Multi-Loss Sub-Ensembles for Accurate Classification with Uncertainty Estimation

Achrack Omer
Ouriel Barzilay
Kellerman Raizy

NEURAL INFORMATION PROCESSING SYSTEMS

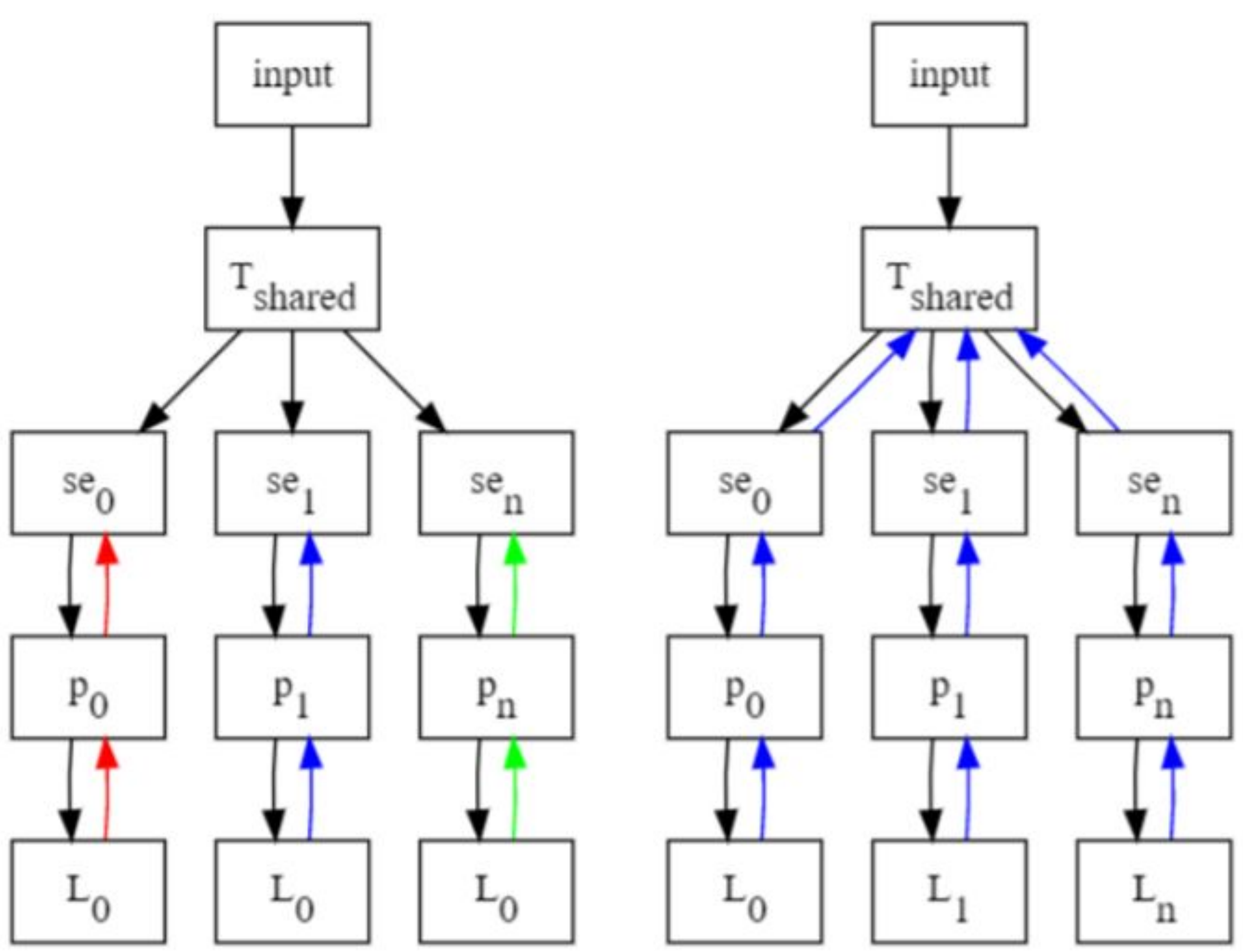
Abstract

We propose an efficient method for uncertainty estimation in DNNs achieving high accuracy. We simulate the notion of multi-task learning on single-task problems by producing parallel predictions from similar models differing by their loss. This multi-loss approach allows one-phase training for single-task learning with uncertainty estimation. We keep our inference time relatively low by leveraging the advantage proposed by the Deep Sub-Ensembles method. The novelty of this work resides in the proposed accurate variational inference with a simple and convenient training procedure, while remaining competitive in terms of computational time and with a good amount of breathing room in between each.

