

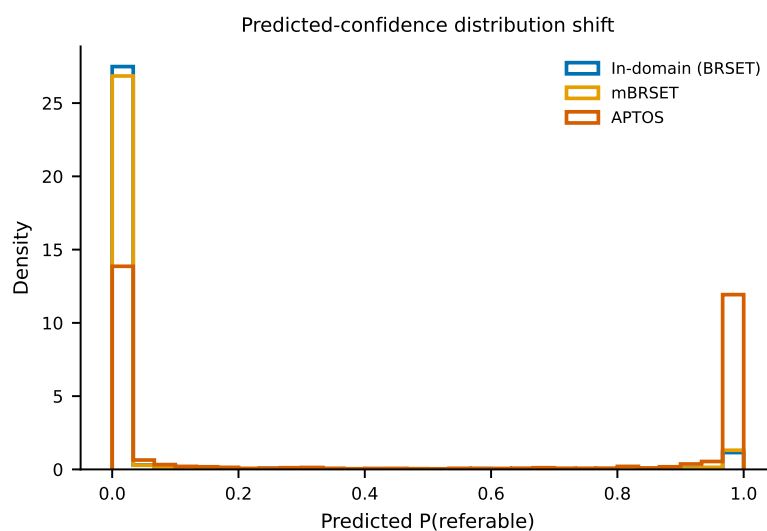
Supplementary Information

Discrimination transfers but calibration does not: prevalence-dominated miscalibration of diabetic retinopathy screening across fundus cameras

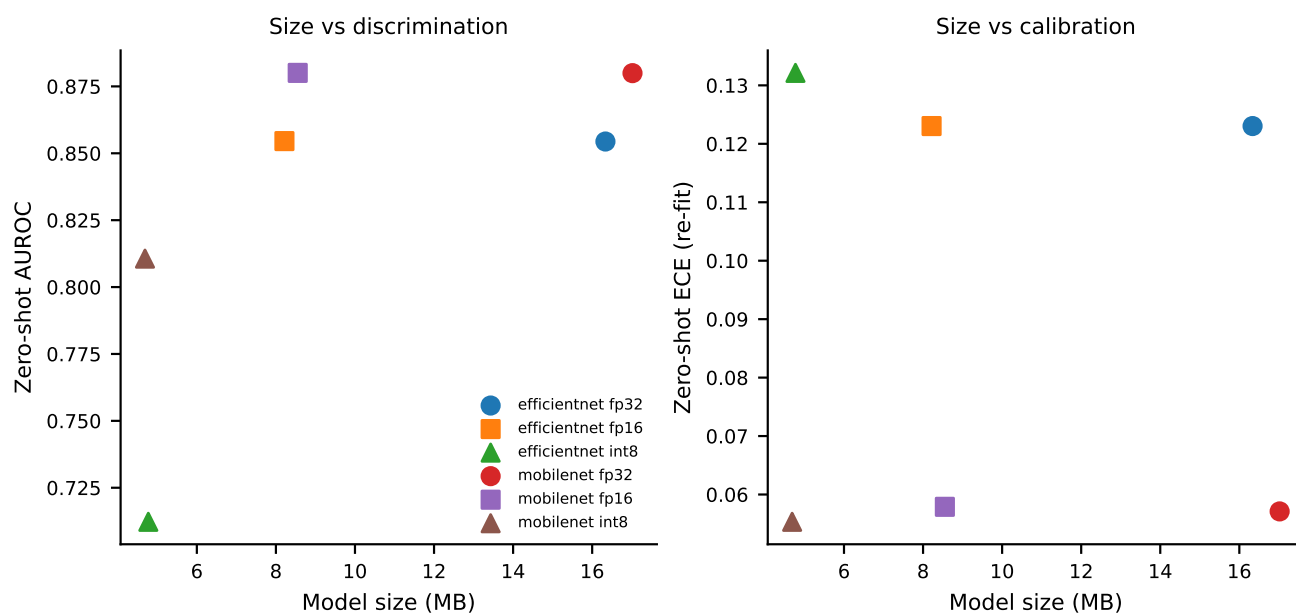
Sairam Tabibu, Mugdha Abhyankar

ABSTRACT

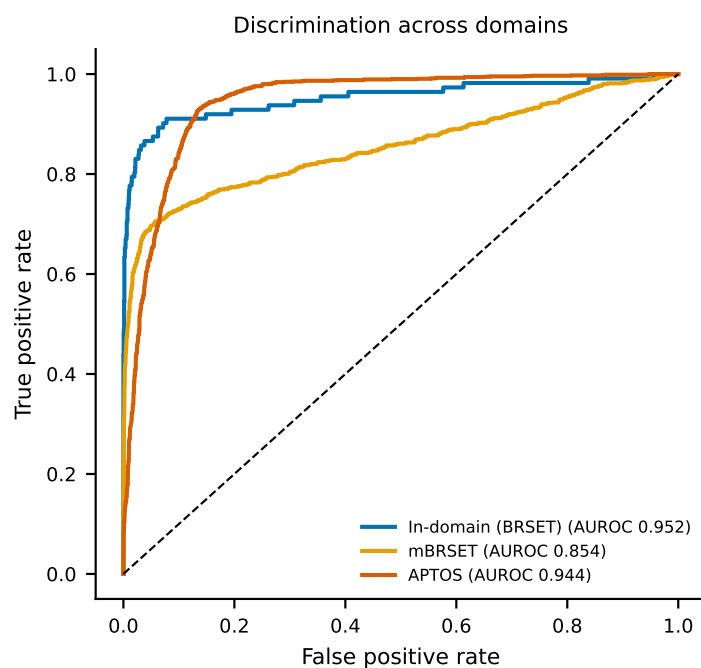
This file contains Supplementary Figures S1–S3 referenced in the main text.



Supplementary Figure 1. Predicted-confidence distribution shift. Density of the predicted probability of referable DR for the in-domain (BRSET) test set and the two zero-shot targets (mBRSET, APTOS), EfficientNet-B0, representative seed. The score distribution shifts markedly across acquisition settings, consistent with the calibration break quantified in the main text.



Supplementary Figure 2. Precision–reliability trade-off. Zero-shot discrimination (left) and calibration (right, re-fit calibrator) versus on-disk model size for FP32, FP16 and INT8. FP16 preserves both properties at half size; INT8 sharply reduces discrimination (which monotone recalibration cannot recover) even after re-fitting the calibrator.



Supplementary Figure 3. Discrimination across domains. Receiver-operating- characteristic curves for the in-domain (BRSET) model applied in-domain and zero-shot to mBRSET and APTOS, EfficientNet-B0, representative seed. Areas under the curve match Table 1 of the main text; discrimination is largely preserved on both external targets.