Tampereen yliopisto
Tampere University

AARHUS
UNIVERSITY

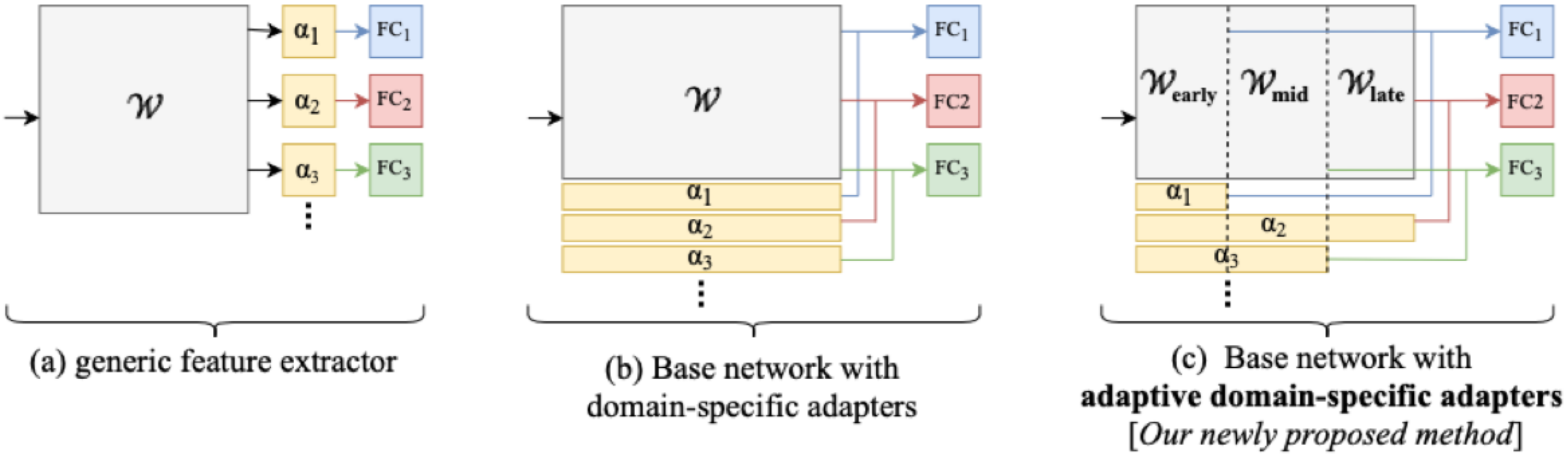
ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

Not all domains are equally complex: Adaptive Multi-Domain Learning

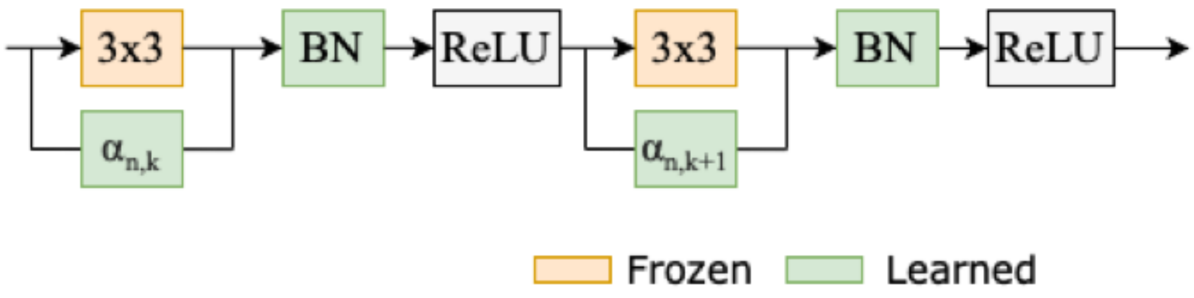
Ali Senhaji*, Jenni Raitoharju†, Moncef Gabbouj* and Alexandros Iosifidis‡

*Tampere University, Finland | †Finnish Environment Institute, Finland | ‡Aarhus University, Denmark

