MTL-NAS: Task-Agnostic Neural Architecture Search towards General-Purpose Multi-Task Learning

Yuan Gao[1*], Haoping Bai[2*], Zequn Jie[1], Jiayi Ma[3], Kui Jia[4], Wei Liu[1]

1 Tencent AI Lab

2 Carnegie Mellon University

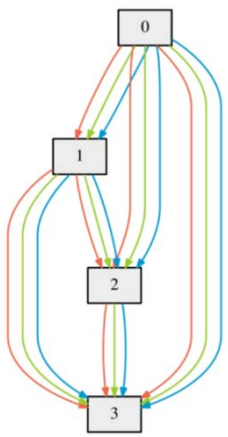
3 WUHAN UNIVERSITY

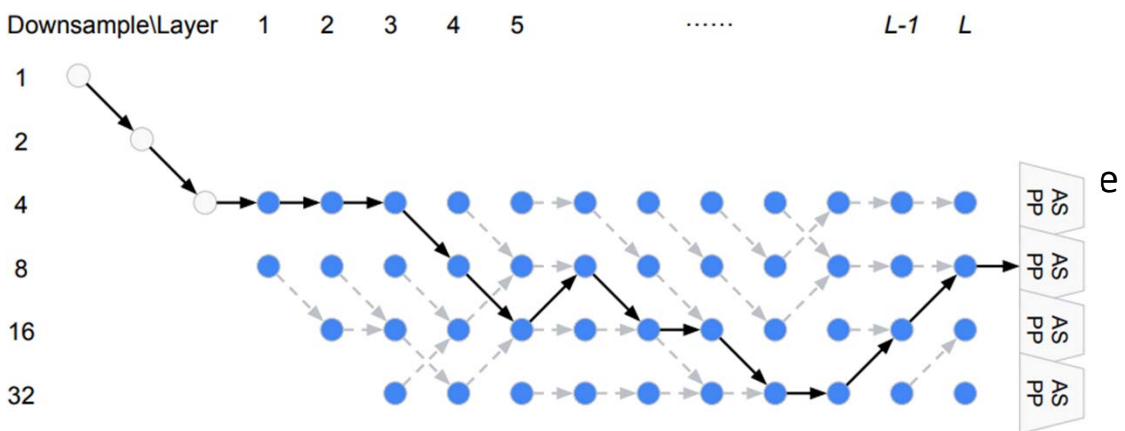
4 South China University of Technology

Motivation

- Integrating Multi-task Learning (MTL) and Neural Architecture Search (NAS) for further improved performance of CNNs.

- Difficulty:
  NAS is typically task specific, which encodes task prior into search space design [1, 2]:





e

Therefore, it is difficult to design a search space that can adapts to any tasks for general-purpose MTL (GP-MTL).

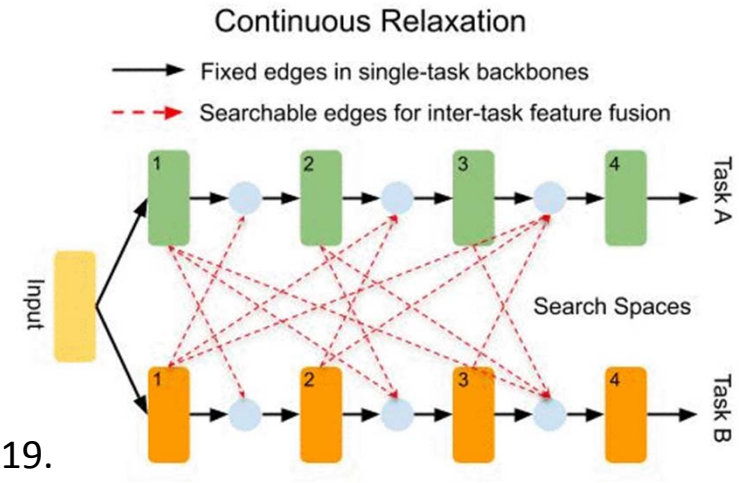
[1] Liu, et al., DARTS: Differentiable Neural Architecture Search. ICLR 2019

[2] Liu, et al., Auto-Deeplab. ECCV 2018.

Our Solution:

- For the *Search Space*, we formulates the GP-MTL problem as:
  1. Single-task branches for each task which encodes task priors.
  2. General feature fusing scheme across different branches [3].
- Those enables us to design a general search space for any task combinations.

- We use the single-shot gradient-based methods (e.g., DARTS) as the *Search Algorithm.*
- We imposes MinEntropy regularization on the architecture weights, which facilitates the algorithm converges to a single model that can be directly used for evaluation.



This is a gif image, please also check our teaser image.

[3] Gao, et al., NDDR-CNN. CVPR 2019.

Results:

- Our method is evaluated on various:
  1. Task Combinations, i.e.,
     Pixel Labeling Tasks: Semantic Segmentation + Surface Normal Prediction.
     Image Level Tasks: Object Classification + Scene Classification.
  2. Network Backbones, i.e., VGG-16 and ResNet-50.
- All the experiments demonstrate significant improvements w.r.t. SOTA GP-MTL.

VGG-16, Seg. + Normal Tasks

	Surface Normal Prediction					Semantic Seg.	
	Err (↓)		Within $t°$ (%) (↑)			(%) (↑)	
	Mean	Med.	11.25	22.5	30	mIoU	PAcc
Single	15.6	12.3	46.4	75.5	86.5	33.5	64.1
Multiple	15.2	11.7	48.4	76.2	87.0	33.4	64.2
C.-S.	15.2	11.7	48.6	76.0	86.5	34.8	65.0
NDDR	13.9	10.2	53.5	79.5	88.8	36.2	66.4
MTL-NAS	**12.6**	**8.9**	**59.1**	**83.3**	**91.2**	**37.6**	**67.9**

ResNet-50, Seg. + Normal Tasks

	Surface Normal Prediction					Semantic Seg.	
	Err (↓)		Within $t°$ (%) (↑)			(%) (↑)	
	Mean	Med.	11.25	22.5	30	mIoU	PAcc
Single	16.2	13.6	41.6	74.1	86.5	34.5	65.5
Multiple	16.6	14.2	39.2	73.8	86.5	34.8	65.1
C.-S.	16.6	14.3	39.1	73.7	86.5	34.8	65.7
NDDR	16.4	**12.8**	42.6	73.3	**86.6**	36.7	66.7
MTL-NAS	**16.2**	**12.8**	**44.8**	**73.9**	85.7	**38.6**	**68.6**

VGG-16, Obj. + Scene Cls. Tasks

	Object		Scene	
	RecRate (%) (↑)		RecRate (%) (↑)	
	Top 1	Top 5	Top 1	Top 5
Single	33.8	63.0	37.8	70.5
Multiple	34.1	66.1	37.8	71.2
Cross-Stitch	33.2	65.2	34.0	70.3
NDDR	32.1	57.7	37.9	71.8
MTL-NAS	**34.8**	**67.0**	**38.2**	**72.5**

Thanks