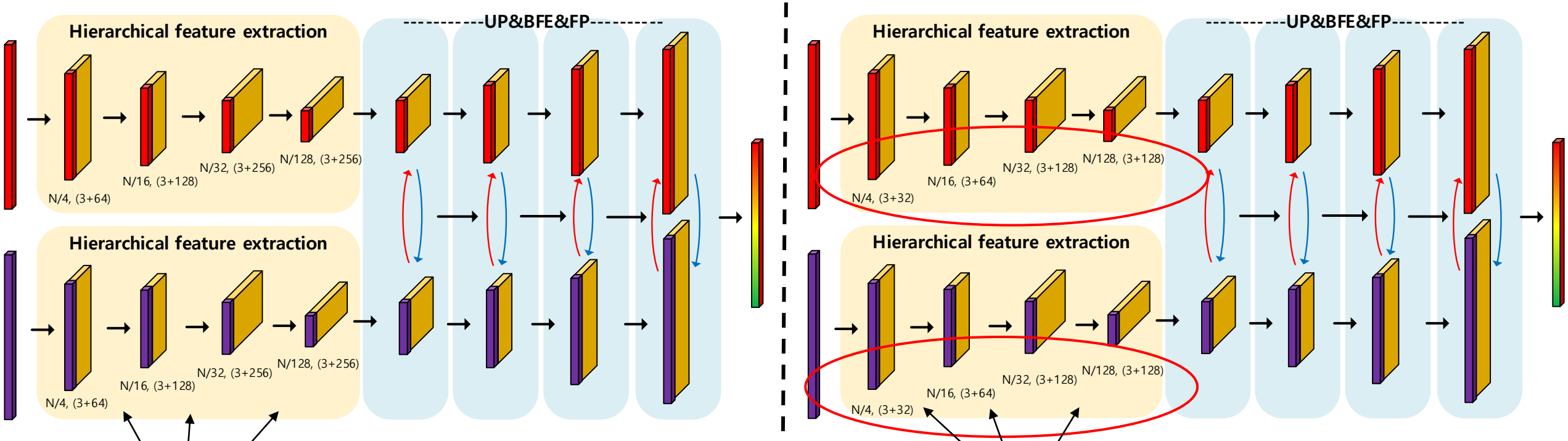
Bi-PointFlowNet

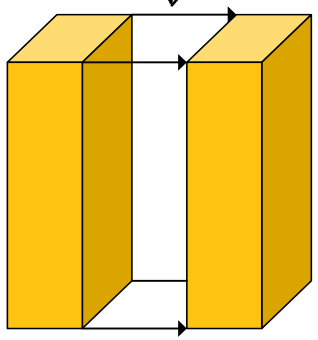
3차 보고

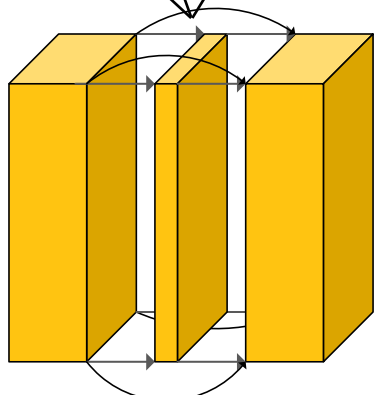
Student Architecture



Teacher Model Architecture(Baseline)

Student Model Architecture





	Teacher		Student
FLOPs	13.3G	41.4% ↓ →	7.8G
Param	7.98M	53.5% ↓ →	3.71M

Experiment

FlyingThings-Teacher: FlyingThings 데이터셋으로 학습한 선생 모델(논문 모델)

EPE3D(↓)	ACC3DS(↑)	ACC3DR(↑)	Outliers3D(↓)	EPE2D(↓)	ACC2D(↑)
0.030	0.920	0.960	0.141	1.056	0.949

KITTI-Teacher: KITTI 데이터셋으로 학습한 선생 모델

EPE3D(↓)	ACC3DS(↑)	ACC3DR(↑)	Outliers3D(↓)	EPE2D(↓)	ACC2D(↑)
0.075	0.617	0.873	0.323	3.648	0.686

