

Challenge Overview



Task1 → Task2 → Task3 → Task4

Batch/Task	Illumination	Occlusion	Clutter
1	Strong		
2	Weak		
3		50%	
4			high

Training fast

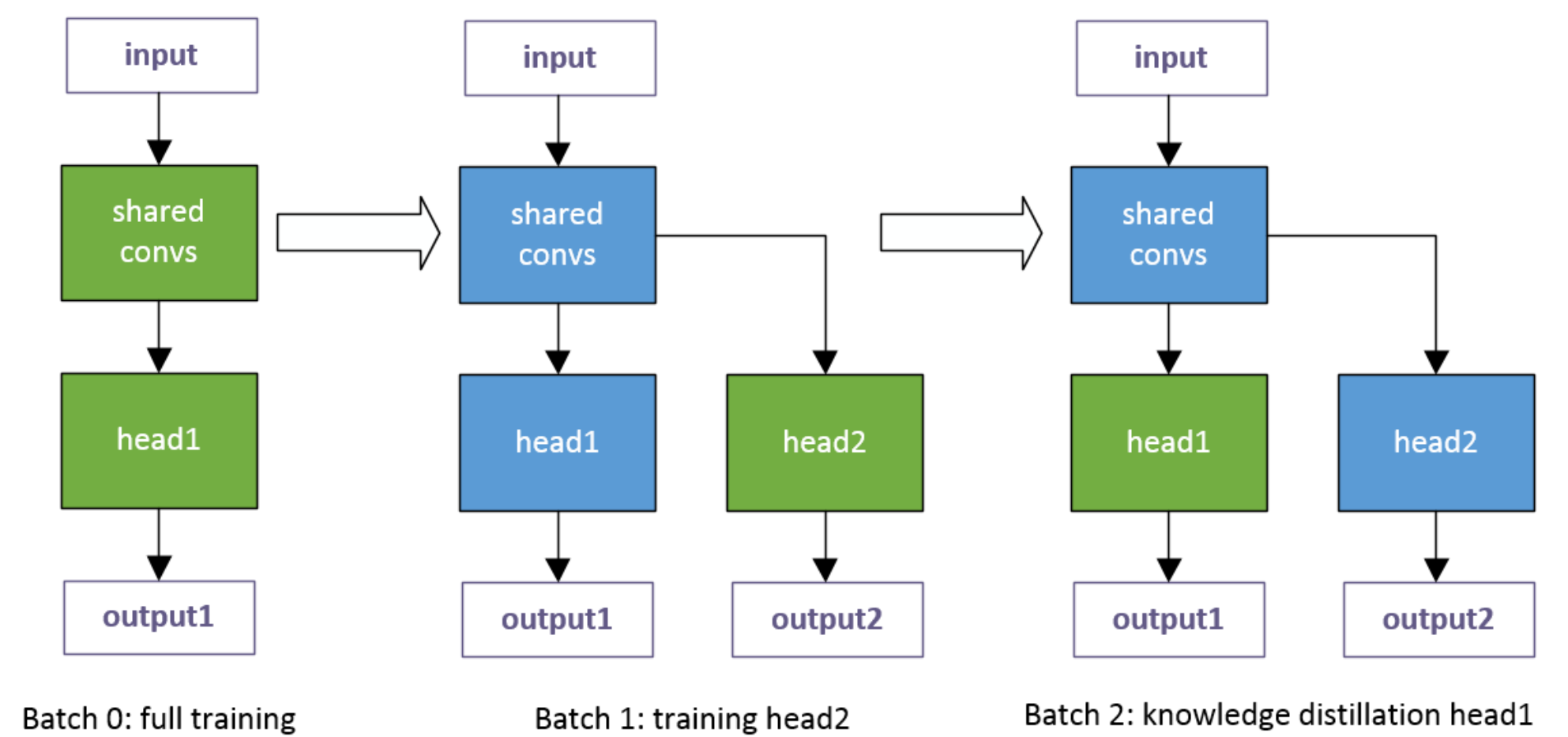
GOAL

Less storage and memory costs

Good performance without forgetting

Our Method

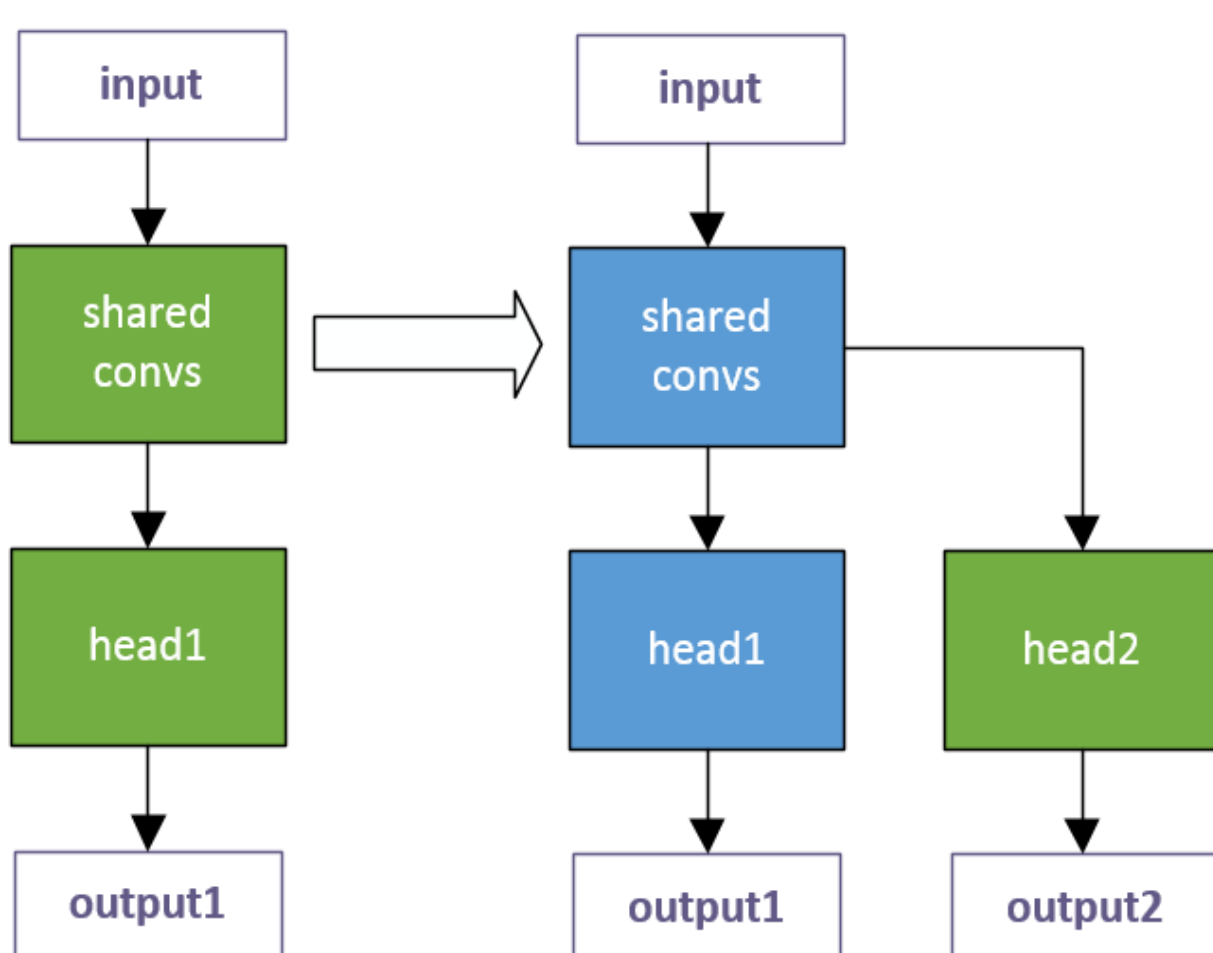
- Dynamic network expansion for data across dissimilar domain
- Knowledge distillation for data in similar domain



- Features**
- Combined method of network expansion and feature regularization
 - No need for previous data
 - Dynamic network expansion to alleviate domain gap

Dynamic network expansion

- Freeze shared conv layers
- Network expansion for severe domain gap (bad accuracy)



