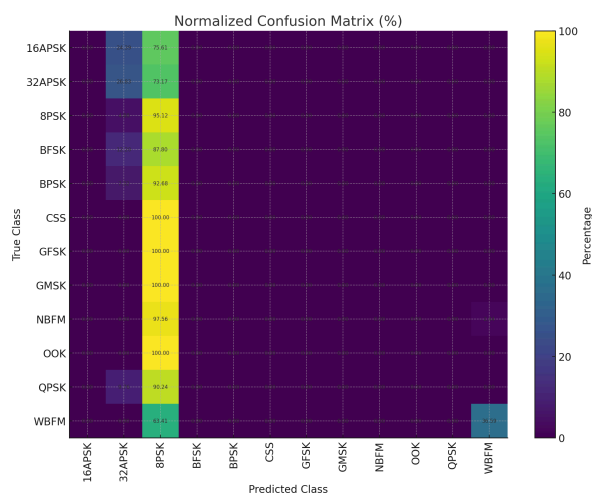


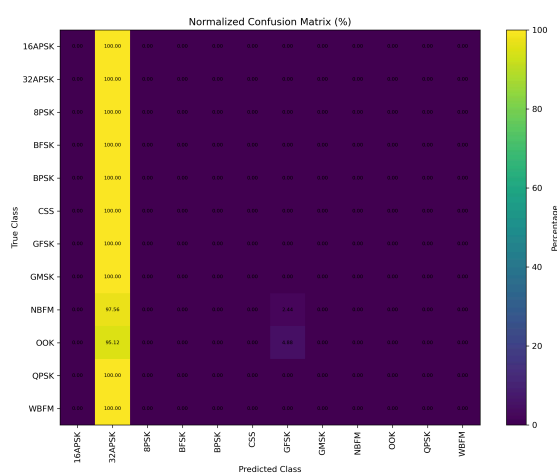
# CLIP 평가 결과

- Model: clip-vit-base-patch32
- GPU: A5000 (24GB memory)
- Dataset
  - Training
    - image\_clean\_signal\_training0704
    - image\_clean\_signal\_training0709
  - Inference: image\_clean\_signal\_inference0707

## 1. Vanilla



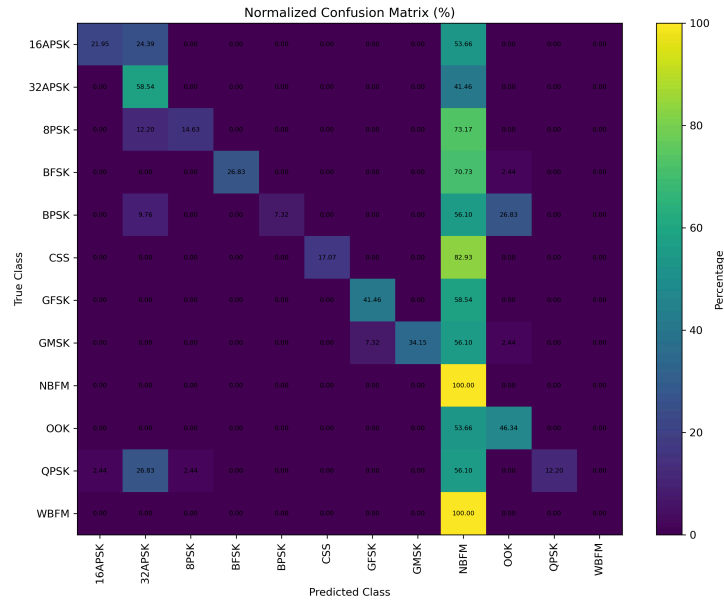
Vanilla 모델 결과 (Confusion Matrix)



Vanilla 모델 결과 (Confusion Matrix)