Multi-domain Learning Paradigms

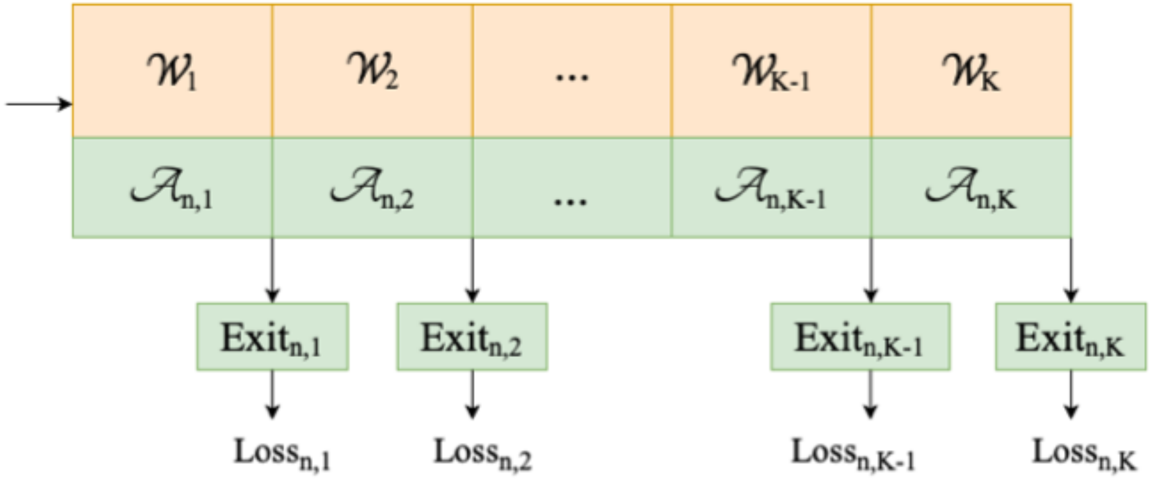


Adaptive Multi-domain Learning [Method]



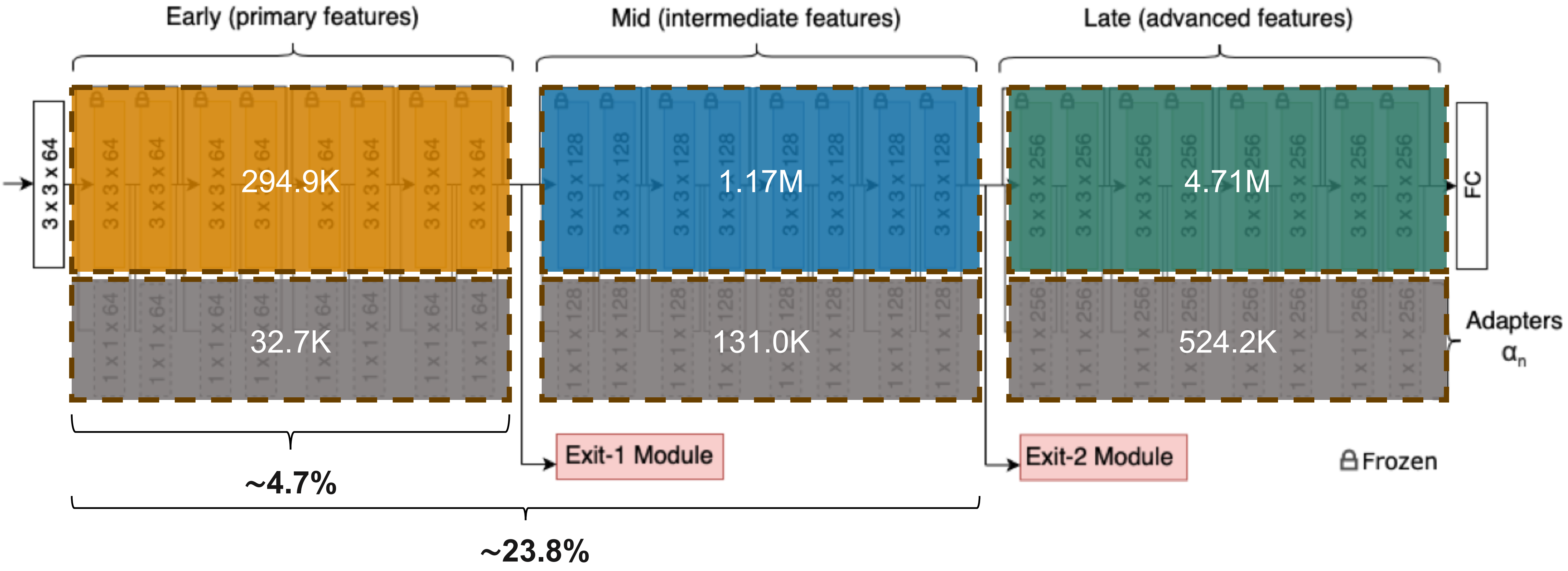
The parallel configuration of domain specific adapters within two consecutive convolution layers.