Batch 0: full training

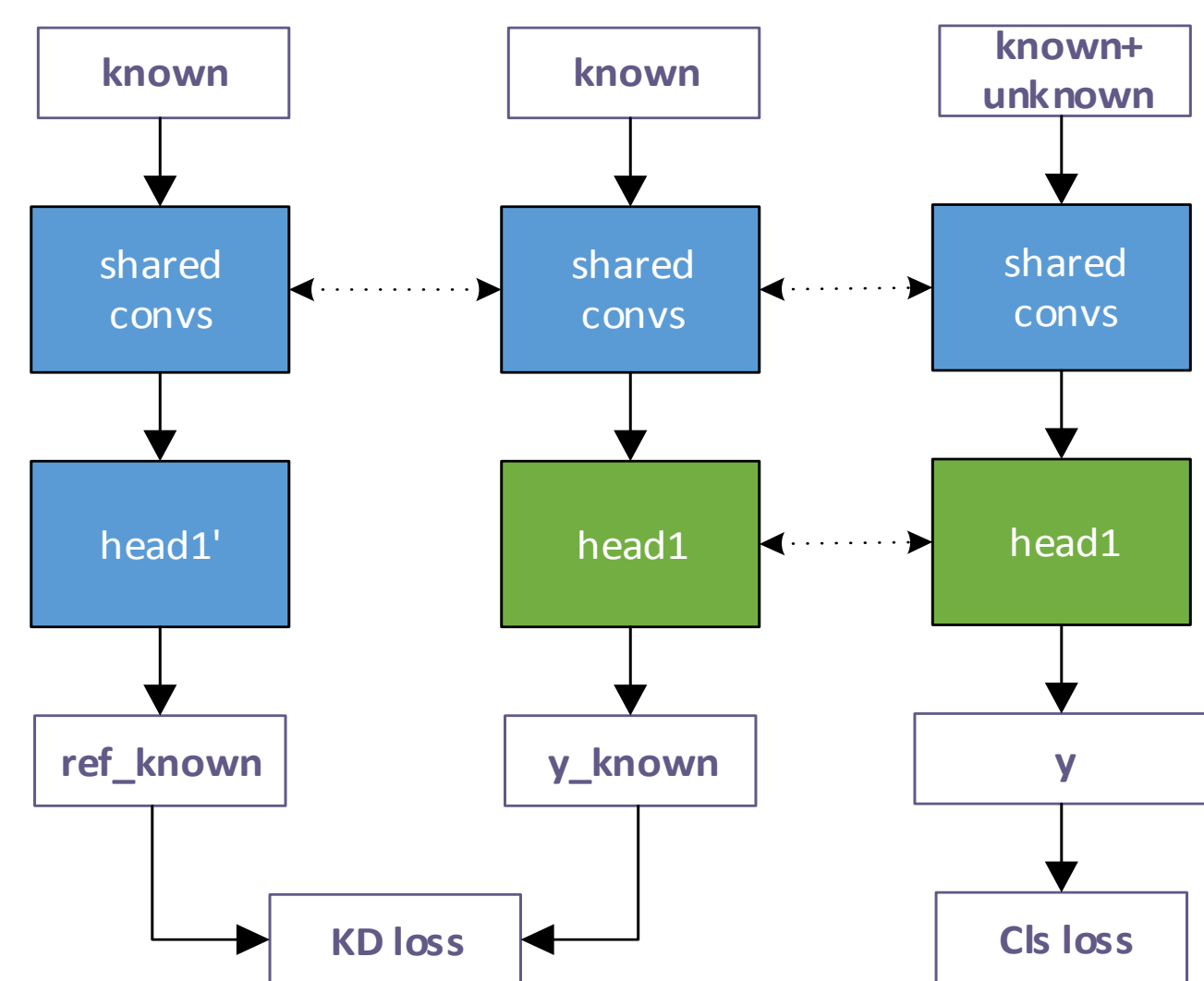
Batch 1: training head2

Tricks for generalization:

- For shared convs, imagenet pre-trained model
- For heads, more data augmentation and more batches to train head1

Knowledge distillation

- Replace BatchNorm with GroupNorm
- Mining known instances in new task and distill on best head



