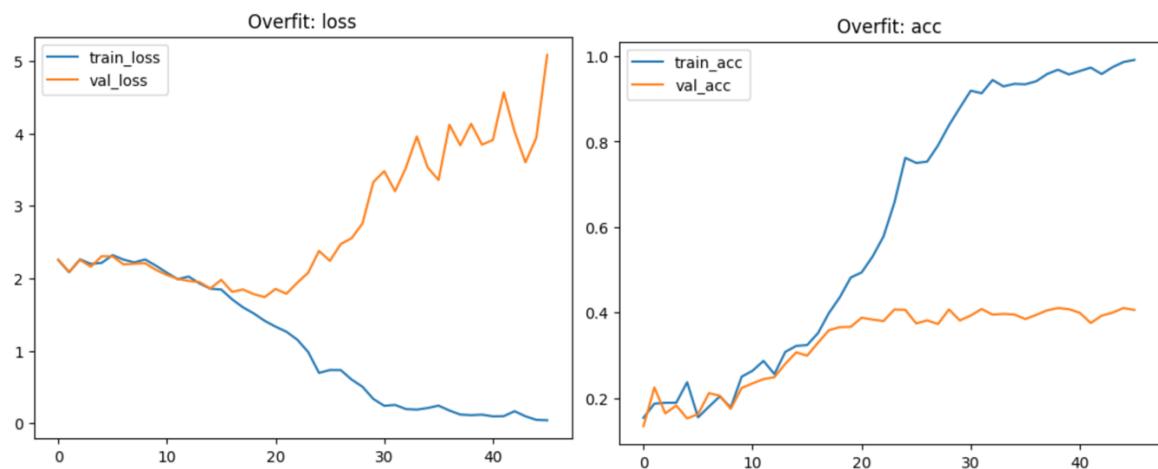


## [Overfit vs Regularization 분석 리포트]

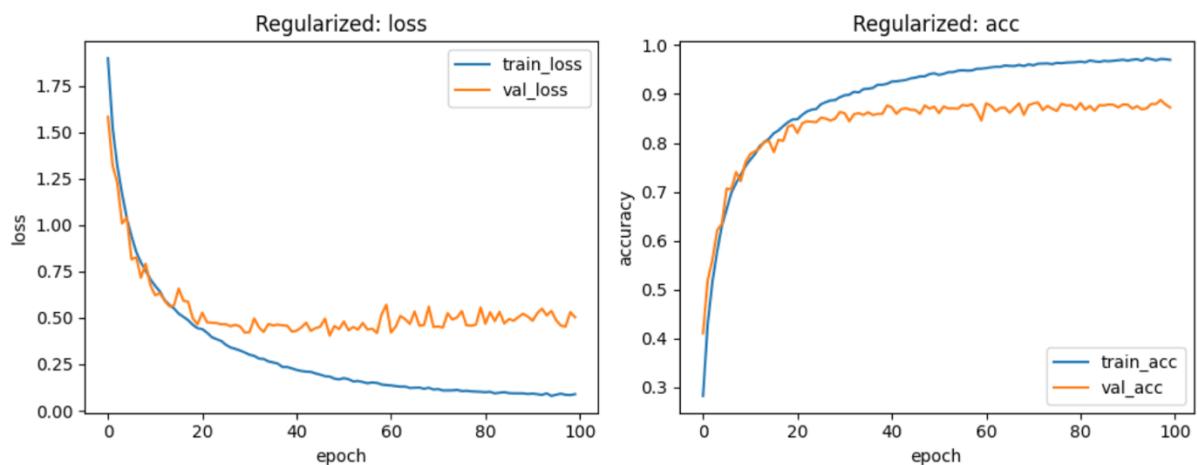
### 1. Overfitted model

- CIFAR-10 train 데이터 중, 일부 데이터(1,000장)로 학습
- 전체 train에서 10%를 분리하여 만든 validation 데이터로 검증
- 데이터 증강, 규제 기법 미사용
- CIFAR-10 입력에 맞춘 AlexNet style CNN
- Conv – ReLU – MaxPool + FC 분류기
- 학습 설정
  - optimizer : SGD + Momentum
  - Learning rate : 0.05
  - Momentum : 0.9
  - Batch size : 128
  - 최대 Epoch : 200
  - Loss : Cross Entropy Loss
  - Weight initialization : Kaiming
  - early stopping : train acc가 0.99 이상이면 종료
- 작은 데이터로 훈련을 진행하여, 과적합 됨
- train acc와 val acc의 차이가 극단적임
- train acc의 성능은 매우 높고, val acc의 성능은 상대적으로 아주 낮음
- 일반화 성능이 아주 낮을 것으로 예상됨



