotonia Institute for Immuno-oncology, The Ohio State University Comprehensive Cancer Center, Columbus, OH 43210, USA. * These authors contributed equally to the paper as first authors[$] To whom correspondence: Dr. Zihai Li. Email: zihai.li@osumc.edu. Dr. Dong Xu. Tel: (573)-882-2299; Email: xudong@missouri.edu. Dr. Qin Ma. Tel: (706)-254-4293; Email: qin.ma@osumc.edu

Introduction

Background:

- Tissue architecture is the biological foundation of spatial heterogeneity within complex organs.
- Spatial transcriptomics (e.g., Visium) provide spatial information and gene expression to explore tissue domains and its biological consequences.

Challenges:

How can we visualize the tissue heterogeneity?

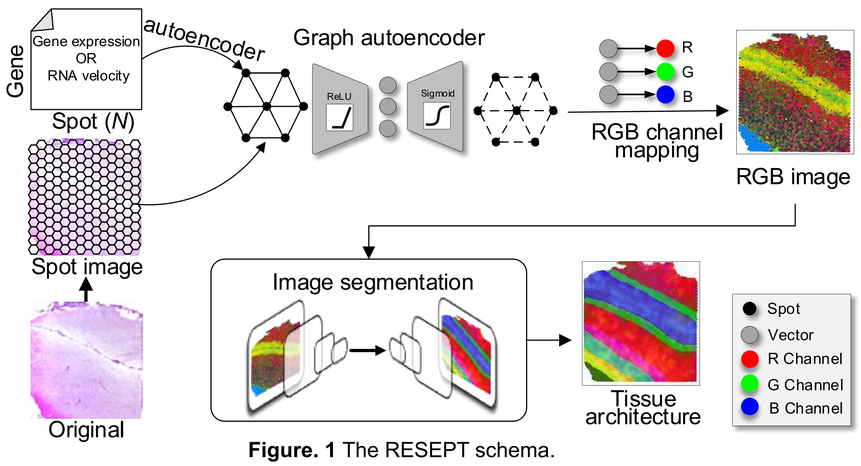

Figure. 1 The RESEPT schema.

Solution: We develop a novel deep-learning framework, RESEPT, for characterizing, visualizing, and interpreting pathological tissue architecture from spatial transcriptomics by reconstructing and segmenting a transcriptome-mapped RGB image.

Highlights and impacts:

- An RGB image represents various spatial contexts together with expression abundance faithfully, and RESEPT resists robustly to noises due to limitations of measuring technology.
- An RGB image is segmented to predict spatial architecture using a pre-trained segmentation deep-learning model and an optional segmentation quality assessment protocol.
- RESEPT can effectively distinguish cortex layers, localize essential cell types, and provide critical insights into the identification of amyloid-beta plaques in AD based on panel genes well-defined by layer-specific markers or pathological signatures.

Results

- 16 training datasets for l6-fold Jackknife cross-validation.
- S5 and S6 (healthy samples) were used for generating simulation data.
- S1 was a glioblastoma sample and was used for the case study.
- RESEPT outperforms six existing tools.
- In the downsampling read depth gradients (simulation data) from very low depth to full depth, RESEPT shows a stable performance.
- RESEPT also applies the SpaGCN embedding method as alternative spatial information retained embedding method.
- RGB images generated from RNA velocity can reveal clear spatial separation between segments from the identified architecture on the sample S4.

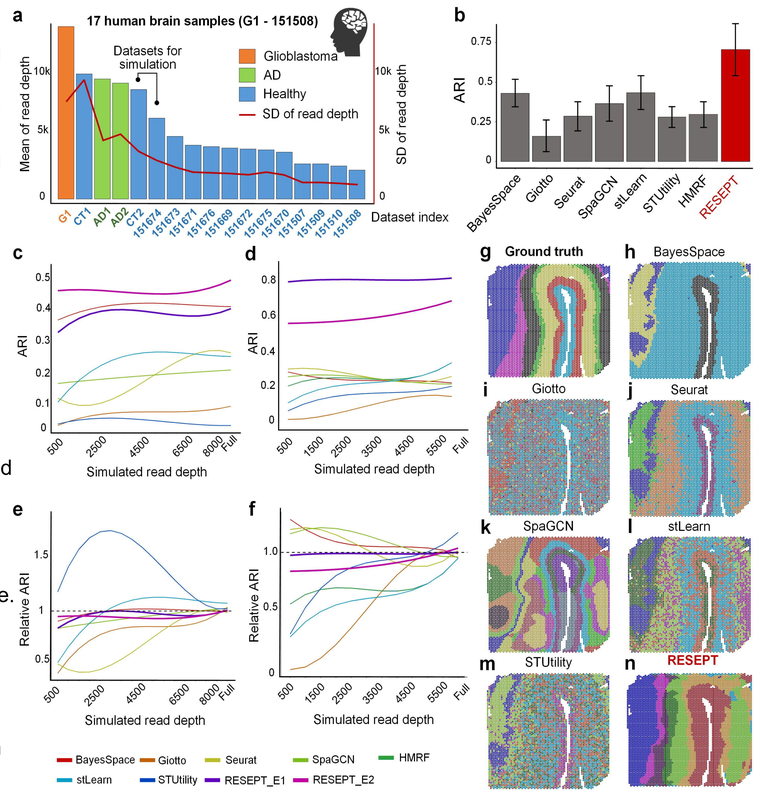

Figure. 2 Dataset and results.

