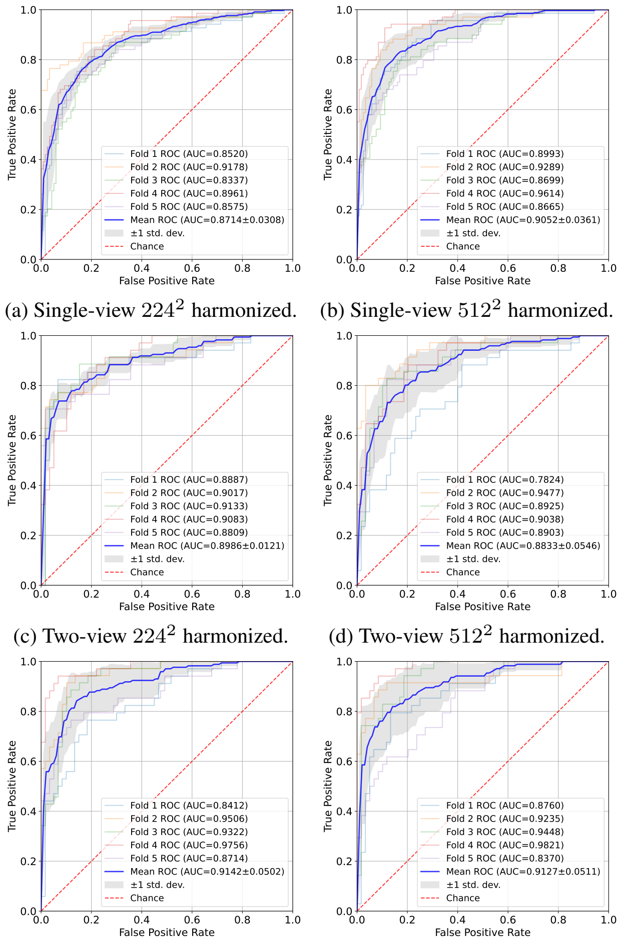
LUMINA: A Multi-Vendor Mammography Benchmark with Energy Harmonization Protocol

Supplementary Material



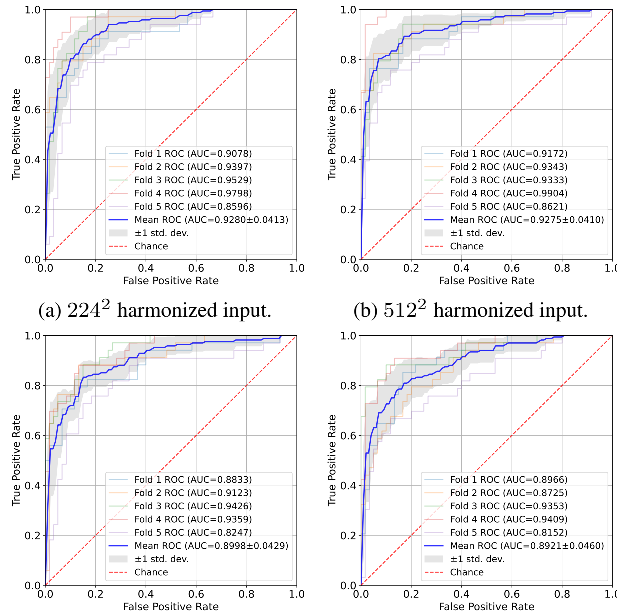
(a) Single-view 224^2 harmonized. (b) Single-view 512^2 harmonized.

(c) Two-view 224^2 harmonized. (d) Two-view 512^2 harmonized.

(e) Two-view 224^2 raw input. (f) Two-view 512^2 raw input.

Figure 9. **ROC curves of the EfficientNet-B0 model for breast cancer diagnosis.**

Fig. 9 shows the ROC curves of EfficientNet-B0 for the breast cancer diagnosis task, where we plot the ROC for each fold and the mean ROC. The two-view model consistently outperforms the single-view variant, demonstrating the benefit of combining CC and MLO information. Higher input resolution (512^2) further improves AUC compared to the 224^2 setting. In all configurations, object-histogram-based harmonization yields a noticeable upward shift in the ROC curves, confirming its effectiveness in stabilizing image appearance and enhancing diagnostic performance. Fig. 10 shows the ROC curves of EfficientNet-B0 for the two-class BI-RADS classification task. It also supports that higher resolution and harmonization improves the results.



