Supplementary Information

Boundary-Stabilized Oversampling for Interpretable Clinical Prediction in Imbalanced and Data-Scarce Regimes

Seong-Hyun Yu[1,4,†], **Minha Kim**[2,†], **Inhyung Lee**[2], **Won-Gyun Son**[2,*], **and Ji-Hoon Jeong**[1,3,4,*]

[1]Sungkyunkwan University, SAIHST, Dept. of Medical Device Management and Research, Seoul, 06355, Republic of Korea
[2]Seoul National University, Dept. Veterinary Clinical Sciences, College of Veterinary Medicine, Seoul, 08826, Republic of Korea
[3]Samsung Medical Center, Research Institute for Future Medicine, Seoul, 06355, Republic of Korea
[4]Sungkyunkwan University, Medical Imaging Insight Institute, Seoul, 06355, Republic of Korea
[*] *Corresponding authors:* carpeego@snu.ac.kr, jihoon.jeong@skku.edu
[†] *These authors contributed equally to this work*

Supplementary Table S1

This supplementary table presents the complete results of the 10-way ablation study evaluating all combinations of five augmentation strategies (No Augmentation, Standard SMOTE, CTGAN, VAE, and BS-SMOTE) with fixed and fine-tuned GradientBoosting classifiers under stratified 5-fold cross-validation. Performance metrics including F1, AUC, Cohen's κ, sensitivity, and specificity are reported as mean values with 95% confidence intervals. These results support the comparative analysis presented in the main text and demonstrate that BS-SMOTE with fine-tuned GradientBoosting (BS-SMOTE + GB F/T) consistently achieved the highest performance across all primary metrics.

Supplementary Table S1. Ten-way ablation study comparing five augmentation strategies across fixed and fine-tuned GradientBoosting configurations under stratified 5-fold cross-validation. F1 and AUC are reported as mean with 95% confidence interval. GB: fixed hyperparameters ($n_{\text{estimators}} = 200$, max_depth$= 3$, lr$= 0.05$, subsample$= 0.8$). GB F/T: fine-tuned via nested 3-fold grid search ($n_{\text{estimators}} \in \{100, 200, 300\}$, max_depth $\in \{2, 3, 4\}$, lr $\in \{0.01, 0.05, 0.1\}$, subsample $\in \{0.7, 0.8, 1.0\}$).

Method	**F1**	**F1 (95% CI)**	**AUC**	**AUC (95% CI)**	**Kappa**	**Kappa (95% CI)**	**Sens.**	**Spec.**
No Augmentation								
No Aug + GB	0.100	0.096–0.296	0.723	0.610–0.835	0.064	0.138–0.266	0.067	**0.976**
No Aug + GB F/T	0.080	0.077–0.237	0.709	0.618–0.800	0.046	0.101–0.193	0.067	**0.976**
Standard SMOTE								
SMOTE + GB	0.147	0.031–0.324	0.706	0.604–0.809	0.096	0.069–0.261	0.133	0.959
SMOTE + GB F/T	0.147	0.031–0.324	0.696	0.588–0.803	0.088	0.083–0.259	0.133	0.951
CTGAN								
CTGAN + GB	0.100	0.096–0.296	0.608	0.458–0.759	0.057	0.150–0.264	0.067	0.968
CTGAN + GB F/T	0.260	0.066–0.586	0.705	0.642–0.769	0.232	0.100–0.563	0.200	0.984
VAE								
VAE + GB	0.100	0.096–0.296	0.663	0.581–0.745	0.052	0.155–0.259	0.067	0.968
VAE + GB F/T	0.213	0.056–0.482	0.697	0.646–0.749	0.171	0.095–0.437	0.200	0.967
BS-SMOTE (Proposed)								
BS-SMOTE + GB	0.280	0.031–0.529	0.709	0.619–0.799	0.221	0.032–0.474	0.267	0.951
BS-SMOTE + GB F/T	**0.307**	**0.016–0.598**	**0.728**	**0.655–0.801**	**0.256**	**0.045–0.557**	**0.267**	0.959

† CI lower bounds for F1 and Kappa may extend below zero due to normal approximation under small-sample conditions.