

INTERVENTION-BASED MIXUP FOR MULTIPLE INSTANCE LEARNING IN WHOLE-SLIDE SURVIVAL ANALYSIS

Author's: Subash K¹, N. Vanjulavalli², K. Sujith³

Abstract

Whole-Slide Image (WSI) survival analysis has emerged as a powerful computational pathology approach for predicting patient prognosis directly from digitized histopathology slides. However, WSIs are extremely high-resolution images that are typically partitioned into thousands of smaller patches, creating a weakly supervised learning scenario in which only slide-level survival labels are available. Multiple Instance Learning (MIL) frameworks have become a standard solution for this setting, where each slide is treated as a bag of instances (patch embeddings) and the model learns to aggregate informative regions for risk prediction. Despite recent advances in attention-based MIL models, a key limitation remains: conventional data augmentation strategies, including standard mixup, operate uniformly across instances and may dilute critical tumor-related signals by blending salient and non-salient patches indiscriminately. To address this challenge, we propose an Intervention-Based Mixup Augmentation framework tailored specifically for MIL-based WSI survival modeling. The core idea is to incorporate structural interventions at the instance level guided by learned attention weights. Instead of mixing entire bags uniformly, the proposed method selectively mixes or swaps subsets of instances based on their semantic importance. In particular, we introduce three intervention mechanisms: (1) top-k salient patch mixup, which mixes highly attended tumor-relevant patches between slides; (2) bottom-k background mixup, which perturbs non-informative regions while preserving critical signals; and (3) salient-instance swapping, which performs targeted structural interventions to encourage robustness against regional variability. By operating in this attention-aware manner, the augmentation preserves meaningful prognostic patterns while improving model generalization. The survival objective is optimized using the Cox proportional hazards partial likelihood, enabling direct modeling of time-to-event data with censoring. Synthetic WSI bag data are generated to simulate heterogeneous tumor fractions, attention distributions, and survival risk associations. Experimental evaluation demonstrates that intervention-based mixup enhances concordance index (C-index) performance compared to baseline MIL models without augmentation. Furthermore, attention analysis confirms improved alignment between high-attention regions and prognostically relevant patches, indicating better interpretability.