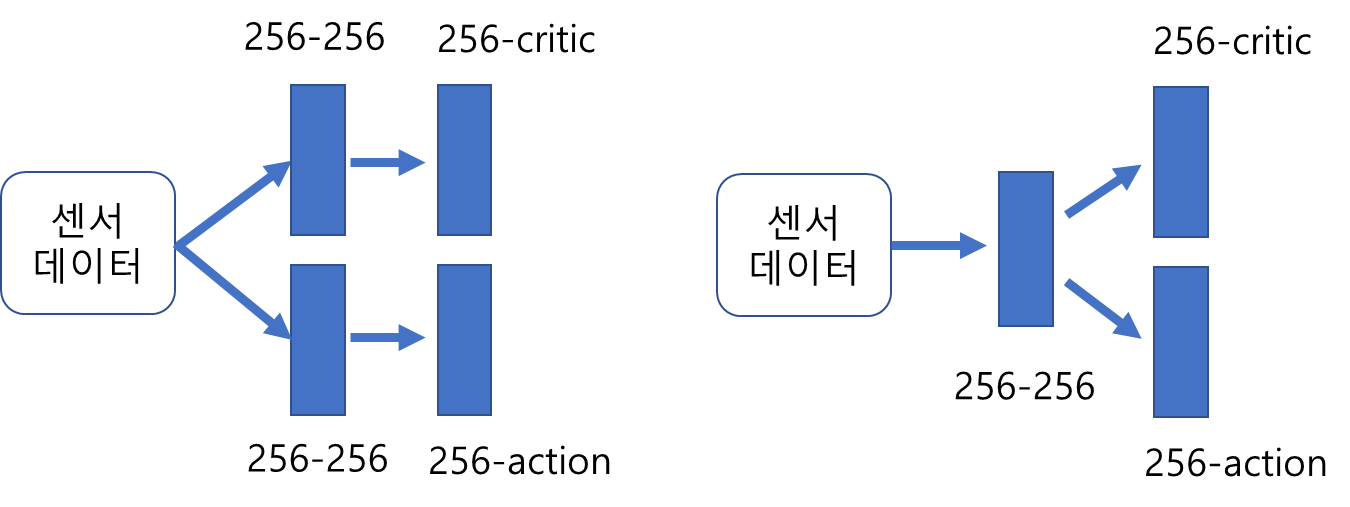
1) RL with 왜곡수준예측 모델 성능 고도화 필요

* 현재 왜곡수준예측이 잘 되지 못함
* RL 모델 부터 256-256 형태로 키울필요가 있을까?
* 왜곡수준예측 레이어를 어디서 브랜치 할 것인가?

\* 10월 24일 해볼 것으로,

Offline data + SAC 할 때 모델 사이즈를 256-256 으로 학습이 되는지 확인.

아래 두 모델을 각각 학습 시켜볼 수 있을지?



2) RL 학습 시 현재는 호버링에 집중.  
 오프라인 데이터도 다양한 목표지점에 대해 멀리 이동하는 데이터를 수집.

이후 SAC 학습 시에도 여러 목표지점에 대해서 이동하는 방식으로 확장

3) CalQL (오프라인 RL) 학습은 왜 안되었을까?   
 학습 후 평가시에 deterministic policy 설정은 했는가?

4) HER 학습도 왜 안되었을까?

5) Stable Baseline 3 는 왜 SAC 학습이 안될까?

6) RL & 왜국수준예측 시 LSTM 또는 self-attention 적용해보자

7) 현재는 왜곡이 적어도 정확히 목표지점에 위치하지 못하고

2m 정도를 왔다갔다하고 있음. reward 함수를 수정?