Learning to Navigate Unseen Environments: Back Translation with Environmental Dropout

- R2Rタスクにおいて見た事がない環境での性能向上
- 模倣学習 (IL) と強化学習 (RL) を混合してエージェントを学習
- environmental dropout を用いた半教師あり学習
  - ある特定のクラスのオブジェクト（例えば、椅子）を取り除く

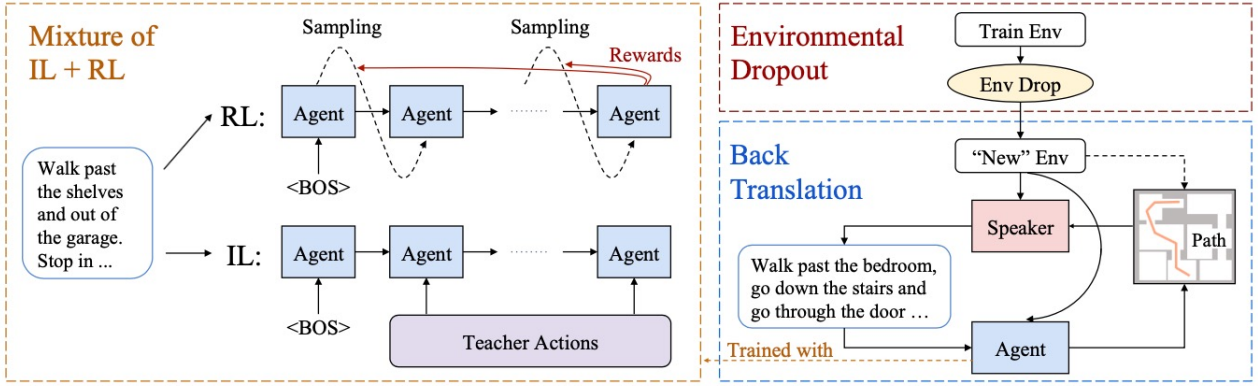

Figure 2: Left: IL+RL supervised learning (stage 1). Right: Semi-supervised learning with back translation and environmental dropout (stage 2).

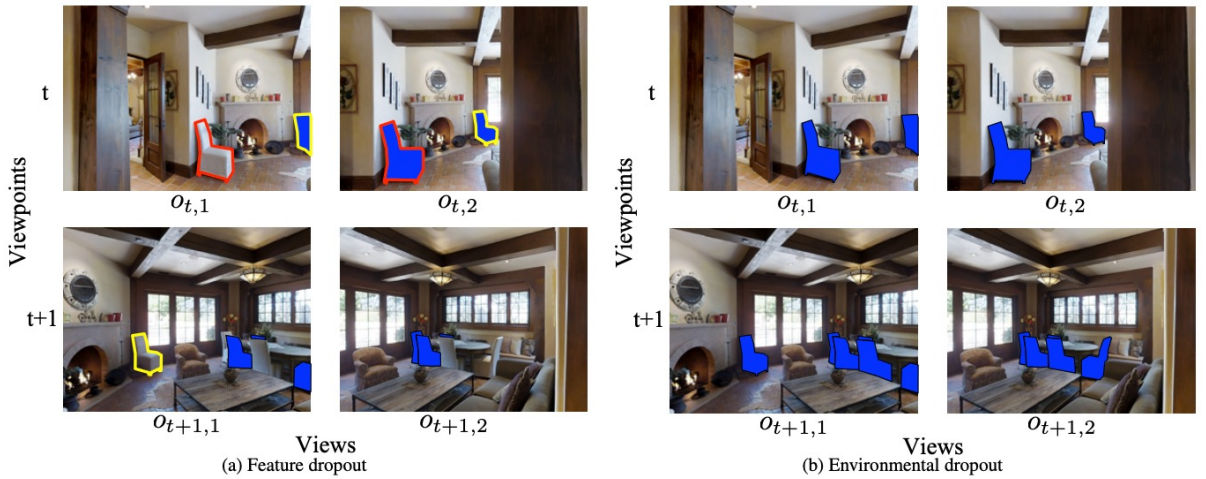

Figure 3: Comparison of the two dropout methods (based on an illustration on an RGB image).

Hao Tan, Licheng Yu, Mohit Bansal. Learning to Navigate Unseen Environments: Back Translation with Environmental Dropout. NAACL 2019.

Large-Scale Adversarial Training for Vision-and-Language Representation Learning

- 視覚＋言語の学習にAdversarial Training (AT)を導入
- Pre-training, finetuningの両方にATを採用
  - Pre-training：タスク非依存
  - Finetuning：タスク依存

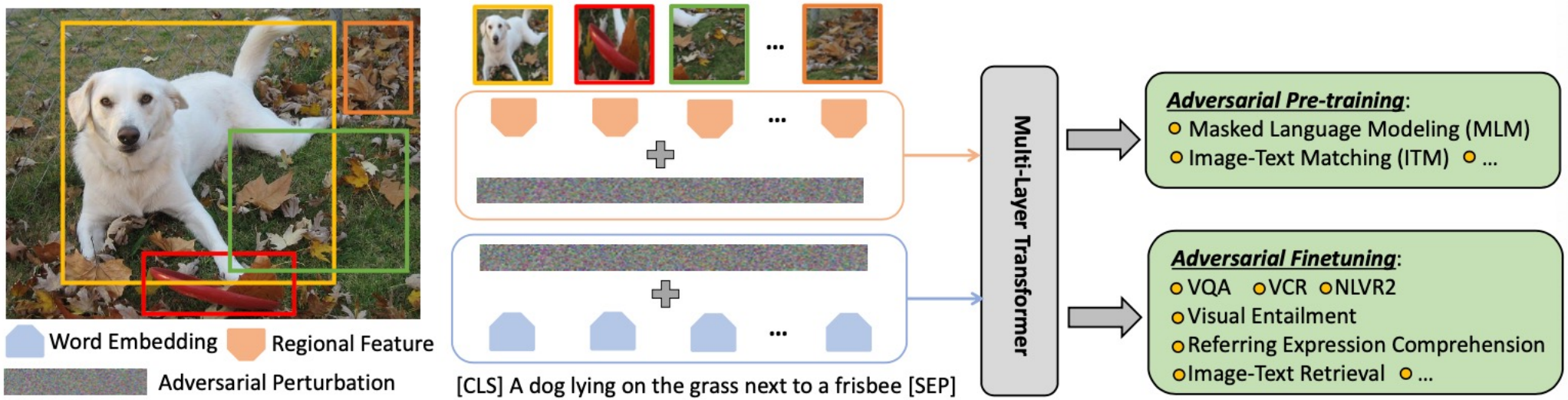

Figure 1: Overview of the proposed VILLA framework for vision-and-language representation learning.

Zhe Gan, Yen-Chun Chen, Linjie Li, Chen Zhu, Yu Cheng, Jingjing Liu.Large-Scale Adversarial Training for Vision-and-Language Representation Learning. NeurIPS 2020