Case one

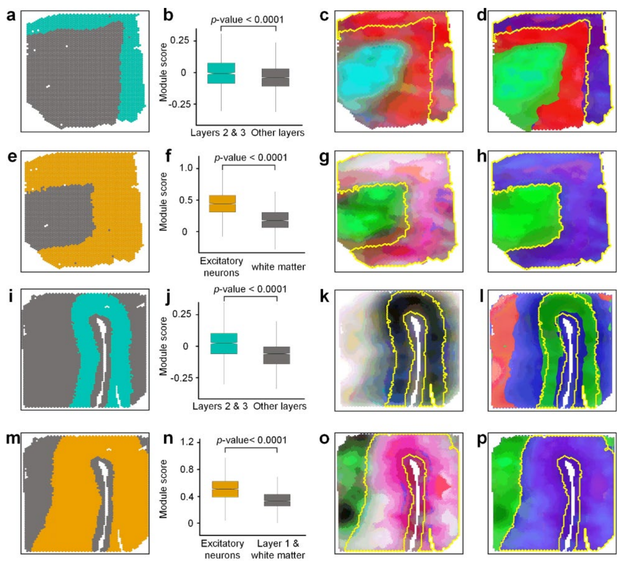

Figure. 3 Results on AD sample.

- Layer 2 and 3 can be detected via reconstructing RGB images based on pre-defined layer 2 and 3 markers.
- Excitatory neuron can be localized via reconstructing RGB image based on pre-defined excitatory neuron.
- Amyloid-beta plaques region can be predicted by RESEPT.

Conclusion:

Using in-house AD samples and well-annotated ground truth of human cortex structures, RESEPT can effectively distinguish cortex layers, localize essential cell types, and provide critical insights into the identification of amyloid-beta plaques in AD.

Case two

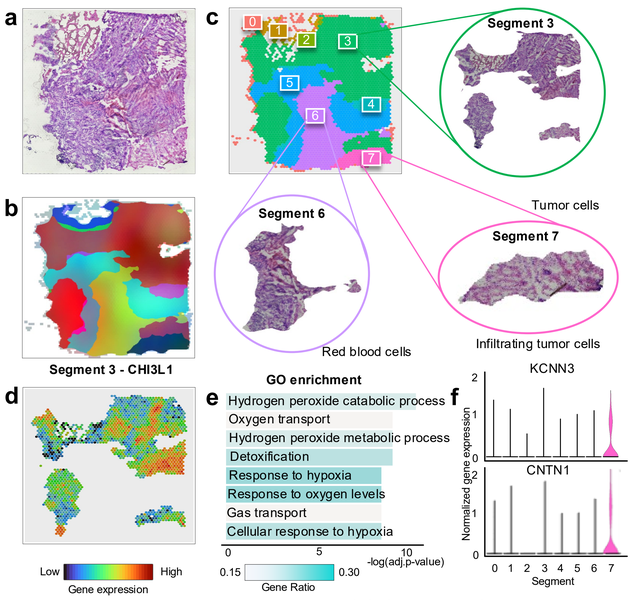

Figure. 4 Results on glioblastoma . (a) Original H&E staining image. (b) and (c) show the segmentation results. (d) Glioblastoma marker CHI3L1 expression distributes on segment 3. (e) GO pathway analysis result on segment 6. (f) two infiltering tumor markers express on segment 7.

Conclusion:

With the support of two clinical pathologists, RESEPT successfully distinguished tumor-enriched, non-tumor, and regions of neuropil with infiltrating tumor cells in a glioblastoma sample, which shows the value for clinical and prognostic applications.

Discussion and future work

- Can be used for spatial transcriptomics data analysis, visualization, and interpretation.
- RESEPT allows taking one of the two types of input, gene expression or RNA velocity.
- The RGB embedding can be used for feature-centric analysis using SpaGFT (QR code on the right), such as correlating genes and RGB channels at Fourier Space.

Paper Links

SCAN ME

GitHub Links

SCAN ME

Acknowledgement

Funding: This work was supported by an NIH R01 award #R01GM131399-01 and #R35-GM126985; This work used Ohio Supercomputer Center (OSC) and the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by the National Science Foundation grant number ACI-1548562. This work was supported by the Pelotonia Institute of Immuno-Oncology (PIIO). The content is solely the responsibility of the authors and does not necessarily represent the official views of the PIIO.