

Supplements: Identity Curvature Laplace Approximation for Improved Out-of-Distribution Detection

Table 1. Computational costs of ICLA and LLLA. The results for initialization are reported in seconds per initialization and seconds per 1000 batches of size 64 for the inference stage.

	Inference		
	CIFAR-10	CIFAR-100	ImageNet
EF	4.69	10.81	10.68
GGN	4.77	10.51	10.56
KFAC	5.11	30.22	29.31
ICLA	4.65	10.70	10.41
	Initialization		
	CIFAR-10	CIFAR-100	ImageNet
EF	19.71	20.68	21.38
GGN	19.28	32.6	37.07
KFAC	19.38	21.44	20.46
ICLA	0.02	0.02	0.05

A. Marginal Likelihood Algorithm

We present a detailed definition of the `marginal_likelihood` function in Algorithm 1.

Algorithm 1 Marginal Likelihood

Input: Dataset $\mathcal{D} = \{x_i, y_i\}_{i=1}^N$, neural network f , Hessian \mathcal{H} , learning rate α , number of epochs T .

Output: Prior precision λ for Equation 7.

- 1: Initialize λ
 - 2: **for** $t = 1, \dots, T$ **do**
 - 3: $h_t \leftarrow \nabla_{\lambda^2} \log(p(\mathcal{D}|f, \lambda^2))$
 - 4: $\lambda_t^2 \leftarrow \lambda_{t-1}^2 + \alpha h_t$
 - 5: **end for**
 - 6: **return** λ
-

B. Computational performance comparison

We report computational performance comparison in Table 1. All the comparisons are conducted on a single A100 80GB GPU. It can be seen that our ICLA implementation doesn't impose any additional computational overhead on

inference and takes near zero time for initialization (Algorithm 1).

C. Discussion on the connection of ICLA and NECO

NEural Collapse-based Out-of-distribution (NECO [2]) has a theoretical framework based on class separability and analyzing eigenvalues of covariance matrix in embedding space. Although some of these aspects might seem related, our analysis method focuses on connecting class separability with the Fisher matrix structure and model curvature, which allows us to put our approach into the field of Laplace approximation and Bayesian methods.

D. Calibration Metrics

In this section, we list performance measurements for calibration. Firstly, let B_t be a batch of samples, whose confidences lie in the interval $(\frac{t-1}{T}, \frac{t}{T}]$, where T is the number of bins we split the prediction by and m is the bin index. We define accuracy and confidence as

$$\text{acc}(B_t) = \frac{1}{|B_t|} \sum_{i \in B_t} \mathbb{I}(f(x_i) = y_i). \quad (1)$$

$$\text{conf}(B_t) = \frac{1}{|B_t|} \sum_{i \in B_t} p_i. \quad (2)$$

D.1. Expected Calibration Error

The expected calibration error (ECE) can be estimated as

$$\text{ECE}(B_t) = \sum_{t=1}^T \frac{|B_t|}{N} |\text{acc}(B_t) - \text{conf}(B_t)| \quad (3)$$

where N is the number of samples.

D.2. Negative Log-Likelihood

Negative log-likelihood typically coincides with cross-entropy and is computed as

$$\text{NLL} = - \sum_{i=1}^N \log(f(x_i)). \quad (4)$$

D.3. Brier Score

Brier score is another calibration measure and is expressed as

$$\text{Brier} = \frac{1}{N} \sum_{i=1}^N (f(x_i) - y_i)^2. \quad (5)$$

E. Details on Hessian Approximations

In this section, we elaborate on Hessian approximations and their formulations.

E.1. Generalized Gauss-Newton Matrix (GGN)

$$G \triangleq \sum_{i=1}^N J(x_i) (\nabla_{\theta}^2 \log p(y_i | f_{\theta})|_{\theta_{\text{MAP}}}) J(x_i)^T, \quad (6)$$

where $J(x) \triangleq \nabla_{\theta} \log(p(\mathcal{D}|\theta))|_{\theta_{\text{MAP}}}$ is the Jacobian matrix.

