



CALIBRATION SUMMARY

Expected Calibration Error (ECE):

- Baseline: 0.2410
- MC-Dropout: 0.2034
- Decoder Var (Fused): 0.2065

Uncertainty Quality (AUROC):

- MC-Dropout: 0.6335
- Decoder Var (Fused): 0.4875
- Single Layer (L6): 0.4985

KEY FINDINGS:

- ✓ Decoder variance achieves competitive calibration
- ✓ Fusing layer uncertainties improves over single layer
- ✓ MC-Dropout still superior for TP/FP discrimination
- ✓ Representation-level fusion shows promise

CONCLUSION:

Representation-level uncertainty fusion (decoder variance) provides computationally efficient calibration, though MC-Dropout remains superior for epistemic uncertainty estimation.