KITTI-Scratch-Student: KITTI 데이터셋으로 학습한 학생 모델

EPE3D(↓)	ACC3DS(↑)	ACC3DR(↑)	Outliers3D(↓)	EPE2D(↓)	ACC2D(↑)
0.104	0.417	0.741	0.439	4.558	0.552

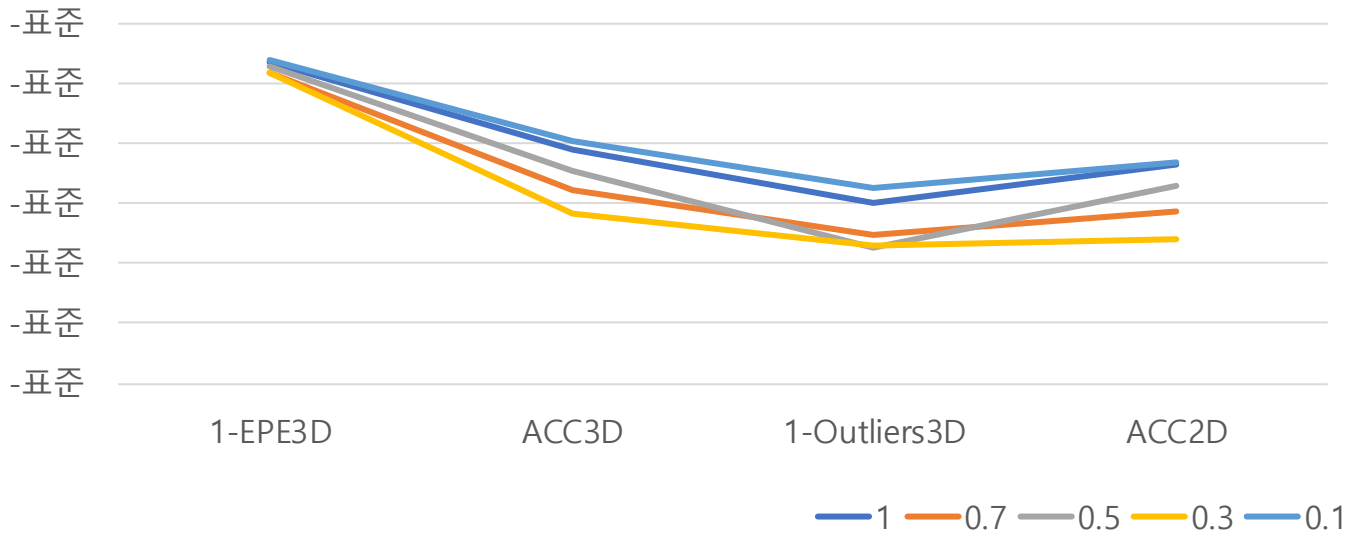
Experiment

EPE3D(↓)	ACC3DS(↑)	ACC3DR(↑)	Outliers3D(↓)	EPE2D(↓)	ACC2D(↑)
0.030	0.920	0.960	0.141	1.056	0.949

1. KITTI-Teacher + KITTI Dataset → **Student**

$$Loss = \gamma L_{KD} + (1-\gamma) L_{GT}$$

Gamma	EPE3D(↓)	ACC3DS(↑)	ACC3DR(↑)	Outliers3D(↓)	EPE2D(↓)	ACC2D(↑)
1	0.066	0.704	0.877	0.298	3.045	0.764
0.7	0.081	0.610	0.835	0.353	4.090	0.688
0.5	0.073	0.647	0.861	0.330	3.357	0.730
0.3	0.081	0.546	0.819	0.372	3.895	0.640
0.1	**0.060**	**0.710**	**0.900**	**0.275**	**2.887**	**0.769**
0	0.104	0.417	0.741	0.439	4.558	0.552



