Histological images by parallel computing on Apache Spark

**Abstract**

Technology is developing rapidly today, including computer science and medicine. Parallel computing has the advantages of high speed, high throughput, and scalability, while histopathological image analysis is the gold standard for cancer identification and diagnosis. This project will be based on Apache Spark, using computer science to help histological images recognition, making doctors' judgment of abnormal tissues more efficient and accurate.

**References**

1. Khened, M., Kori, A., Rajkumar, H. et al. A generalized deep learning framework for whole-slide image segmentation and analysis. Sci Rep 11, 11579 (2021). <https://doi.org/10.1038/s41598-021-90444-8>
2. Qin, P., Chen, J., Zeng, J. et al. Large-scale tissue histopathology image segmentation based on feature pyramid. J Image Video Proc. 2018, 75 (2018). https://doi.org/10.1186/s13640-018-0320-8
3. Gudhe, N.R., Kosma, VM., Behravan, H. et al. Nuclei instance segmentation from histopathology images using Bayesian dropout based deep learning. BMC Med Imaging 23, 162 (2023). https://doi.org/10.1186/s12880-023-01121-3