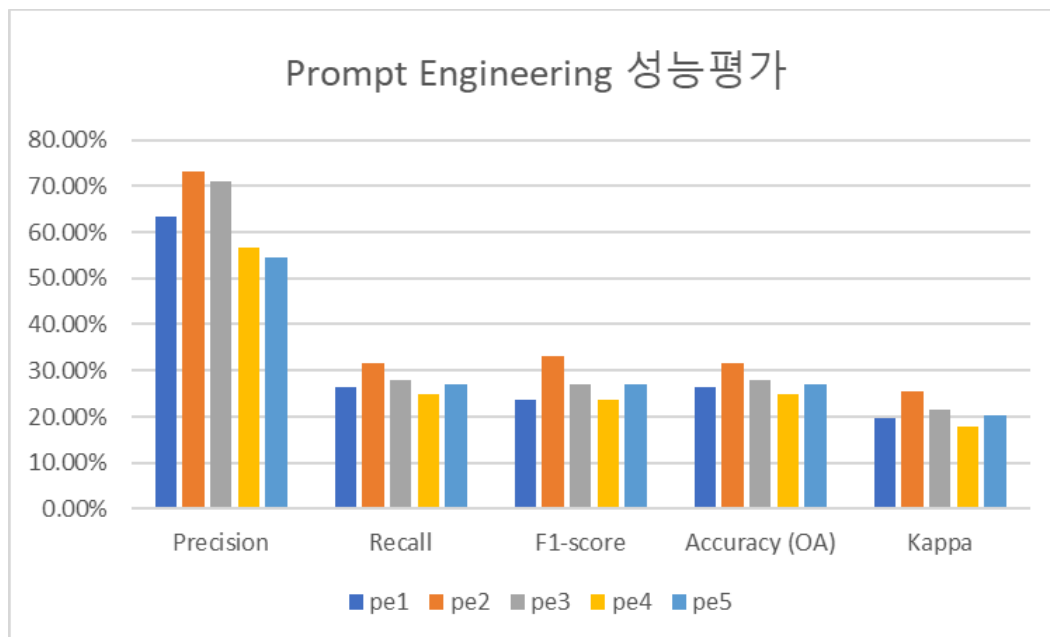
Vanilla CLIP 사용 불가능.

## 2. PE



최고 성능을 보인 Prompt 사용한 모델의 Confusion Matrix

```
prompt_prefix = (
    "These four signal plots help determine modulation. Types include phase-based"
    "frequency-based (FM, FSK, GFSK), APSK, and CSS. This signal most likely uses"
)
```

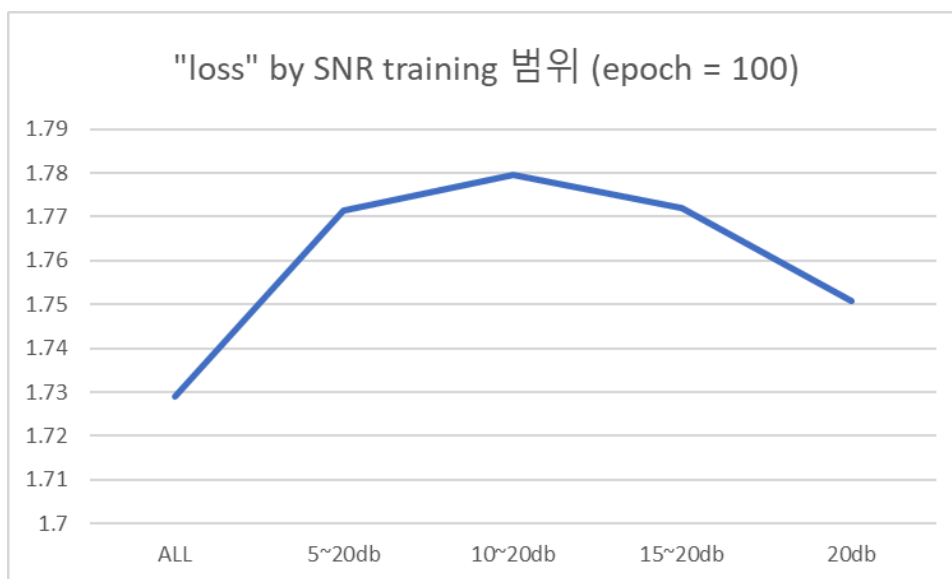


(inference에는 Full fine tuning (Freeze-Backbone) 한 모델 사용.)