Experiment Results on 1st round

Finetune

DynamicNN

Square performance on all tasks

Adam, base lr=0.003 @[50, 90, 120], inc lr=0.003 @[50, 90]

	0	1	2	3	4	5	6	7	8	avg
Task 0	99.71 ± 0.06	/	/	/	/	/	/	/	/	99.69 ± 0.06
Task 1	85.61 ± 0.72	99.91 ± 0.10	/	/	/	/	/	/	/	92.70 ± 0.39
Task 2	97.93 ± 0.40	80.92 ± 2.61	99.88 ± 0.10	/	/	/	/	/	/	92.77 ± 0.91
Task 3	76.87 ± 1.19	68.28 ± 1.04	79.01 ± 1.89	99.99 ± 0.01	/	/	/	/	/	81.16 ± 1.10
Task 4	87.40 ± 0.76	91.02 ± 0.84	86.81 ± 1.08	96.93 ± 0.53	99.95 ± 0.03	/	/	/	/	92.54 ± 0.47
Task 5	88.29 ± 0.85	89.33 ± 0.99	84.16 ± 0.54	81.67 ± 1.70	82.71 ± 2.02	99.90 ± 0.02	/	/	/	87.40 ± 0.94
Task 6	99.00 ± 0.23	95.62 ± 1.04	99.07 ± 0.25	98.75 ± 0.35	94.97 ± 0.86	95.81 ± 0.59	99.72 ± 0.08	/	/	97.46 ± 0.41
Task 7	92.01 ± 0.74	97.01 ± 0.09	89.40 ± 0.95	87.62 ± 2.25	93.58 ± 1.40	99.30 ± 0.01	91.21 ± 1.06	99.91 ± 0.05	/	93.58 ± 0.73
Task 8	98.54 ± 0.28	96.09 ± 0.49	99.53 ± 0.01	92.72 ± 1.88	92.69 ± 1.22	93.54 ± 0.21	95.57 ± 0.89	91.31 ± 0.83	99.82 ± 0.02	95.68 ± 0.51

V2: use worst branch for KD loss, expand @ task [1,3]

	0	1	2	3	4	5	6	7	8	avg
Task 0	99.30 ± 0.18	/	/	/	/	/	/	/	/	99.36 ± 0.17
Task 1	97.13 ± 0.46	98.42 ± 0.39	/	/	/	/	/	/	/	97.79 ± 0.07
Task 2	99.25 ± 0.28	95.92 ± 0.33	97.88 ± 0.30	/	/	/	/	/	/	97.68 ± 0.05
Task 3	96.74 ± 0.57	95.74 ± 0.28	94.69 ± 0.24	99.21 ± 0.06	/	/	/	/	/	96.63 ± 0.12
Task 4	94.50 ± 0.35	97.08 ± 0.08	93.12 ± 0.60	99.31 ± 0.20	97.10 ± 0.51	/	/	/	/	96.30 ± 0.28
Task 5	93.63 ± 0.68	97.25 ± 0.14	92.08 ± 0.93	99.26 ± 0.24	96.82 ± 0.55	90.89 ± 0.42	/	/	/	95.09 ± 0.35
Task 6	96.77 ± 0.93	98.01 ± 0.69	96.15 ± 1.38	99.35 ± 0.20	97.58 ± 0.53	93.41 ± 1.21	98.26 ± 0.49	/	/	96.93 ± 0.76
Task 7	96.82 ± 0.39	98.96 ± 0.24	96.08 ± 0.57	97.83 ± 0.53	97.84 ± 0.43	96.43 ± 0.84	97.85 ± 0.18	96.09 ± 0.75	/	97.29 ± 0.20
Task 8	97.98 ± 0.33	99.10 ± 0.33	98.00 ± 0.35	98.17 ± 0.45	97.94 ± 0.38	95.33 ± 1.06	98.38 ± 0.19	94.97 ± 0.85	97.97 ± 0.26	97.60 ± 0.23

finetune	93.84
DynamicNN(No expand)	94.50
DynamicNN(Expand@1)	95.75
DynamicNN(No expand) + LR trick	96.01