E.2. K-FAC

In some setups, Fisher might still require solid computational resources. Another popular factorization is the Kronecker-factored approximate curvature (K-FAC). It factorizes layer-wise Fisher as a Kronecker product of smaller matrices under the assumption of independence of layer-wise weights.

Given a layer with N hidden units, denote h_n as the n -th hidden vector and g_n as the log-likelihood gradient w.r.t. h_n . Then Fisher can be approximated as

$$F_n \approx \mathbb{E}(h_{n-1} h_{n-1}^T) \otimes \mathbb{E}(g_n g_n^T). \quad (7)$$

F. Embeddings Visualization

In this section, we visualize the embeddings for CIFAR-10, CIFAR-100, C100-5-SP and C100-5-NSP in Figure 1.

G. Additional Calibration Details

Here, we provide a precise comparison between LLLA variations and ICLA for calibration in Table 3.

H. Prior Precision Values

We report the obtained prior precision λ values in Section 4.2: 2.76 for CIFAR-10 and 3.06 for CIFAR-100.

I. Prior Precision Impact on OOD Detection

We demonstrate the relation between the value of prior precision λ and OOD detection AUROC in Table 2. As can be seen, prior precision values affect the OOD detection performance. It makes prior precision optimization sensible in our algorithm.

Table 2. OOD detection AUROC depending on prior precision value λ for CIFAR-10 dataset.

λ	Near OOD	Far OOD
1	90.31	92.20
3	89.89	91.76
5	89.06	91.08
7	88.84	90.86

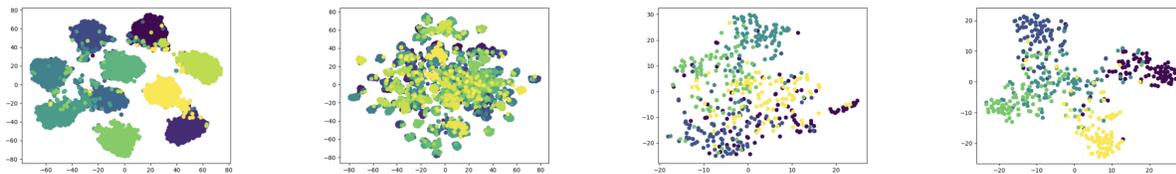


Figure 1. Visualizations of feature embeddings. (First) CIFAR-10. (Second) CIFAR-100. (Third) C100-5-NSP. (Fourth) C100-5-SP. CIFAR-10 and C100-5-SP present more separability, as the clusters of classes overlap less. ICLA performs better on more separable cases, showing the connection between curvature and data separability. See Section 5.3 for details.

Table 3. Precise calibration results between LLLAs and ICLA.

Dataset	Metric	LLLA (GGN)	LLLA (EF)	LLLA (K-FAC)	ICLA
CIFAR-10	ECE	1.23 ± 0.21	10.80 ± 7.10	2.75 ± 0.29	15.15 ± 4.53
	NLL	1.53 ± 0.01	1.62 ± 0.05	1.51 ± 0.03	1.59 ± 0.04
	Brier	0.07 ± 0.01	0.09 ± 0.01	0.08 ± 0.01	0.10 ± 0.01
CIFAR-100	ECE	19.11 ± 1.69	61.84 ± 4.08	6.47 ± 0.86	65.11 ± 0.54
	NLL	4.08 ± 0.01	4.47 ± 0.03	3.89 ± 0.02	4.47 ± 0.03
	Brier	0.37 ± 0.01	0.75 ± 0.05	0.32 ± 0.01	0.80 ± 0.01
ImageNet-200	ECE	3.54 ± 0.02	4.78 ± 0.24	1.77 ± 0.18	4.77 ± 0.26
	NLL	3.53 ± 0.01	3.91 ± 0.01	3.11 ± 0.01	3.94 ± 0.01
	Brier	0.21 ± 0.01	0.22 ± 0.01	0.20 ± 0.01	0.22 ± 0.01