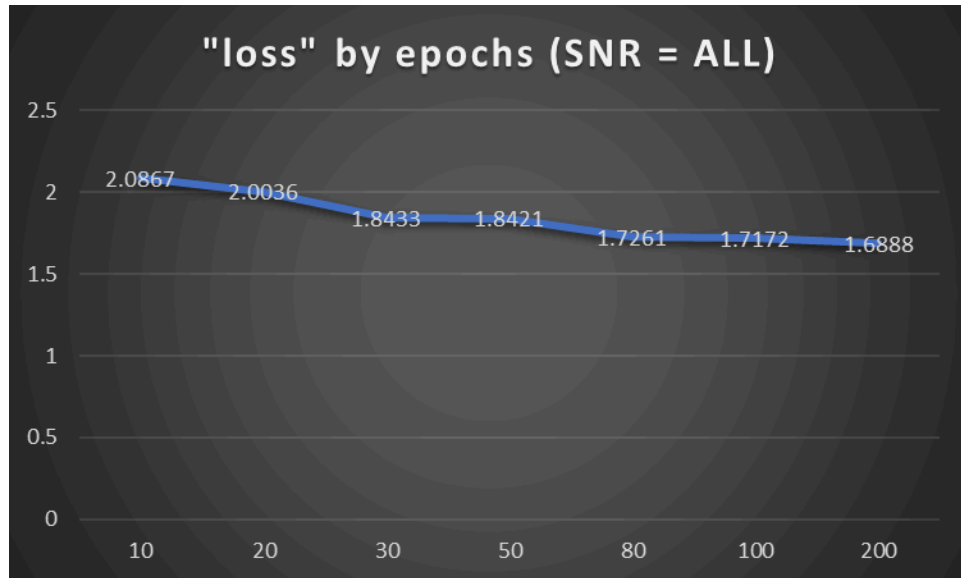
Prompt 를 바꾼다고 해도 성능 상 큰 차이는 없으나 fine-tuning으로 이미 결과가 잘 나오는 모델에 대해서 좋은 prompt를 사용하면 약간의 성능 향상을 기대할 수 있을 것으로 보임.

다만, CLIP 특성 상 긴 prompt를 사용할 수 없었고, 명확한 prompt 생성 기준이 없었기 때문에 향후 모델 테스트 때는 어느 정도 기준을 확보해서 할 예정.

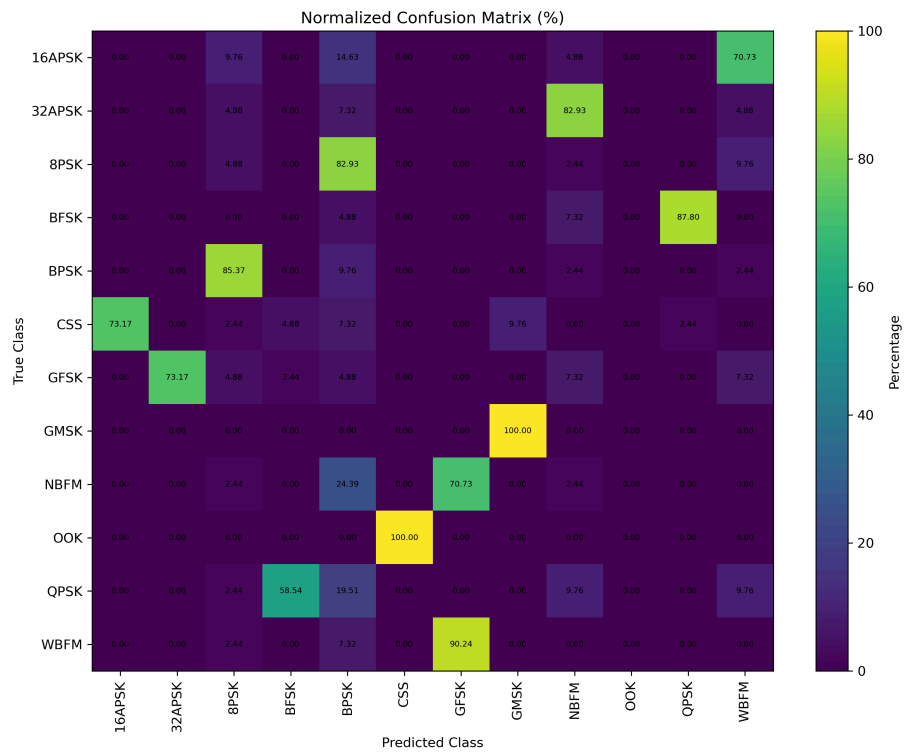
### 3. LoRA



SNR training 범위에 따른 loss 값 (낮을 수 록 좋음)



epoch 수에 따른 loss 값 (낮을 수 록 좋음)

