α, β-CROWN 과제 4 보고서

2022430004 김민영

1. 선택한 모델과 데이터셋, 그리고 그 이유

- **모델:** Oval21 CNN (cifar_base_kw.onnx)
    - VNN-COMP 2021 공식 벤치마크에 포함된 경량 CNN 모델로, 입력 차원이 작고 파라미터 수가 3,100~6,200개 수준으로 가벼움.
    - 이미지 분류용으로 설계되어 있으며, 다양한 신경망 검증 도구의 성능 비교에 널리 사용됨.
- **데이터셋 및 사양:**

  Oval21 벤치마크의 CIFAR-10 기반 이미지와 VNNLIB 사양 파일(cifar_base_kw-img1697-eps0.0014379084967320263.vnnlib)
    - 해당 VNNLIB 파일은 특정 입력 이미지(1697번)와 섭동 크기(ε=0.00144...)에 대한 로컬 강건성 검증 조건을 정의함.
- **선정 이유:**
    - α, β-CROWN 공식 models 디렉토리에 포함되어 있지 않은 외부 모델이며, 국제 대회에서 표준적으로 사용된 검증 사례라 과제 요구조건에 부합함.
    - 모델이 경량이라 실행이 빠르고, 다양한 섭동 조건 실험이 가능함.

2. 호환성을 위해 필요한 수정 또는 전처리 단계

- **모델 및 사양 파일 준비:**

- VNN-COMP 2021 벤치마크 저장소에서 ONNX 모델(cifar_base_kw.onnx)과 VNNLIB 파일(cifar_base_kw-img1697-eps0.0014379084967320263.vnnlib)을 다운로드했음.
- 별도의 변환이나 추가 전처리는 필요하지 않으며, α, β-CROWN에서 바로 지원하는 포맷.

- **환경 구축:**
  - Miniconda/Anaconda를 설치한 후, α, β-CROWN 저장소의 environment_win.yml 파일로 conda 환경을 생성함.
  - Python, PyTorch, auto_LiRPA 등 필수 패키지가 자동으로 설치됨.
- **폴더 구조 및 경로:**
  - 모델, 사양 파일, 설정 파일의 경로를 YAML 설정에 맞게 지정함.

3. α, β-CROWN 실행 및 모델 검증 단계

3.1 환경 구축 및 파일 준비

```
git clone https://github.com/Verified-Intelligence/alpha-beta-CROWN.git
cd alpha-beta-CROWN
conda env create -f complete_verifier/environment.yml --name alpha-beta-crown
conda activate alpha-beta-crown

git clone https://github.com/stanleybak/vnncomp2021.git
```

- 모델: vnncomp2021/benchmarks/oval21/nets/cifar_base_kw.onnx
- 사양: vnncomp2021/benchmarks/oval21/vnnlib/cifar_base_kw-img1697-eps0.0014379084967320263.vnnlib.

3.2 YAML 설정 파일 예시

```
model:
  onnx_path: ../../vnncomp2021/benchmarks/oval21/nets/cifar_base_kw.onnx
specification:
  vnnlib_path: ../../vnncomp2021/benchmarks/oval21/vnnlib/cifar_base_kw-img1697-
eps0.0014379084967320263.vnnlib
general:
  root_path: ../../vnncomp2021/benchmarks/oval21/
```

3.3 실행 명령어

```
python abcrown.py --config exp_configs/vnncomp21/oval21_base_01.yaml
```

3.4 검증 결과 예시 및 해석

- 콘솔 출력 예시:

```
PGD attack succeeded!
Checking and Saving Counterexample in check_and_save_cex
verified_status unsafe-pgd
verified_success True
Result: sat
Time: 4.431538105010986
```

  - sat: 반례 존재(강건성 미보장), verified일 경우 강건성 보장.

4. 주요 코드/명령어 요약

단계	명령어/코드 예시
환경 구축	conda env create -f complete_verifier/environment.yml --name alpha-beta-crown
환경 활성화	conda activate alpha-beta-crown
모델/사양 준비	git clone https://github.com/stanleybak/vnncomp2021.git
검증 실행	python abcrown.py --config exp_configs/your_config.yaml

5. 재현성 및 환경 정보

- **환경:**
  - OS: Windows/Linux/Mac
  - Python: 3.8~3.10
  - Conda, PyTorch, auto_LiRPA 등
- **환경 파일:**
  - environment.yml (α, β-CROWN 저장소 내)
- **폴더 구조/README:**
  - 각 파일 위치, 실행법, 주요 선택 이유 등 간단 설명 포함

6. 결론

Oval21 CNN과 공식 VNNLIB 사양 파일을 활용하여 α, β-CROWN 신경망 검증을 수행하였으며, 별도의 변환 없이 공식 벤치마크 파일을 그대로 사용해 빠르고 정확한 실험이 가능함을 확인함.

파일/폴더	내용/용도
assignment4보고서.pdf	본보고서(1~2페이지)
alpha-beta-CROWN/complete_verifier/exp_configs/vnncomp21/oval21_base_01.yaml	YAML 설정 파일
alpha-beta-CROWN/vnncomp2021/benchmarks/oval21/nets/cifar_base_kw.onnx	ONNX 모델 파일
alpha-beta-CROWN/vnncomp2021/benchmarks/oval21/vnnlib/cifar_base_kw-img1697-eps0.0014379084967320263.vnnlib	VNNLIB 사양 파일
environment_win.yaml	Conda 환경 파일