S. Rebuffi, H. Bilen, and A. Vedaldi. Learning multiple visual domains with residual adapters. In Proc. NIPS, 2017.

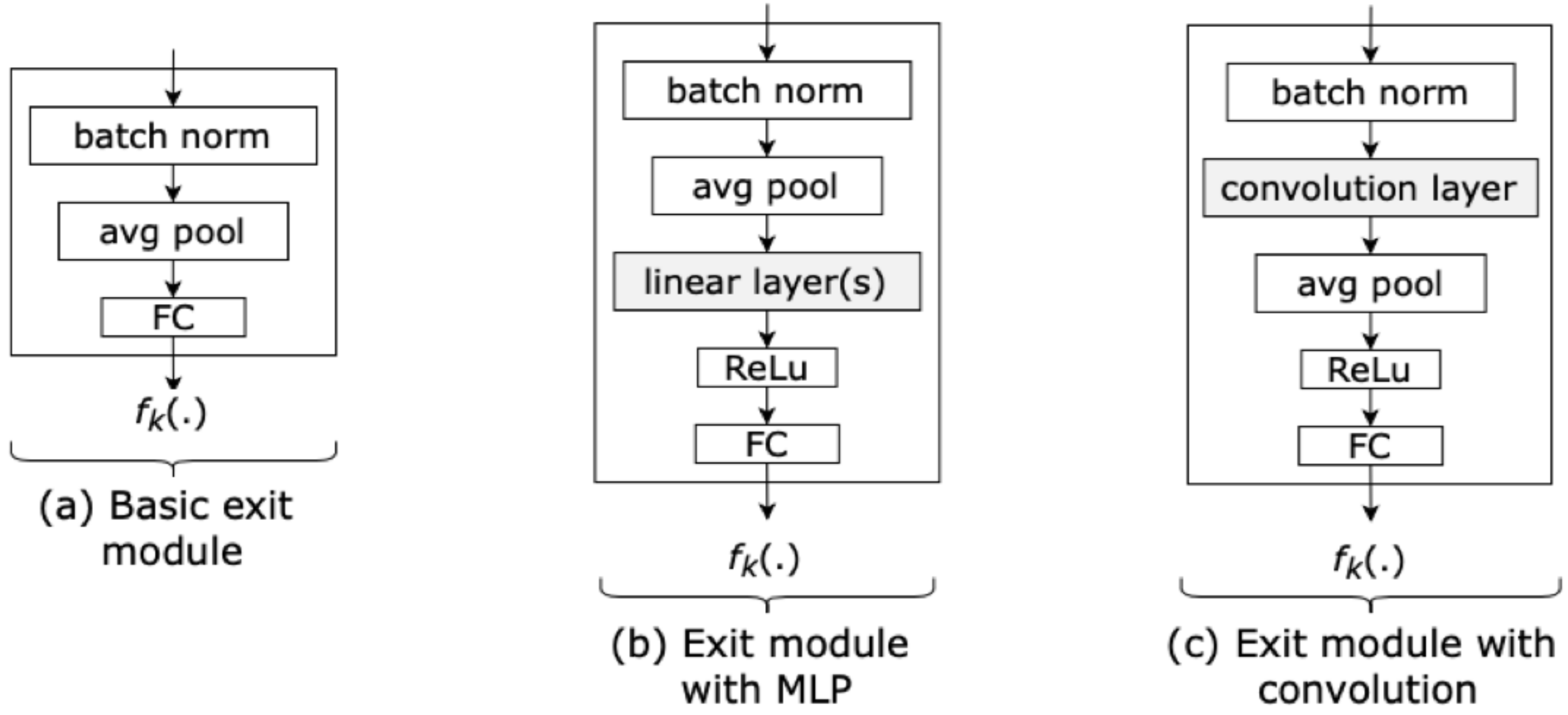


The partition of the base model into K non-overlapping blocks of layers adapted with A_n. Exit modules are placed after every portion of the base model.

Adaptive Multi-domain Learning [Method]



Exit Modules [Method]



