Nonlinear Compression Augmentation for Generalizable 12-Lead ECG Classification Across Multi-Site Datasets

Juhee Lee[1], Minseo Go[2], Bomin Kim[3], Dillon J Dzikowicz[4], Ben Bailey[4], Junmo An[4]

[1]Soonchunhyang University, Asan, Republic of Korea
[2]Chosun University, Gwangju, Republic of Korea
[3]Korea University, Seoul, Republic of Korea
[4]Philips, Cambridge, MA, USA

Data augmentation can improve automatic electrocardiogram (ECG) classification by enhancing model generalizability and robustness to variations across datasets. Among nonlinear compression methods, prior work has primarily focused on sigmoid-based transformations, with limited investigation of alternative nonlinear compression strategies. While sigmoid compression has shown improved performance over baseline models, the broader effectiveness of diverse nonlinear compression methods, particularly for cross-dataset generalization, remains underexplored. This study systematically evaluates multiple nonlinear compression strategies for both in-domain and cross-dataset performance.

We evaluated multiple nonlinear compression methods, including Sigmoid, Tanh, Arctan, Softsign, Logarithmic, Power, Huber-style soft clipping, and Adaptive Tanh compression. A 1D ResNet-based model was trained to classify nine diagnostic classes (NSR, AF, IAVB, LBBB, RBBB, PAC, PVC, STD, STE). In-domain performance was assessed using 10-fold cross-validation on the China-based dataset (CPSC, 6,877 records). For cross-dataset evaluation, models were trained on CPSC and tested on the US-based dataset (Georgia, 4,175 records), restricting evaluation to seven shared classes (excluding STD and STE). Performance was measured using the F1 score.

In-domain, all nonlinear compression methods outperformed the baseline (F1 = 0.737). Tanh achieved the highest F1 (0.755, +2.51%), followed by Power (0.754, +2.36%) and Huber-style clipping (0.753, +2.15%). In cross-dataset evaluation, most methods improved generalization. Sigmoid achieved the highest F1 (0.590, +1.73%), followed by Tanh (0.586, +1.02%). Huber-style compression showed a slight performance degradation (F1 = 0.580, −0.12%), indicating limited transferability.

Nonlinear compression improves ECG classification performance in both in-domain and cross-dataset evaluations; however, performance gains are not consistently aligned across settings. Improvements observed under in-domain evaluation may not reflect performance under distribution shifts, emphasizing the need for cross-dataset validation when developing robust models for real-world clinical deployment. These findings suggest that nonlinear compression improves robustness by reducing amplitude variability and mitigating outlier-driven signal distortions across heterogeneous acquisition systems.