(a) 224^2 harmonized input. (b) 512^2 harmonized input.

(c) 224^2 raw input. (d) 512^2 raw input.

Figure 10. **ROC curves of the EfficientNet-B0 model for two-class BI-RADS classification.**

Table 8 lists the distinguishing characteristics of LUMINA compared to existing mammography datasets and prior harmonization strategies.

Table 9 summarizes the unified hyperparameter settings and pre-processing operations used across all model architectures, respectively. These settings and operations ensured a controlled and fair comparison in our benchmark experiments.

Table 8. **Positioning.** LUMINA provides pathology, BI-RADS, and density on multi-vendor FFDM, plus a simple, effective pixel-space harmonization baseline.

Family	Examples	Labels/Supervision	Architecture/Objective	Limitations vs. LUMINA
Film (SFM) datasets	MIAS [31], DDSM [4], CBIS-DDSM [14]	Pathology (yes), BI-RADS (varies)	Early CNNs / CAD; single-task classification	Film scans, lower resolution; limited modern FFDM relevance.
Digital (FFDM) datasets	INbreast [20], VinDR-Mammo [23], RSNA [6], CMMD [5], KAU-BCMD [1]	Pathology or BI-RADS; density varies	CNNs/ViTs; single- or dual-task	Often single vendor/system; partial labels; limited multi-task scope.
Harmonization (feature space)	ComBat [22]	Batch-effect correction on features	Empirical Bayes (location/scale) on radiomics/latent features	Requires feature extraction; not pixel-space; modality-agnostic assumptions.
Harmonization (federated learning)	HarmoFL [12]	Federated frequency-domain drift normalization	Joint optimization across clients	Heavy infra; task/model coupling; not a simple preproc.
LUMINA (Ours)	Multi-vendor FFDM (6 systems); energy metadata	Pathology + BI-RADS + density	*Foreground-only* pixel-space CDF matching + unified 3-task benchmark; single-/two-/four-view baselines	Vendor/energy diversity; three tasks; consistent harmonization gains; improved attention focus.

Table 9. **Hyperparameter parity across backbones.** Shared schedule and augmentations ensure a fair comparison (see Sec. 6.1).

Backbone	Pre-trained	Learning Rate	Weight Decay	Epochs	Learning Rate scheduler	Batch	Augmentation
ResNet-50	ImageNet-1K	1×10^{-3}	1×10^{-5}	100	$\times$0.1 every 30	32	flip, resize
DenseNet-121	ImageNet-1K	1×10^{-3}	1×10^{-5}	100	$\times$0.1 every 30	32	flip, resize
EfficientNet-B0	ImageNet-1K	1×10^{-3}	1×10^{-5}	100	$\times$0.1 every 30	32	flip, resize
Swin-T	ImageNet-1K	1×10^{-5}	1×10^{-5}	100	$\times$0.1 every 30	32	flip, resize

Table 10. **Shared pre-processing.** Ensures identical inputs across backbones; isolates model differences.

Stage	Operation (applied to all models equally)
DICOM handling	Convert MONOCHROME1 to MONOCHROME2; remove text burn-ins.
Foreground mask	Define $M = \{(x,y) \mid \mathbf{I}(x,y) > 0\}$; exclude background from histograms.
Harmonization	Foreground CDF matching to low-energy reference (Eqs. 1–8); training-free, model-agnostic.
Resize/replicate	Resize to 224^2 or 512^2; replicate grayscale to 3 channels for ImageNet backbones (Sec. 6.1).
Augment	Random horizontal flip (all backbones, all tasks).
Normalization	As in backbone defaults; no per-model special tuning beyond LR noted in Sec. 6.1.