Bi-PointFlowNet

1차 보고

김민제

Bi-PointFlowNet 경량화 기법

- Knowledge Distillation 기법을 이용하여 경량화하고자 함.

- **Knowledge Distillation?**

Teacher Model의 지식을 Student Model이 학습하는 경량화 방법

Loss Function에

Loss = alpha*(Student Prediction – Tch answer) + (1-alpha)*(Student Prediction – Ground Truth)

Bi-PointFlowNet 경량화

- Student Model

  전체적인 channel depth를 절반으로 감소

  - 3->64->128->256->512-> Up Layer…
  - 3->32->64 ->128->256 -> Up Layer…

+ feature extracting layer의 경우 BottleNeck Layer를 이용해서 추가적으로 사이즈 축소

기존 모델 대비 **약 53%의 사이즈 감소**

***.pth [31.2MB -> 14.7MB]**

Number of Param [7.97M -> 3.72M]

Bi-PointFlowNet 경량화

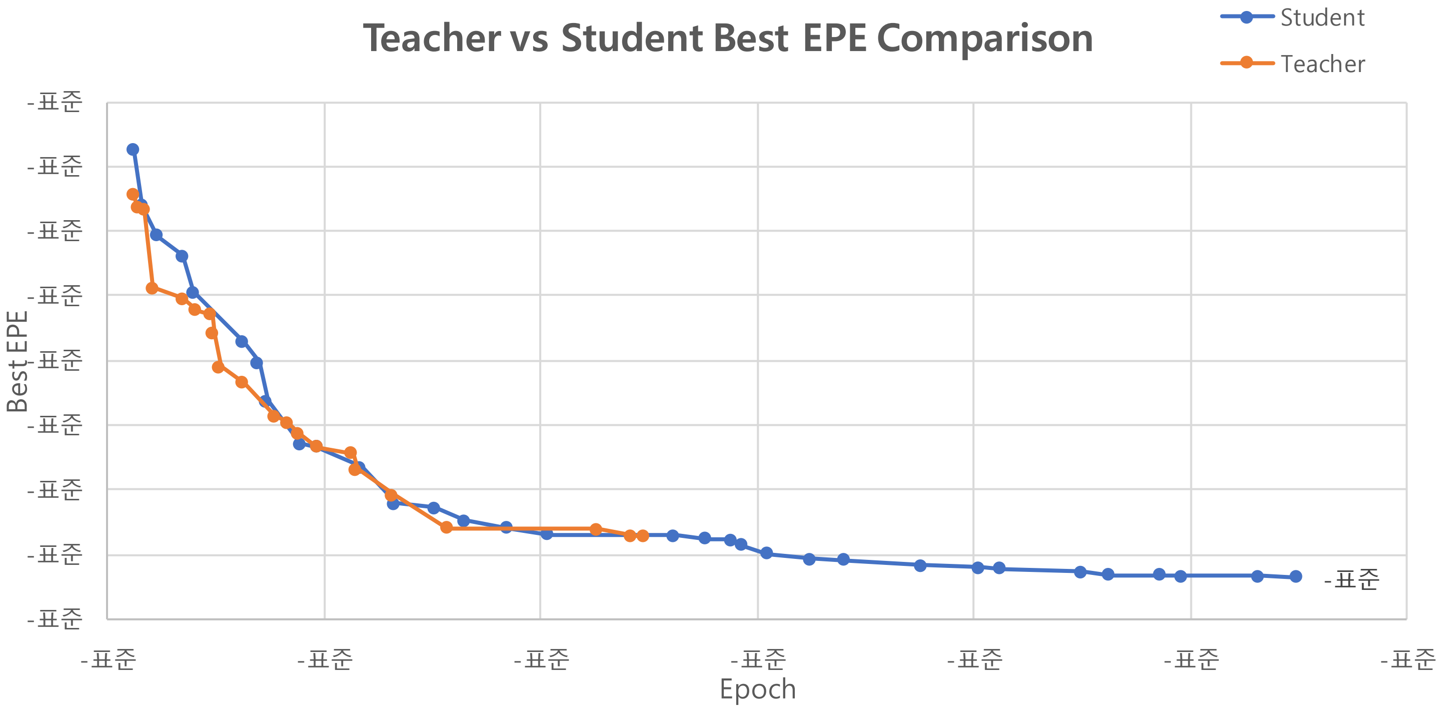
- Teacher Training Dataset을 이용해서 학습 진행 시 **(KITTI)**

Student Model : Epoch=560, LR=Scheduler(0.001, step_size=150, gamma=0.5)

(Teacher Model: Epoch=560, LR=Scheduler(0.001, step_size=80, gamma=0.5))

Pretrained Teacher Model(Epoch:536): 0.0359
Student Model (Epoch:550): 0.0323

더 경량화 가능성



Bi-PointFlowNet 경량화

Model	EPE3D(↓)	ACC3DS(↑)	ACC3DR(↑)	Outliers3D(↓)	EPE2D(↓)	ACC2D(↑)
Teacher	0.050	0.869	0.940	0.177	1.771	0.896
Student	0.033	0.910	0.987	0.165	1.574	0.924
Learned from Teacher	0.102	0.392	0.798	0.420	4.23	0.535
Learned from Student	0.067	0.589	0.919	0.284	3.116	0.680

Bi-PointFlowNet 경량화

- Problem

기존에 구현한 Teacher Knowledge를 7:3의 비율로 포함 시 Best EPE 0.167의 성능이 나옴

Loss Function을 재정의할 필요가 있음.

추후 연구

- Relational Knowledge Distillation

  [Wonpyo Park, Dongju Kim, Yan Lu, Minsu Cho, Relational Knowledge Distillation, CVPR, 2019]

- Self Knowledge Distillation

  [Linfeng Zhang, Jiebo Song, Anni Gao, Be your Own Teacher: Improve the Performance of Convolutional Neural Networks via Self Distillation, ICCV, 2019]

- Online Knowledge Distillation

  [Qiushan Guo1, Xinjiang Wang2, Yichao Wu2, Zhipeng Yu2, Ding Liang2, Xiaolin Hu3, Ping Luo4, Online Knowledge Distillaiton via Collaborative Learning, CVPR, 2020]