## 2. Regularized model

- CIFAR-10 데이터 중 전체 훈련 데이터로 학습
- data augmentation 기법 2가지 적용 (학습 데이터)
  - RandomCrop
  - RandomHorizontalFlip
- regularization 기법 2가지 적용
  - weight decay : 5e-4
  - dropout : 0.5
- 학습 설정:
  - Learning rate : 0.01
  - Momentum : 0.9
  - Batch Size : 128
  - 최대 epoch : 100
  - Loss : Cross Entropy Loss
  - Weight initialization : Kaiming
- val accuracy 최고 갱신 시, best checkpoint에 저장
- 데이터 증강과 규제 기법을 적용하여, 일반화 성능 강화



### 3. Overfitted model vs Regularized model

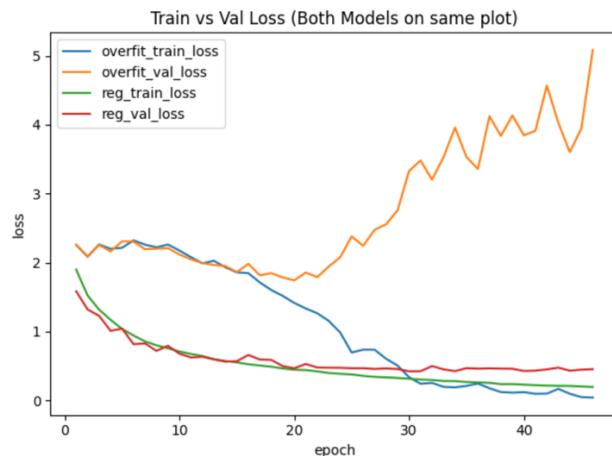
<Quantitative comparison>

(final training, validation, test accuracy for both models)

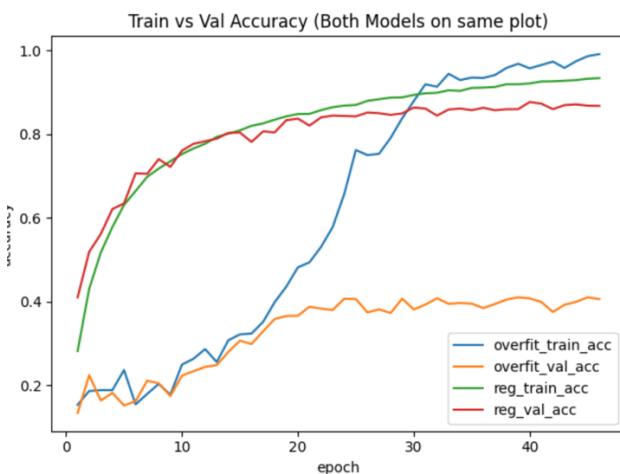
== Final Accuracy Summary (SAME val/test; Overfit-train is on its 1000-sample subset) ==				
Model	Train Acc	Val Acc	Test Acc	Train-Val Gap
Overfit	0.9930	0.4062	0.4026	0.5868
Regularized	0.9849	0.8880	0.8822	0.0969

- overfit model의 경우 train acc는 매우 높지만, val/test acc는 상대적으로 아주 낮으며, train-val gap이 큰 것을 확인할 수 있음
- regularized model의 경우 val/test acc가 overfit model에 비해 크게 증가했으면 train-val/test gap이 아주 작아진 것을 확인할 수 있음

(training vs validation loss)



(training vs validation accuracy)



- loss curve에서 비교해보면 :
  - overfit model의 경우 train loss는 계속해서 감소하는 반면에, val loss는 초반에는 감소하다가 일정 시점 이후에 계속 증가하는 것을 확인 : 이는 훈련 데이터에 과적합 되었음을 시사
  - 반면 regularized model은 대체적으로 train loss와 val loss가 함께 꾸준히 감소하면서, 일반화 성능이 강화된 것을 확인 가능
- acc curve에서 비교해보면 :
  - overfit model의 경우 train acc는 계속해서 증가하는 반면에, val acc는 초반에 증가하다가 더 이상 증가하지 않고 횡보 혹은 감소하는 것을 확인
  - 반면 regularized model은 대체적으로 train acc와 val acc가 함께 꾸준히 증가 – 일반화 성능이 강화되었음을 확인 가능함

### <Qualitative comparison>

overfitted model :

(correctly classified)

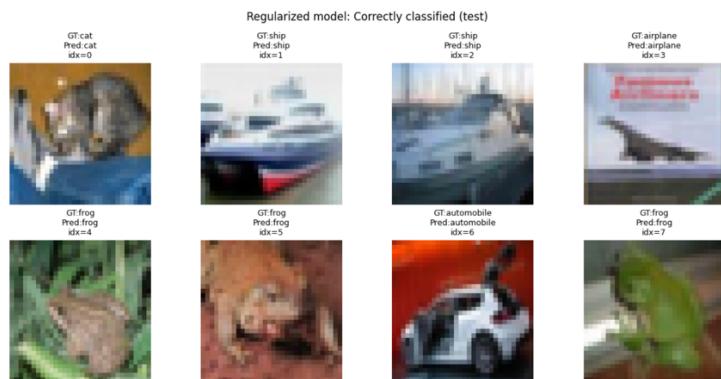


(miss classified)



regularized model:

(correctly classified)



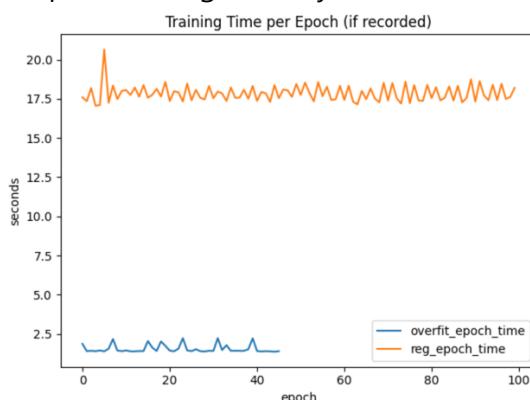
(miss classified)



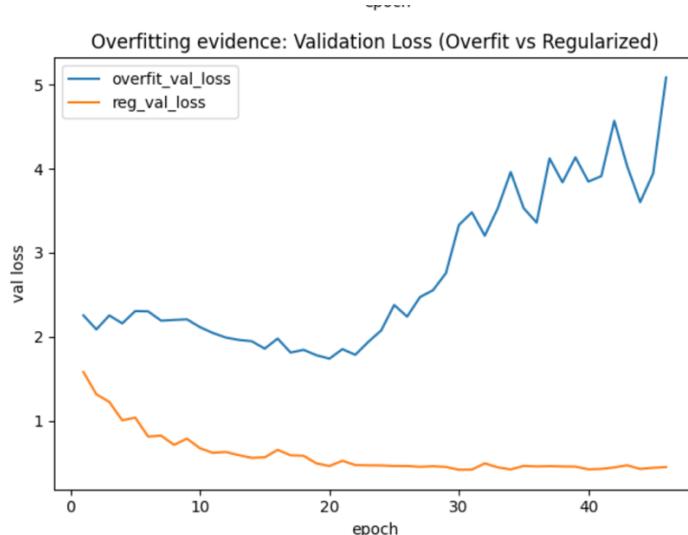
discussion :

- Training Time

- 데이터 증강 (crop, horizontal flip)이 적용된 regularized model의 경우, 입력 변환 비용이 추가되어 epoch 당 학습 시간이 데이터 증강이 적용되지 않은 overfitted model에 비해 증가함
- dropout과 weight decay 자체는 연산량 증가가 크지 않은 것으로 확인됨



- Overfitting Behavior
  - overfitted model :
    - ◆ train 데이터(1000장)에서 train 정확도가 매우 빠르게 성장하여 거의 100%에 근접함
    - ◆ 하지만 val, test 정확도는 그에 비례해 증가하지 않거나 정체됨
    - ◆ train loss는 계속 감소하는 반면, val loss는 일정 시점 이후 증가
    - ◆ train 데이터에 대해서만 성능이 뛰어나고, 검증 및 테스트 데이터에 대한 성능은 매우 떨어지며 train-val acc gap이 크게 나타남
    - ◆ 전형적인 과적합이 발생하였다고 볼 수 있음
  - regularized model :
    - ◆ 데이터 증강과 규제를 적용한 모델은 훈련 정확도가 상대적으로 낮게 유지됨
    - ◆ 하지만 val, test 정확도가 크게 개선되었고, train-val acc gap이 overfitted model에 비해 크게 감소함
    - ◆ overfitted model에 비해 val, test 데이터에서 보다 좋은 성능을 보이며, 일반화 성능이 강화된 것을 확인 가능



- Training Dynamics
  - $\Delta\text{val\_acc}$  검증 정확도의 에폭별 변화량으로, 초반 큰 양수 스파이크는 빠른 학습, 후반 0 근처 진동은 수렴/정체를 의미함
  - Overfit 모델 :  $\Delta\text{val\_acc}$ 의 진동(특히 음수 하락)이 더 커서 검증 성능이 불안정한 경향을 보임
  - regularized 모델 : 변화 폭이 작아 학습이 더 안정적으로 진행되는 경향을 보임



#### - Final Generalization Performance

- overfitted model :
  - ◆ train acc는 매우 높았지만, val/test acc는 낮은 성능을 보여주며 train-val/test acc gap이 매우 크다는 것을 확인함
  - ◆ 최종적으로 일반화 성능이 매우 떨어짐을 확인함
- regularized model :
  - ◆ train acc는 overfitted model에 비해 소폭 하락했지만, val/test acc가 크게 상승하여 train-val/test acc gap이 많이 줄어든 것을 확인함
  - ◆ 최종적으로 일반화 성능이 매우 강화되었음을 확인함
  - ◆ 데이터 증강이 입력 분포를 다양화하고, 규제 기법이 모델이 특정 샘플에 과도하게 적합하는 것을 억제한 것에 대한 효과로 보여짐