- Attractive in runtime
- Simple training procedure
- Improved accuracy
- Competitive uncertainty measures

Deep sub ensembles vs. our approach

Here we demonstrate the major difference between us and Deep-Sub-Ensembles(DSE) method. In our method the training is done in an end-to-end manner using different loss functions. In DSE method training is done in phases using same loss function.



Implementation details

Table with all losses per architecture as well as split point choice. we added also other training details like number of epochs.

Table 2: Implementation details

Arch/Data-set	Method	Epochs	Loss	Split
VGG16/SVHN	MC-dropout	300	[36]	-
	DE	300+30	[36]	-
	DSE	300+80x4	[36]	Blocks 3-5
	SWA	300+30	[36]	-
	Ours	300	[36], [42], [47],MSE	Blocks 3-5
MobileNet-v2/CIFAR10	MC-dropout	300	[36]	-
	DE	300+30	[36]	-
	DSE	300+80x4	[36]	Last 2 blocks
	SWA	300+30	[36]	-
	Ours	300	[36], [42], MAE,MSE	Last 2 blocks
ResNet50/CIFAR100	MC-dropout	300	softmax	-
	DE	300+30	softmax	-
	DSE	300+80x3	softmax	Last block
	SWA	300+30	softmax	-
	Ours	300	softmax,[42],MSE	Last block
Xception/ImageNet	MC-dropout	150	[36]	-
	DE	150+30	[36]	-
	DSE	150+50x2	[36]	From 7th middle-flow block onward
	SWA	150+50	[36]	-
	Ours	150	[36], [42]	From 7th middle-flow block onward

Results

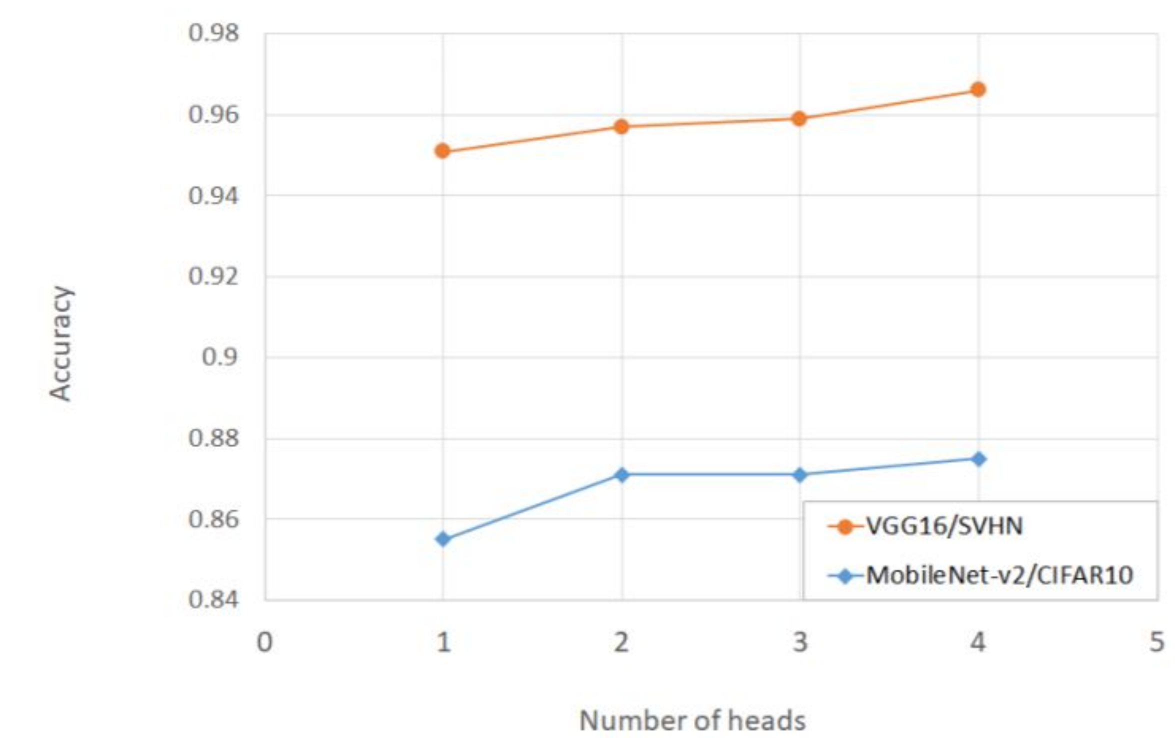
Here we present our main results. For each architecture/Data-set pair we measure accuracy as well as Brier Score and expected calibration error (ECE) for all methods we compare with as well as our method.

