

# Overview of the potentials of multiple instance learning in cancer diagnosis: Applications, challenges, and future directions

Tagne Poupi Theodore Armand\*, Subrata Bhattacharjee\*\*, and Hee-Cheol Kim\*

\*Institute of Digital Anti-Aging Healthcare, Inje University, Gimhae 50834, Republic of Korea

\*\*Department of Computer Engineering, u-AHRC, Inje University, Gimhae 50834, Republic of Korea

poupiarmand2@gmail.com, subrata\_bhattacharjee@outlook.com, heeki@inje.ac.kr

**Abstract**— *The outcome of cancer patients mostly depends on the diagnosis process and the treatment strategies. Computer-aided diagnosis (CAD) methods have demonstrated the potential to handle accurate diagnostics using artificial intelligence techniques such as machine learning and deep learning. The nature of the data used in training the AI-based model determined the paradigm, often classified as supervised and unsupervised learning for scenarios with labeled and unlabeled data, respectively. Due to the cost of time and resources, most datasets are nowadays partially labeled and used for training. The weakly supervised learning approach enables the AI models to be trained with incompletely labeled, noisy, or imbalanced data. In recent years, multiple instance learning (MIL) has emerged as a promising weakly supervised learning approach in many fields, including cancer diagnosis. Unlike traditional supervised learning methods, MIL allows the classification of groups of instances, known as bags, where only the bag's label is available. This comprehensive review aims to provide an in-depth analysis of the applications of MIL in cancer diagnostic tasks, highlighting its advantages, challenges, and future directions. By examining these advantages, challenges, and future trends, the review aims to contribute to advancing MIL as a powerful tool in improving cancer diagnostic accuracy and patient outcomes.*

**Keywords**— *Bag, Cancer diagnosis, Instance, Multiple Instance Learning, Weakly supervised learning.*



**Tagne Poupi Theodore Armand** was born in Cameroon, received Msc in information System and networking at ICT University USA, Cameroon Campus in 2021. Currently, he is a Ph.D. research scholar at the Institute of Digital Anti-aging and healthcare at Inje University. His research interest field includes image processing with a focus on medical image analysis, Deep Learning, Machine Learning and Metaverse.



**Subrata Bhattacharjee** is pursuing his M.S. leading Ph.D. from the Graduate School of Computer Engineering, Inje University, Korea. He is a Research Scholar and Teaching Assistant at the Medical Image Technology Laboratory (MITL) at Inje University. His research interests include multimodal medical imaging, tissue and cell image analysis, digital pathology, image reconstruction, cell and gland segmentation, machine learning, and deep learning. He has some SCIE paper publications and has participated in International and Domestic conferences.



Hee-Cheol Kim received his BSc at the Department of Mathematics, MSc at the Department of Computer Science at SoGang University in Korea, and Ph.D. in Numerical Analysis and Computing Science at Stockholm Stockholm University in Sweden in 2001. He is a professor at the Department of Computer Engineering and Head of the Institute of Digital Anti-aging Healthcare Inje University in Korea. His research interests include machine learning, deep learning, Computer vision, and medical informatics.