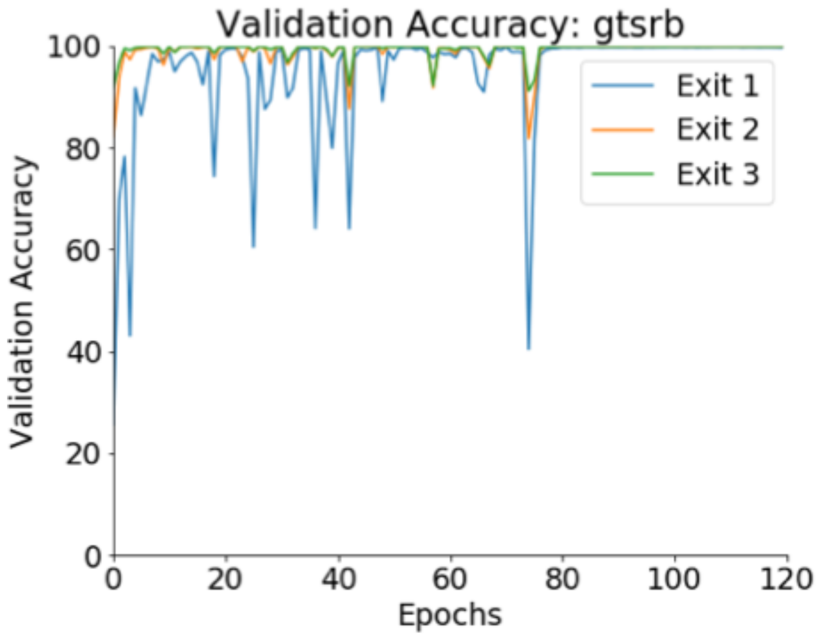
Datasets

Dataset	no. classes	training	validation	testing
Aircraft	100	3334	3333	3333
CIFAR-100	100	40000	10000	10000
Daimler Ped	2	23520	5880	19600
D. Textures	47	1880	1880	1880
GTSRB	43	31367	7842	12630
ImageNet	1000	1232167	49000	48238
Omniglot	1623	19476	6492	6492
SVHN	10	47217	26040	26032
UCF101 Dyn	101	7629	1908	3783
VGG-Flowers	102	1020	1020	6149