Table 1: Basic accuracy and calibration comparison

Arch/Data-set	Method	Accuracy	ECE	Brier Score
VGG16/SVHN	MC-dropout	0.965	0.05	0.005
	DE	0.963	0.037	0.006
	DSE	0.962	0.652	0.051
	SWA	0.962	0.033	0.006
	Ours	0.966	0.262	0.013
MobileNet-v2/CIFAR10	MC-dropout	0.823	0.055	0.025
	DE	0.873	0.029	0.019
	DSE	0.861	0.034	0.02
	SWA	0.871	0.048	0.022
	Ours	0.875	0.1	0.023
ResNet50/CIFAR100	MC-dropout	0.669	0.029	0.022
	DE	0.709	0.074	0.019
	DSE	0.609	0.101	0.028
	SWA	0.713	0.021	0.019
	Ours	0.679	0.04	0.022
Xception/ImageNet	MC-dropout	0.581	0.011	0.023
	DE	0.801	0.028	0.022
	DSE	0.802	0.030	0.042
	SWA	0.816	0.027	0.022
	Ours	0.832	0.029	0.030

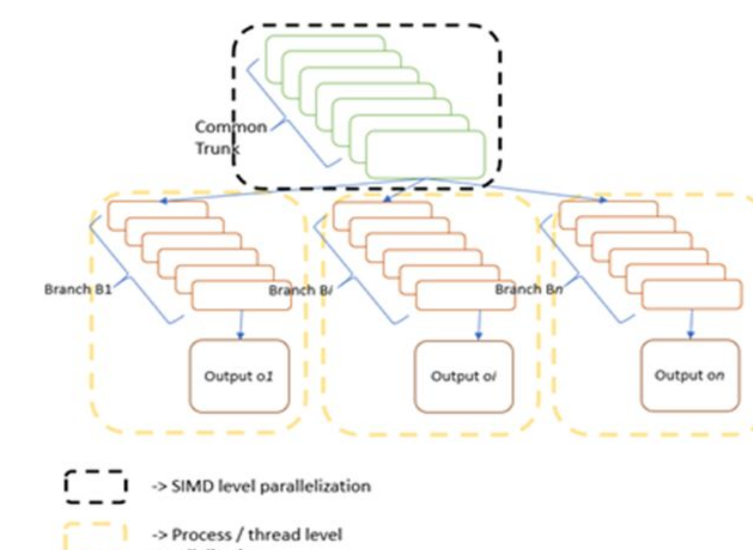
Accuracy vs. number of heads

Here we demonstrate an important behavior of our work. as number of heads increase, the accuracy should be increase also.



Engineering point of view

From engineering point of view, using an appropriate parallelization mechanism one can extract the feature from the shared trunk and in parallel execute each head. This might lead to another runtime reduction.



References

Matias Valdenegro-Toro. "Deep Sub-Ensembles for Fast Uncertainty Estimation in Image Classification". In: arXiv preprint arXiv:1910.08168 (2019)

Marge Simpson (2010). "Blue hair looks nice.". In: Nature communications 1, p. 622.

Murat Sensoy, Lance Kaplan, and Melih Kandemir. "Evidential deep learning to quantify classification uncertainty". In: Advances in Neural Information Processing Systems. 2018, pp. 3179–3189.

Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell. "Simple and scalable predictive uncertainty estimation using deep ensembles". In: Advances in neural information processing systems. 2017, pp. 6402–6413.