

LS+: Informed Label Smoothing for Improving Calibration in Medical Image Classification

Abhishek Singh Sambyal(✉)¹, Usma Niyaz¹, Saksham Shrivastava¹,
Narayanan C. Krishnan², and Deepti R. Bathula¹

¹ Department of Computer Science & Engineering, Indian Institute of Technology
Ropar, Punjab, India

{abhishek.19csz0001, usma.20csz0015, 2022aim1012, bathula}@iitrpr.ac.in

² Department of Data Science, Indian Institute of Technology Palakkad, Kerala, India
ckn@iitpkd.ac.in

1 Supplementary Material

1.1 Retention Curves for ResNet-50

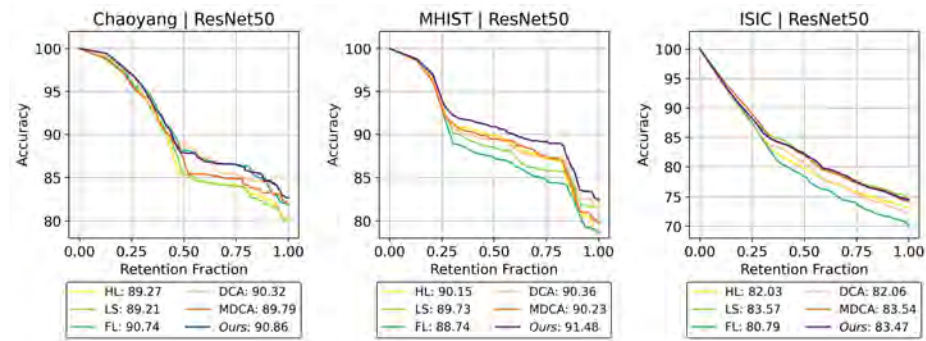
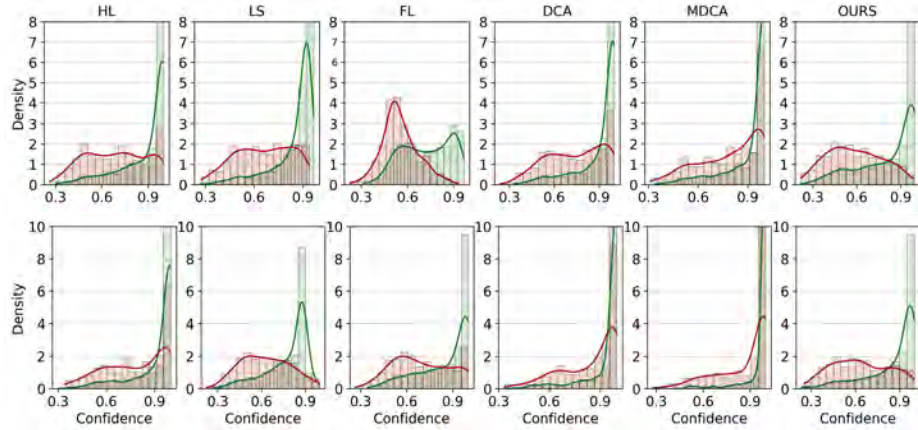
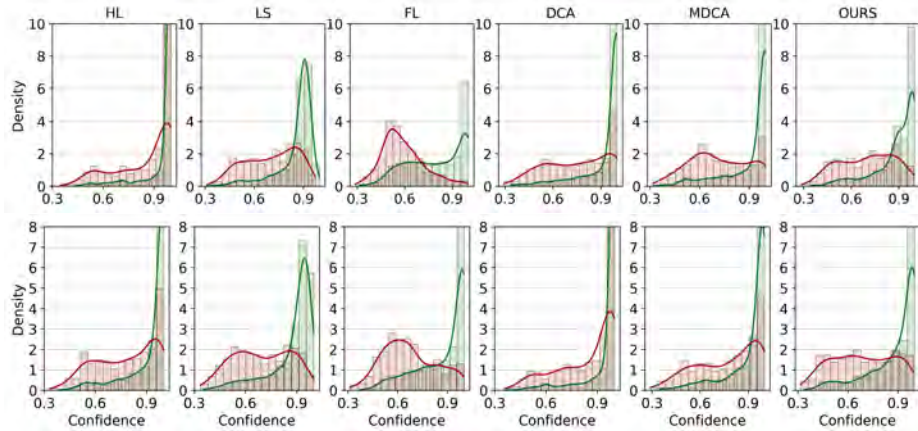


Fig. 1: **Retention Curves.** Accuracy as a function of retention fraction along with the area under the retention curve (R-AUC) values using ResNet-50 for all three datasets. HL - *Hard Labels*, LS - *Label Smoothing*, FL - *Focal Loss*, DCA - *Difference between Confidence and Accuracy* and MDCA - *Multi-class Difference in Confidence and Accuracy*.

1.2 Density Plots for Correct and Incorrect Classification Confidences



(a) ISIC Dataset.



(b) Chaoyang Dataset.

Fig. 2: Density plots for correct (green) and incorrect (red) classification confidences for ResNet-34 (top) and ResNet-50 (bottom) on ISIC and Chaoyang datasets. For incorrect predictions, LS, FL and Ours provide low confidence which is desirable. However, methods like HL, DCA and MDCA exhibits higher confidence even when they are wrong making them unreliable.