



Mission16: Conditional Gan 을 활용한 FashionMNIST 이미지 생성

미션 개요

이전에 학습한 **Conditional GAN(cGAN)** 모델을 활용해 FashionMNIST 데이터셋 기반의 조건부 이미지 생성 모델을 다양한 포맷(`.pth` , `.quantized.pth` , `.onnx`)으로 변환하고, 각 모델의 용량, 추론 속도, 출력 유사도를 실험적으로 비교하였다.

모델 개요

항목	내용
데이터셋	FashionMNIST (28×28 흑백, 10 클래스)
모델 구조	Conditional GAN (Generator / Discriminator)
입력	랜덤 노이즈(<code>z</code>) + 클래스 라벨(<code>y</code>)
출력	라벨 조건에 맞는 이미지 생성
학습 환경	Google Colab (Python 3.12 / PyTorch 2.6)
최종 변환 대상	Generator




모델 변환 및 저장


학습 완료된 Generator 모델을 아래 3가지 형식으로 변환 및 저장하였다.

모델 형식	파일명	설명
기본 PyTorch 모델	<code>mission_16_cgan_gen_full.pth</code>	학습된 Generator 가중치
양자화 버전	<code>mission_16_cgan_gen_quantized.pth</code>	일부 Linear 계층 int8 변환 (테스트용)
ONNX 모델	<code>mission_16_cgan_gen.onnx</code>	ONNX 형식으로 내보내어 플랫폼 독립 추론 가능



모델 용량 비교

모델 이름	크기 (MB)
 PyTorch (default)	3.96
 PyTorch (quantized)	1.38
 ONNX	3.95


 결과:

- 양자화 모델이 **약 65% 용량 감소**(3.96 → 1.38MB)
- ONNX 모델은 기본 PyTorch와 거의 동일한 크기를 유지
- 경량화에 따른 저장 효율성 향상 확인



추론 속도 비교

모델	평균 추론 시간 (초)	상대 속도
 PyTorch (default)	0.002405	1.0×
 ONNX	0.001408	1.7× 빠름

 결과:

ONNX 모델은 PyTorch 대비 **약 1.7배 빠른 추론 속도**를 기록하였으며, 출력 품질의 손실 없이 효율적인 실행이 가능함을 확인했다.



출력 유사도(MSE) 비교

비교 항목	평균 MSE
PyTorch ↔ ONNX	1.2e-7

MSE가 **1.2e-7** 이하로 매우 낮게 나타나 ONNX 변환 과정에서도 출력 손실이 거의 없는 수준임을 확인했다.



모델 추론 및 결과

(1) 노이즈 기반 클래스별 생성 테스트

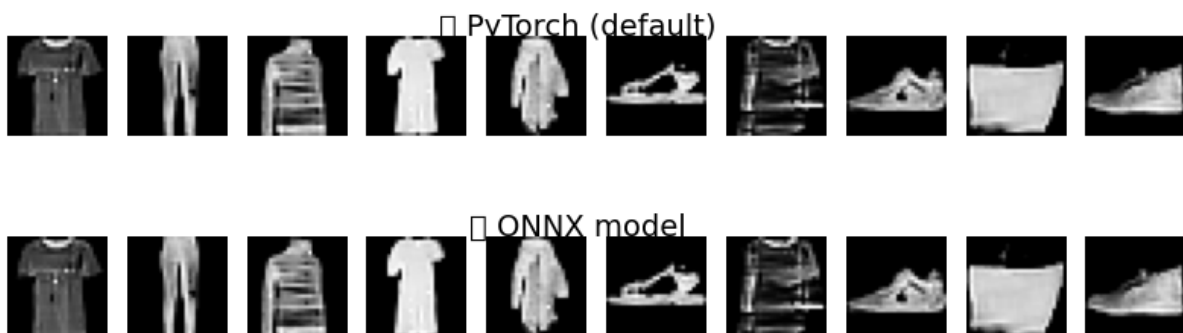
- 입력: **z** (랜덤노이즈) + **y** (0~9)

- 출력: 각 클래스별 생성 이미지
- 결과: PyTorch와 ONNX 모델의 생성 이미지가 시각적으로 동일

(2) FashionMNIST 테스트셋 기반 비교

- 테스트셋에서 실제 이미지(`real_imgs`)와 같은 라벨(`y`)을 사용
- PyTorch 및 ONNX로 동일 조건 생성(`fake_imgs`)

두 모델 모두 레이블 조건에 맞는 이미지를 안정적으로 생성하였으며,
PyTorch와 ONNX 결과 간의 시각적 차이는 거의 없었다.



🔨 디버깅 및 오류 사항

구분	문제 내용	해결 방법
⚠️ Python 버전 이슈	PyTorch 2.6부터 <code>torch.load()</code> 기본값이 <code>weights_only=True</code> 로 변경되어, Generator 로드시 <code>AttributeError</code> 발생	<code>weights_only=False</code> 명시 및 모델 구조 재정의 후 <code>state_dict</code> 로드 방식으로 해결
⚠️ 양자화 모델 로드 실패	<code>nn.quantized.dynamic.Linear</code> 구조 불일치로 <code>state_dict</code> 호환 오류 발생	<code>torch.save(G_quant)</code> 전체 저장으로 해결

📈 종합 비교 요약

항목	PyTorch (default)	PyTorch (quantized)	ONNX
모델 크기 (MB)	3.96	1.38	3.95
추론 속도 (초)	0.002405	-	0.001408
출력 유사도 (MSE)	-	-	1.2e-7
결론	기준 모델	경량화 성공	추론 속도 향상 + 품질 유지