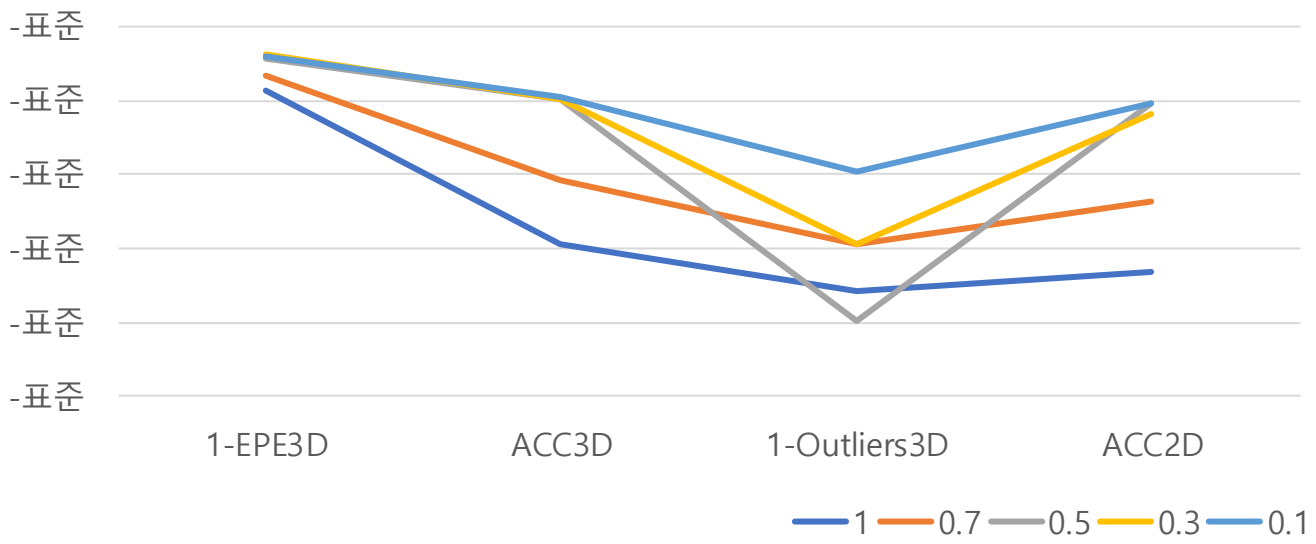
Experiment

EPE3D(↓)	ACC3DS(↑)	ACC3DR(↑)	Outliers3D(↓)	EPE2D(↓)	ACC2D(↑)
0.030	0.920	0.960	0.141	1.056	0.949

2. FlyingThings-Teacher + KITTI Dataset → **Student**

$$Loss = \gamma L_{KD} + (1-\gamma) L_{GT}$$

Gamma	EPE3D(↓)	ACC3DS(↑)	ACC3DR(↑)	Outliers3D(↓)	EPE2D(↓)	ACC2D(↑)
1	0.087	0.589	0.825	0.356	4.119	0.668
0.7	0.066	0.704	0.882	0.295	3.092	0.763
0.5	0.041	0.857	0.946	0.198	1.960	0.896
0.3	**0.039**	**0.842**	**0.961**	**0.195**	**1.937**	**0.881**
0.1	0.041	0.860	0.952	0.195	1.961	0.896
0	0.104	0.417	0.741	0.439	4.558	0.552



Experiment

3. **KITTI-Teacher** + **FlyingThings3D** → Student

4. **FlyingThings-Teacher** + **FlyingThings3D** → Student

5. **FlyingThings3D-Scratch**-Student

→ 연구실 컴퓨터를 이용해서 연구를 진행해야 할 것 같음

Experiment

정리

1. KITTI-Teacher + KITTI Dataset의 경우 선생 모델보다 좋은 성능을 보여줄때도 있었음

2. FlyingThings-Teacher + KITTI Dataset으로 Knowledge Distillation 했을 경우 가장 좋은 성능을 보여줌

Experiment

2. FlyingThings-Teacher + KITTI Dataset -> Student

BIG-Dataset-Teacher + **Small Dataset** -> Student

모델 경량화 + 학습 경량화

기존 Training Time (7일 + a) -> Training Time (12시간)

560 Epoch 1200Epoch

현재 가장 일반적인 KD방법을 사용
성능을 더 높일 수 있는 방법들을 찾아보고 연구