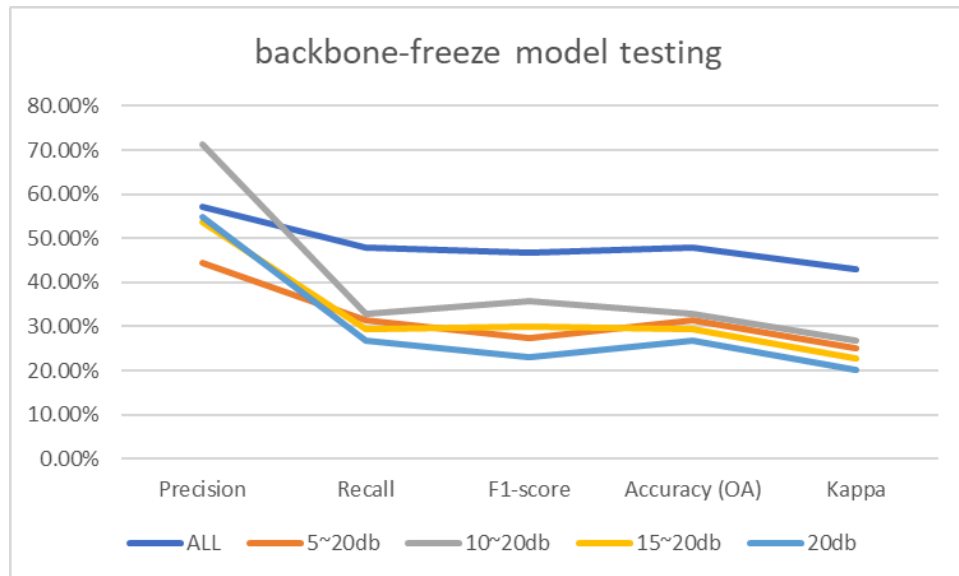


loss 값이 작게 나온 모델의 confusion matrix

LoRA training으로 정상 사용 가능한 모델 없음.

앞선 결과들을 통해 Vanilla, LoRA training 으로는 제대로된 결과가 나오지 않음을 확인. (CLIP 모델 크기 자체가 너무 작아서 training이 제대로 되지 않는다고 예상.)

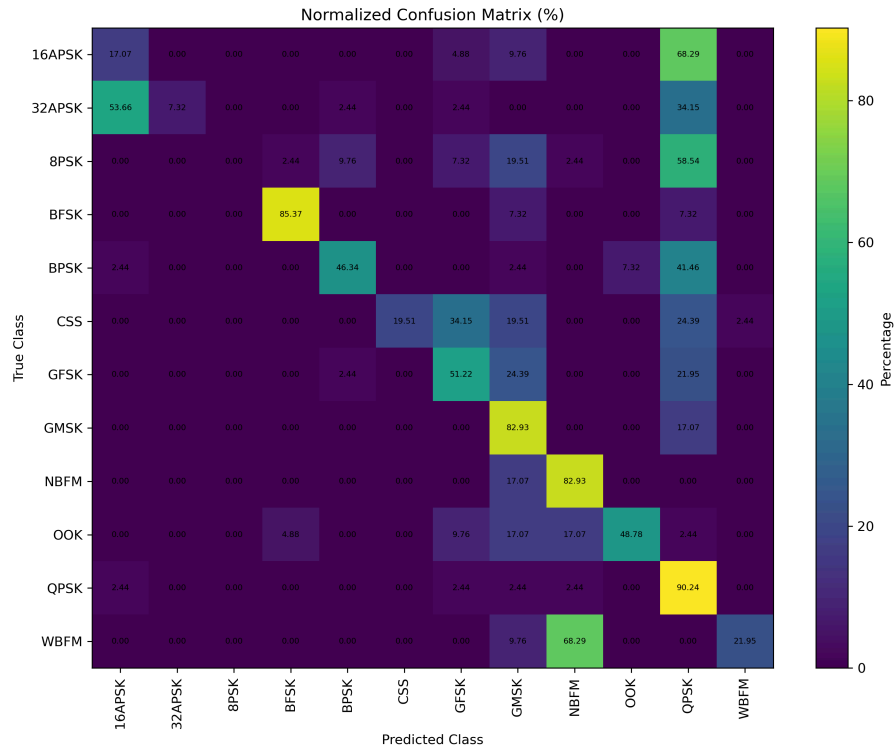
#### 4. Full fine-tuning (backbone freeze)



SNR training 별 평가 지표 그래프



"SNR\_ALL" 모델의 epoch 에 따른 loss 변화



최고 성능을 보인 모델 결과 Confusion Matrix (아래 평가 지표 표 참고)

Precision	Recall	F1-score	Accuracy (OA)	Kappa
60.41%	46.14%	42.55%	46.14%	0.4124

대부분의 지표에서 모든 SNR 값의 이미지를 training 한 모델이 가장 성능이 좋았으나, 사용할 수 있을 정도의 성능 향상은 어려웠음.

**결론: "CLIP 은 AMC 용 VLM으로 사용 부적합."**