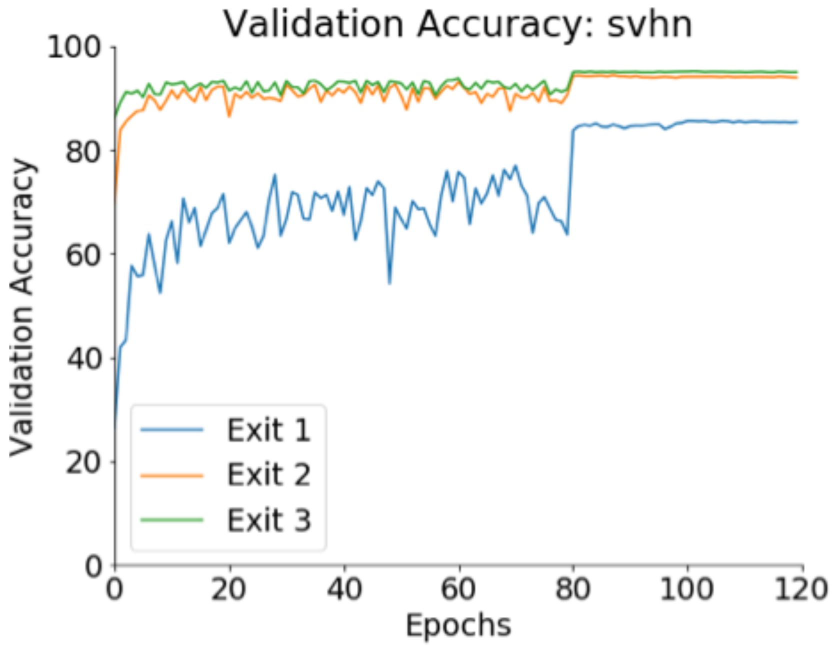
Experimental Results

Method	%P.	#P.	ImNet	Airc.	C100	DPed	DTD	GTSR	Flwr	OGlt	SVHN	UCF	mean
Finetune [8]	10x	61.94M	59.87	60.34	82.12	92.82	55.53	97.53	81.41	87.69	96.55	51.20	76.51
A. Series [9]	2x	12.38M	59.67	56.68	81.20	93.88	50.85	97.05	66.24	89.62	96.13	47.45	73.88
A. Paral [8]	2x	12.38M	60.32	50.29	81.01	90.57	51.65	99.02	70.24	87.42	95.84	48.01	73.44
No-adp E-1	0.05x	296.6K	-	10.12	25.97	66.58	24.74	41.18	33.03	41.34	27.39	14.99	31.71
No-adp E-2	0.24x	1.47M	-	7.30	20.08	64.64	17.88	29.55	21.18	49.12	25.05	11.48	27.37
No-adp E-3	1x	6.19M	60.32	4.33	9.25	63.52	11.71	19.02	10.61	21.67	20.84	6.96	22.83
Basic E-1	0.09x	591.5K	-	16.30	53.85	87.05	39.26	96.28	47.87	81.79	89.72	31.09	60.36
Basic E-2	0.47x	2.95M	-	42.55	72.20	89.86	48.20	97.88	56.31	**87.13**	**95.35**	43.30	70.31
Basic E-3	2x	12.38M	60.32	46.18	78.00	89.84	49.53	97.86	59.19	87.41	96.06	46.95	71.14
MLP128 E-1	0.09x	599.7K	-	24.07	53.65	**87.33**	43.73	**97.00**	54.18	80.95	**90.04**	30.19	**62.35**
MLP128 E-2	0.48x	2.95M	-	45.04	71.85	90.80	**49.15**	98.52	63.14	86.44	**95.24**	**43.12**	**71.48**
MLP128 E-3	2x	12.38M	60.32	49.36	78.55	90.93	49.90	98.86	68.06	88.32	96.21	**48.64**	72.92
MLP128-B E-1	0.09x	599.7K	-	23.23	56.90	**87.72**	39.85	95.77	46.32	80.71	90.16	33.15	61.54
MLP128-B E-2	0.48x	2.95M	-	43.63	71.88	89.74	47.56	96.87	58.42	85.70	95.02	43.33	70.24
MLP128-B E-3	2x	12.38M	60.32	**50.29**	**81.01**	90.57	51.65	99.02	**70.24**	87.42	95.84	48.01	73.44
MLP512 E-1	0.10x	624.3K	-	24.40	54.60	87.08	40.64	95.81	53.67	80.97	87.25	31.96	61.82
MLP512 E-2	0.48x	3.01M	-	45.37	71.67	90.78	47.40	97.78	61.87	86.46	95.30	44.20	71.21
MLP512 E-3	2x	12.38M	60.32	50.32	78.18	90.93	48.73	97.78	64.73	87.78	96.28	46.90	72.20
MLP-2L E-1	0.10x	616.1K	-	23.17	53.04	86.86	41.07	95.54	53.93	74.62	87.67	30.06	60.67
MLP-2L E-2	0.48x	2.98M	-	44.74	70.49	90.46	47.98	96.88	63.97	83.80	94.78	42.64	70.64
MLP-2L E-3	2x	12.38M	60.32	**50.62**	77.94	90.68	48.78	97.46	69.85	88.34	96.06	**49.04**	72.91
CL E-1	0.09x	595.6K	-	23.47	52.55	86.56	41.18	96.16	51.66	80.79	88.60	30.91	61.32
CL E-2	0.47x	2.96M	-	43.21	71.34	89.89	**49.53**	97.76	59.53	86.20	95.04	43.33	70.65
CL E-3	2x	12.38M	60.32	47.74	77.50	90.14	49.10	97.81	63.23	87.94	96.16	47.11	71.71
Best T=3.5%	1.53x	9.50M	60.32	50.62	81.01	87.72	49.53	97.00	70.24	87.13	95.35	49.04	72.79

