

Supplementary Material

Impatient DNNs – Deep Neural Networks with Dynamic Time Budgets

Manuel Amthor
manuel.amthor@uni-jena.de

Erik Rodner
erik.rodner@uni-jena.de

Joachim Denzler
joachim.denzler@uni-jena.de

Computer Vision Group
 Friedrich Schiller University Jena
 Germany
www.inf-cv.uni-jena.de

Abstract

The following documents contain additional results related to our submission.

1 Individual performance of EP layers

In the following section, we analyze the accuracies of each early prediction layer individually for different architectures and datasets. As can be seen in Table 1, the specified budget distribution has a significant impact on the final model performance. Especially in early as well as late layers the influence of extreme loss weights becomes most evident. As for example of an Impatient AlexNet trained on MIT-67 (third row), an increasing polynomial weighting scheme **POLY** yields an accuracy of 2.83% for the very first predictor EP1, where in contrast an inverted weighting **IPOLY** provides a substantial improvement to 29.85%. On the other hand, the inverted polynomial weighting scheme **IPOLY** shows an accuracy of the last layer EP6 of only 28.81% whereas its counterpart with increasing polynomial weighting provides an accuracy of 57.46%. Without any prior knowledge of the available time budget, the equally distributed weighting scheme offers a decent trade-off.

Impatient AlexNet trained on 15-Scenes								
EP	EQ	LIN	POLY	ILIN	IPOLY	NORM	t_B	t_A
EP 1	71.32	65.42	15.81	73.23	73.50	63.91	0.329	0.325
EP 2	82.78	82.31	72.12	83.58	82.94	83.24	0.441	0.454
EP 3	84.92	85.66	84.42	84.92	84.92	85.22	0.566	0.664
EP 4	86.23	87.23	87.03	86.76	84.89	86.49	0.732	0.772
EP 5	87.33	87.80	87.53	87.26	84.22	87.26	0.867	0.993
EP 6	88.04	88.37	87.80	87.87	67.67	87.70	1.077	1.233
Impatient VGG19 trained on 15-Scenes								
EP	EQ	LIN	POLY	ILIN	IPOLY	NORM	t_B	t_A
EP 1	61.57	53.50	10.75	65.59	69.94	52.52	0.357	0.350
EP 2	79.93	77.52	60.80	81.17	81.77	79.63	0.531	0.552
EP 3	87.57	88.17	86.59	87.70	87.00	87.73	0.889	0.936
EP 4	90.01	90.05	90.18	90.41	89.07	89.91	1.189	1.326
EP 5	92.49	92.19	91.42	92.46	89.54	92.39	1.710	1.815
EP 6	92.12	92.46	91.45	92.39	74.63	91.95	1.903	2.034
Impatient AlexNet trained on MIT-67								
EP	EQ	LIN	POLY	ILIN	IPOLY	NORM	t_B	t_A
EP 1	28.58	24.03	2.83	29.03	29.85	23.06	0.329	0.325
EP 2	41.64	41.34	31.49	41.57	42.01	41.94	0.441	0.454
EP 3	50.37	49.85	49.55	49.48	48.51	50.82	0.566	0.664
EP 4	52.91	53.58	53.58	53.13	50.60	53.96	0.732	0.772
EP 5	53.73	54.70	54.63	53.58	47.01	53.58	0.867	0.993
EP 6	57.46	58.51	57.46	55.00	28.81	54.92	1.077	1.233
Impatient VGG19 trained on MIT-67								
EP	EQ	LIN	POLY	ILIN	IPOLY	NORM	t_B	t_A
EP 1	25.74	21.41	2.68	26.34	27.76	21.56	0.357	0.350
EP 2	39.62	35.97	22.16	40.44	40.44	38.50	0.531	0.552
EP 3	52.91	52.16	49.70	52.61	51.34	52.83	0.889	0.936
EP 4	62.31	63.13	63.58	60.89	55.82	62.31	1.189	1.326
EP 5	69.40	70.52	71.79	68.20	58.20	68.88	1.710	1.815
EP 6	67.23	69.47	71.71	68.50	36.64	67.83	1.903	2.034

Table 1: Comparison of different weighting schemes for different time-budget distributions. Accuracy on the testset is shown for all available early prediction layers (EP) with corresponding runtimes for both scenarios (a priori given budget t_B and interruptable prediction t_A).