**[BIOINFO 2024]**

# ConFuseIT: Contrastive Learning-based Pretraining Model for Integrating Image and Transcriptomics for Multi-Task Prediction in Breast Cancer

Suwan Yu[1], Yooeun Kim[1], Hoeyoung Kim[1], Sangseon Lee[2] and Kwangsoo Kim[3,4,*]

[1] *Interdisciplinary Program in Bioinformatics, College of Natural Sciences, Seoul National University*
[2] *Department of Artificial Intelligence, Inha University*
[3] *Department of Medicine, College of Medicine, Seoul National University,*
[4] *Department of Transdisciplinary Medicine, Institute of Convergence Medicine with Innovative Technology, Seoul National University Hospital*
* *Corresponding author: kwangsookim@snu.ac.kr*

This study introduces a multimodal deep learning model to predict survival and explore molecular characteristics in breast cancer (BRCA) patients by integrating pathology images and transcriptomic data. Utilizing TCGA datasets, the model employs pre-training with a contrastive learning model based on CLIP, followed by fine-tuning to predict survival status and classify PAM50 subtypes. By comparing survival prediction performance across single-modal (Whole Slide Image or RNA-seq) and multimodal data, the study demonstrates a clear trend of improved performance with multimodal integration. Additionally, an attention-based multiple instance learning (ABMIL) model incorporating cell type-related, deep learning-derived features enhances survival prediction and molecular classification performance. This approach overcomes significant limitations of current multimodal models, which typically depend on the availability of both data modalities and labeled survival information to perform effectively. By leveraging pre-training to capture data distributions without relying on labels, the developed model can be applied to single-modal datasets, enhancing its adaptability and practical utility across diverse research contexts. These advancements establish a robust framework for improving diagnostic models and survival predictions in oncology, highlighting the potential of multimodal data to enhance personalized cancer prognosis.