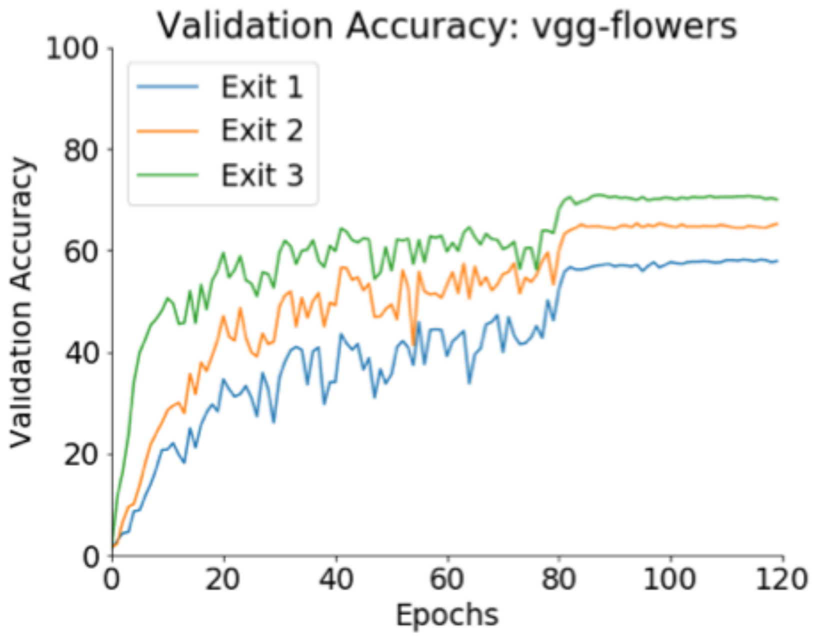
Experimental Results



Easy Domain



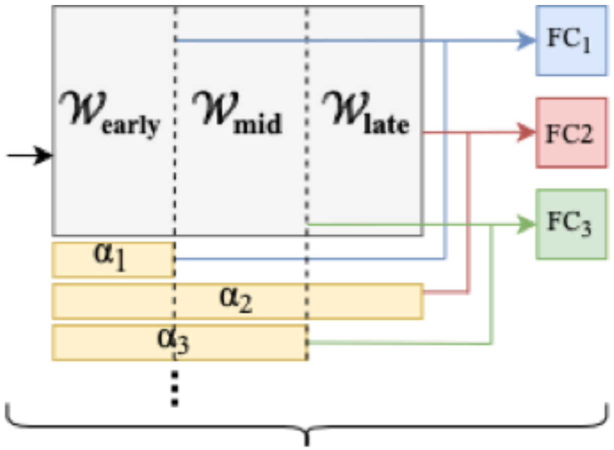
Intermediate Domain



Challenging Domain

Conclusions

- We proposed an adaptive method for incremental multi-domain learning for reducing the required base model parameters and the domains specific parameters added for domains with different levels of complexity.
- We showed through experimentation that the same levels of performance can be achieved with only a portion of the base network and its corresponding adapters for some less complex domains.
- Our approach encourages an efficient adaptation of multi-domain paradigms using domain agnostic parameters.



Base network with
adaptive domain-specific adapters
[*Our newly proposed method*]